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

Giovanni Gallavotti
Fisica and I.N.F.N. Roma 1
October 19, 2006

Abstract: *Boltzmann's discrete approach to the theory of ensembles and the ergodic hypothesis and modern applications.*

I. BOLTZMANN'S AND HELMOLTZ' APPROACH TO ENSEMBLES.

Among the many BOLTZMANN's contributions to statistical mechanics is the theory of statistical ensembles, [1], perhaps the conceptually deepest.

At the beginning of his work (1866) kinetic theory was quite clearly founded: although the pioneer works of BERNOULLI, HERAPATH, WATERSTONE had not received the deserved attention (or no attention at all, [2]), AVOGADRO's law, the establishment of the first law of Thermodynamics (MEYER, JOULE), and the identification of average kinetic energy with absolute temperature (KRÖNIG, CLAUSIUS) proposed the challenge of founding on a mechanical basis the newly discovered second law and entropy. The young BOLTZMANN immediately started on this endevour which was to occupy most of his attention in the rest of his studies.

The ergodic hypothesis emerged almost immediately as a tool to insure existence of the average values of functions on the phase space of the systems considered. At first in a rather naive form "*... 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*", [3, p.20]. Later it evolved, [4, p.237], into an assumption for the internal motion of atoms in single molecules of a gas: "*it is clear that the various gas molecules will go through all possible states of motion*" which, however, could possibly be understood from the context to be different from the ergodic hypothesis, see [2], because the molecules undergo from time to time collisions, see [5, p.96],[6, p.xxxvii].

Considering a collection of molecules alike to a large molecule, [7, p.284], the same assumption became what is often referred as the ergodicity property of the entire gas.

Certainly the ergodic hypothesis could not possibly say that every point of the energy surface in phase space visits in due time (the *recurrence time*) every other, [2, 505]. But this statement was attributed to BOLTZMANN and criticized over and over again (even by Physicists, including in the influential book [8], although enlightened mathematicians could see better, [9, 385]).

However for BOLTZMANN phase space was discrete and points in phase space were *cells* Δ with finite size, $h > 0$, as made clear by: *Die Zahl der lebendigen Kräfte ist eine diskrete*, [10, 167]. A viewpoint that permitted to imagine that time evolution was a permutation of cells: ergodicity meant therefore that the permutation was a *one cycle permutation*. On this assumption BOLTZMANN derived, in a series of papers the canonical and microcanonical distributions and the related second law, [4, 7, 11]. And in [10] he explicitly used this discrete conception of phase space to find an alternative "combinatorial" derivation of the canonical distribution.

In [12, Sect.III] appears a simple example of "mechanical analogy" Thermodynamics: this concept was later explicitly introduced by HELMHOLTZ for general systems admitting only periodic motions (called *monocyclic*), [13, 14].

This meant existence of quantities associated with the phase space of a mechanical equation of motion (typically defined as time averages over the solutions of the equations of motion), which could be given thermodynamic names like equilibrium state, pressure p, volume V, energy U, temperature T, entropy S *and be related* by the thermodynamic relations that are expected to hold between the physical quantities bearing the same name, namely

$$dS = \frac{dU \,+\, p\, dV}{T} \tag{1}$$

The proposal provided a new perspective and generated the new guiding idea that thermodynamic relations would hold in *every mechanical system*, from the small and simple to the large and complex: in the first cases the relations would be trivial identities of no or little interest, just *thermodynamic analogies*, but in the large systems they would become nontrivial and interesting being relations of general validity. In other words Thermodynamic would be a kind of symmetry property of Hamiltonian Mechanics.

HELMHOLTZ' papers do not mention a relation with the earlier [12]: but the relation was clear to BOLTZMANN who immediately published the key work [15] where the idea was taken up. It naturally allowed him to fit into a single general theory the ergodic hypothesis, the theory of ensembles and their equivalence. The realization that Thermodynamics was just what a general property of mechanical systems became when considered for large assemblies of particles reaches its final and definitive form in this work.

In the simplest case consider a one dimensional system confined by a potential $\varphi_V(q)$, depending on a generic parameter V, the average kinetic energy $T = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau K(t))dt$ is identified as proportional to temperature and energy is identified with the total energy U. Then if pressure p is defined as the time average $\lim_{\tau\to\infty} -\frac{1}{\tau}\int_0^\tau \partial_V\varphi_V(q(t))dt$, the quantities T, p become functions $p(U,V)$,$T(U,V)$ of the parameters U, V determining the state of the system, identified with the peri-

odic trajectory of energy U between extremes $q_\pm(U,V)$, and Eq.(1) should hold.

Indeed it is $dt \equiv \frac{dq}{|\dot q|} = \frac{dq}{\sqrt{2(U-\varphi_V(q))/m}}$ and the period of the oscillations is given by $\tau_0 = \tau_0(U,V) = 2\int_{q-}^{q+} \frac{dq}{\sqrt{2(U-\varphi_V(q))/m}}$, hence ([15, p.127] and [16, Ch.I]),

$$\begin{aligned} T &= \frac{2}{\tau_0}\int_{q-(U,V)}^{q+(U,V)} \frac{m}{2}\sqrt{\frac{2}{m}(U-\varphi_V(q))}dq, \\ p &= \frac{2}{\tau_0}\int_{q-(U,V)}^{q+(U,V)} \frac{\partial_V\varphi_V(q)}{\sqrt{\frac{m}{2}(U-\varphi_V(q))}}dq \end{aligned} \tag{2}$$

and it is immediate to check, as in [12], that Eq.(1) is fulfilled by setting

$$S(U,V) = 2\int_{q-(U,V)}^{q+(U,V)} \sqrt{2m(U-\varphi_V(q))}\,dq \tag{3}$$

Assuming the ergodic hypothesis, the motion of an assembly of particles could be regarded as performing a periodic motion filling the energy surface; the potential confining the particles in their container could be considered as a function of the volume V and all observables F would admit a time average $\langle F\rangle$ equal to their average over the constant energy surface, *i.e.* our "microcanonical average".

Setting $p(U,V) = -\langle \partial_V\varphi_V\rangle$, the quantity $p$can be seen to admit the interpretation of pressure (defined also as average force per unit container surface), [16, App.9.3] so that Eq.(1) was not just showing that the ergodic hyptothesis led to a "model of Thermodynamics" but it gave the "actual Thermodynamics" of the system, provided the average kinetic energy was identified as proportional to the temperature.

In the same paper Boltzmann extends Helmholtz' notion of thermodynamic analogy by proposing to identify general families of thermodynamic equilibrium states with families $\mathcal{E}$ of *invariant probability distributions* on phase space, depending on two parameters α,β (in the case of simple systems). Then a family $\mathcal{E}$ would be called an *orthode* if the average kinetic energy $T = T(\alpha,\beta)$, the average energy $U = U(\alpha,\beta)$, the average force per unit surface $p = p(\alpha,\beta)$ (the average of $-\partial_V\varphi_V$) and the volume $V = V(\alpha,\beta)$ had the property that varying α,β by $d\alpha, d\beta$ the differential form in the *r.h.s* of Eq.(1) would turn out to be exact (*i.e.* the differential of some function $S = S(\alpha,\beta)$ of the parameters α,β).

And Boltzmann proves that the collection of the microcanonical distributions $\mu_{U,V}$ (parameterized by energy and volume, $(\alpha,\beta) = (U,V)$) is an example of orthode (deducing this property without using the ergodic hypothesis, which would have equally implied the result): which he named *ergode*. He also proves orthdicity of the collection of the canonical distributions $\mu_{T,V}$ (parameterized by temperature and volume, $(\alpha,\beta) = (\frac{1}{T},V)$), naming it *holode* (deriving this property without using the ergodic hypothesis).

Since orthodicity of the canonical and microcanonical ensembles could be derived without having recourse to dynamical assumptions a new issue arose: could it be that for the same system one could define several Thermodynamics (*e.g.* different equations of state)? this was not even raised as a problem because in the same paper Boltzmann proves the equivalence of the Thermodynamics defined by the microcanonical and canonical ensembles.

II. LIOUVILLE THEOREM AND PHASE SPACE CONTRACTION

Given the striking success of Boltzmann's program of deriving equilibrium Thermodynamics several attempts have been made to extend the method at least to stationary nonequilibrium phenomena.

The first difficulty is that systems out of equilibrium, even if attention is only confined to stationary states, are no longer described by Hamiltonian equations because nonconservative forces are supposed to act on the particles and the work that they produce has to be dissipated in the form of heat.

Equations of motion will be symbolically written as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{4}$$

and we shall assume that $\mathbf{f}$ is smooth as a function of $\mathbf{x}$ and of the external parameters, like volume or strength of the external forces.

Since we imagine that the system is subject to nonconservative forces the phase space volume (or any probability distribution with density with respect to the volume) will not be preserved by the evolution and the divergence

$$\sigma(\mathbf{x}) = -\sum_i \partial_{x_i} f_i(\mathbf{x}) \tag{5}$$

will be, in general, *different* from 0.

We shall suppose that the phase space visited by trajectories is bounded (at fixed external or control parameters) and that, aside for a set of initial data occupying 0 volume, the time averages of all smooth observables $\mathbf{x}\to F(\mathbf{x})$, namely

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t F(S_\tau\mathbf{x})\,d\tau = \int F(\mathbf{y})\,\mu(\mathbf{y}) \tag{6}$$

exist and define a unique *statistics* μ *i.e.* a unique probability distribution, invariant with respect to the evolution in Eq.(4).

Then one can try to follow Boltzmann's method and derive consequences of physical interest by conceiving phase space as discrete and time evolution as a permutation of its "discrete" volume elements.

A key obstacle is that in all interesting cases the time average of the phase space contraction, Eq.(6) with $F(\mathbf{x}) \equiv \sigma(\mathbf{x})$, that will be denoted σ_+ will be a *positive*

quantity, [17], implying that phase space will indefinitely contract as time evolves and no invariant distribution will be described by a density function: *hence* the statistics μ in Eq.(6) will be concentrated on a set of 0-volume, the *attracting set*.

Here the fundamental difference between equilibrium and non-equilibrium becomes manifest. If phase space is discretized then a faithful representation of the dynamics cannot be, not even approximately, a simple permutation of discrete volume elements unless special, non trivial, attention is devoted to their definition.

In equilibrium, *i.e.* when the system is described by Hamiltonian equations, the notion of *discrete element* of phase space and that of *coarse grained cell* of phase space are often confused: an identification that is probably induced by the validity of Liouville's theorem. However most Hamiltonian systems, and all the ones of interest in systems consisting of assemblies of interacting particles, while conserving volume strongly deform any phase space set as time evolves. And the same can be said, of course, when nonconservative forces act.

Therefore, if by discrete volume element is intended a cube of size small enough so that all the *few* observables of interest in a macroscopic thermodynamical description of the equilibrium (or stationary) states, are constant, one cannot say that evolution is simply a permutation of phase space cells.

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

In a numerical simulation the microscopic configurations of the N particles constituting the system are represented by small parallelepipeds *defined by the machine precision*: their number is clearly huge, even in computers operating with few bits chips. Evolution is by construction a map of such "microscopic cells" into themselves. *However*, as anyone who has ever attempted a simulation knows, the map is far from being a permutation. Unless very special care, see [18] for the only nontrivial example that I know, is devoted to the program design, it is unavoidable that the programmed map merges distinct microscopic cells (a violation of the uniquess inherent in the differential equation being simulated).

Thus, even when simulating Hamiltonian equations, it is inevitable that some dissipation occurs: microscopic cells will merge and phase space will effectively contract. But contraction cannot go on forever and any map will generate motions which are *eventually* a permutation of a subset of the microscopic cells, [16, 19]. This suggests that it should be possible to *unify* the treatment of equilibrium and nonequilibrium by regarding phase space as discrete in both cases and the discretized microscopic cells so small that motion is well represented by a map which, after a finite time, becomes a permutation of the microscopic cells lying on the attracting set.

The difference between equilibrium and nonequilibrium will then simply become that while in equilibrium the microscopic cells that are not on the attracting set can be neglected and the motion can be effectively regarded as a permutation, 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.

In both cases an *ergodic hypothesis* can be formulated: the permutation of the microscopic cells that are on the attracting set is *a one cycle permutation*: every microscopic cell visits any other.

III. COARSE GRAINING

To see that the above viewpoint really unifies equilibrium and nonequilibrium the *chaotic hypothesis* has been introduced which also allows us to give a precise definition of *coarse grained cells*. According to the hypothesis *the asymptotic motions of a confined chaotic mechanical system develop on an attracting set on which motion can be considered mixing and hyperbolic*, [19–23].

To enter into some detail it is convenient to look at the time evolution by drawing a surface Σ transversal to the phase space trajectories, and such that the trajectories cross Σ over and over again (*i.e.* each trajectory crosses Σ infinitely many times both in the future and in the past). The surface is usually called a *Poincaré's section* (it is not necessarily connected and it may consist of a few disjoint surface elements) defining the *timing events*.

Let S be the map which transforms a point $\xi \in \Sigma$ into the point $S\xi$ where the orbit of ξ meets again, for the first time, Σ. The points in phase space can therefore be described by pairs $\mathbf{x} = (\xi, \theta)$ if ξ is the point in Σ last visited by the trajectory starting at x and θ is the time elapsed since that moment: θ varies between 0 and the time $\tau(\xi)$ necessary for the crossing next to ξ to happen.

In what follows we shall discuss properties of the action of the map S on Σ: the statistics for this motion and the one for the original continuous time dynamical system in Eq.(4) are directly related by $\mu(d\xi d\theta) = \frac{d\theta}{\tau(\xi)}\,\overline{\mu}(d\xi)$ if $\overline{\mu}$ is the statistics of the map S (defined as in Eq.(6) with t integer and the integral replaced by a sum over integr τ).

It is possible, *under the chaotic hypothesis*, to partition Σ into regions $P_1, P_2, \ldots, P_n$ with the property that the symbolic dynamics histories $\boldsymbol{\sigma} = \{\sigma_i\}_{i=-\infty}^{\infty}$ on the sets P_σ, $\sigma = 1, \ldots, n$, has a *Markov property*, in the sense that

(1) there is a suitable matrix $M_{\sigma,\sigma'}$ with entries 0 or 1 and "transitive", in the sense that there is $m > 0$ for which $(M^m)_{\sigma,\sigma'} > 0$, such that if $M_{\sigma_i,\sigma_{i+1}} \equiv 1$ for all i then there is a unique point $\xi \in \Sigma$ such that $S^i\xi \in P_{\sigma_i}$: the point ξ is said to be "coded" by the sequence $\boldsymbol{\sigma}$. And

(2) calling *compatible* a sequence $\boldsymbol{\sigma}$ 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 relative to Σ the sequence $\boldsymbol{\sigma}$ is unique (*i.e.* much as it is the case in the binary representation of real numbers).

Remark: The above properties are highly nontrivial and the partition exists because of the assumed hyperbolicity properties. Symbolic dynamics built using generic partitions will assign to points of Σ sequences that are subject to long range constraints rather than to the simple "nearest neighbor" ones in (2).

The partition $P_1, P_2, \ldots, P_n$ is then called a *Markov partition*: since the set of exceptions in the correspondence $\xi \longleftrightarrow \boldsymbol{\sigma}$ has zero volume, the volume distribution on phase space Σ (often called "Liouville's distribution"), can be represented as a probability distribution μ_0 over the space of compatible sequences. And the statistics of the evolution of data ξ *chosen at random with respect to the distribution* μ_0, which is the main object of interest, [24], will therefore be represented also by a S–invariant probability distribution on the space Ω of the compatible sequences $\boldsymbol{\sigma}$, [23].

The sets $P_1, P_2, \ldots, P_n$ can be used to define *coarse grained* states of the system: given a precision $h > 0$ it is possible to find N_h such that the sets

$$\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}} = \bigcap_{j=-N_h}^{N_h} S^{-j} P_{\sigma_j} \tag{7}$$

have a diameter $< h$. Therefore the (nonempty among the) sets $\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}}$ can be conveniently used as *coarse cells* to describe the evolution, provided the size h is small enough for considering acceptable to neglect the variations of the (few) interesting observables within the Δ's. Increasing N_h increases the coarse grain precision, but the limit $h \to 0$ is never necessary because in the end only a few observables are necessary for describing macroscopic properties.

The evolution S will stretch Δ along the unstable manifolds of Σ and compress it along the stable ones: it will map $\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}}$ inside the union of the n sets $\cup_\sigma P_{\sigma_{-N_h+1},\ldots,\sigma_{N_h},\sigma}$, with the unstable manifolds "exactly fitting" as illustrated in Fig.1 below as a *consequence of property* (2) above.

We then imagine that the cell Δ is filled by *far smaller* boxes, that will be called *microscopic cells* or simply *microcells*, of equal volume, which under the action of S are transformed into boxes contained in only one of the n sets $\Delta_\sigma = P_{\sigma_{-N_h+1},\ldots,\sigma_{N_h},\sigma}$. Such microcells, which in a simulation could be *identified with the integers defining the system points coordinates and momenta in the computer memory*, should be 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.

The choice of equal volume microcells reflects the special role attributed to the volume measure: in Eq.(6) the statistics is identified *up to a set of exceptional points occupying* 0 *volume*. Under the chaotic hypothesis such statistics μ is determined uniquely.

Since the evolution, in the average, contracts phase space the layers will merge under the action of S so that the number of microcells will initially decrease; but eventually in each cell Δ will survive layers of microcells whose collection will be mapped one to one into themselves: the latter collection of microcells is a representation of the attractor, within the precision h. This is illustrated symbolically in Fig.1.

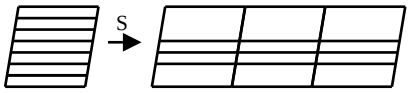

Fig.1: The lines symbolize arrays of microcells Δ: the ones in the left drawing are stretched and merged by the time evolution into arrays that end up exactly fitting into a few new boxes (three in the drawing), a consequence of (2) above.

For consistency the number of microcells that is eventually found in each coarse grain cell Δ is inversely proportional to the expansion coefficient $\Lambda_e(\Delta)^{-1}$ of the surface elements on the unstable manifold in Δ: it will be denoted $\mathcal{N}\pi(\Delta)$,

$$\pi(\Delta) = \frac{\Lambda_e^{-1}(\Delta)}{\sum_{\Delta'} \Lambda_e^{-1}(\Delta')} \tag{8}$$

The time evolution S can then be represented as a permutation of the $\mathcal{N}$ microcells on the attractor, as follows:

(a) give a rule to select the $\mathcal{N}\pi(\Delta)$ microcells in $\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}}$ and to partition them into n_Δ groups labeled by the labels σ of the $n_\delta \le n$ cells that the set $S\Delta$ intersects ("crossing" it by the propertes (1,2,3) above, see also Fig.1) of the microcells selected in Δ and
(b) establish a correspondence $\overline{S}$ between the selected microcells in the group labeled σ and a subset the ones selected in $\Delta_\sigma = P_{\sigma_{-N_h+1},\ldots,\sigma_{N_h},\sigma}$
(c) approximate S by replacing it by $\overline{S}$.

Certainly there is a lot of ambiguity in deciding how to set up the selection and the correspondence: but for the purposes of a description of dynamics with precision h the ambiguity has no consequence. Note that in simulations the microcells selection is implicitly prescribed by the program, and certainly changes quite substantially by any small change of the program or by a change of the computer used. By definition of computer program the evolution S is replaced by a map of microcells (in huge number even in simple simulations): the map is not invertible but being a map of a finite set into itself it is eventually reduced to a permutation of a *subset* of the microcells.

Transitivity of the compatibility matrix M implies that the permutation of the microcells on the attractor can be chosen cyclic: therefore the stationary distribution $\overline{\mu}$

will be approximated by the *uniform distribution* on the attractor: this is a picture which seems close to BOLTZMANN's conception of discretness and extends the ideas behind the ergodic hypothesis to systems more general than Hamiltonian, [16, 25]. Of course this means that the distribution $\overline{\mu}$, or the corresponding continuum time μ, can be seen, *exactly as in the equilibrium cases*, as the "equidistribution" among the cells on the attracting set. The difference is only that in the nonequilibrium case the attracting set is not the full energy surface: but it is a set, rather difficult to describe which, in the considered representation of motion, is a "coarse grained" approximation of a set of zero phase space volume.

A remarkable property of the statistics $\overline{\mu}$ emerges when it is regarded as a probability distribution on the compatible sequences $\boldsymbol{\sigma}$ which code the points $\mathbf{x}$ in phase space Σ. Namely it is a "Gibbs distribution", in the sense of probability theory, with a short range potential: this is, essentially, a Markov process with finitely many states, *i.e.* an object that is very well understood, [26]. This is a surprising consequence, saying that hyperbolic systems are "completely integrable", in the sense that we can compute essentially everything, at least in principle, [23, 27]. They become a *paradigm* for chaotic evolutions in the same way in which harmonic oscillators are a paradigm for ordered motions.

Hyperbolicity, however, is a strong property which in practice is not satisfied in physical systems: the hypothesis set at the beginning of the section has to be interpreted in the same spirit of the viewpoint expressed by RUELLE in [28] "*... 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...*".

The natural distribution $\overline{\mu}$ approximately defined by the prescriptions (a,b,c) above, *i.e.* by Eq.(8) (and strictly defined by it in the limit $h \to 0$) has since been called the *SRB distribution* and RUELLE has proposed, in the very similar context of turbulence theory and at least three years earlier than it appeared in print [22, 29], that in general there should be a unique distribution (or possibly a finite number of them) describing the experimental statistics of motions: and it should be the distribution giving the asymptotic behavior of motions with arbitrary initial data apart from a set of zero volume.

Eq.(8) is a concrete representation of the probability distribution $\overline{\mu}$ that has to be used to compute the statistical averages of the observables in the stationary state reached by the system. It is remarkable not only because it unifies equilibrium and nonequilibrium but also, and perhaps mainly, because it can be used to provide formal expressions for average values of observables. Although the actual averages cannot be computed still their formal expression may allow us to find relations between them, as it is done in equilibrium (where important relations can be derived, like the *heat theorem* where Eq.(1) can be established without actually being able to give the expressions for the functions $p(U, V)$ and $T(U, V)$.

IV. ENTROPY?

The above discussion shows that BOLTZMANN's method of deriving properties of complex systems by regarding them as discrete can be still very fruitful. And it has led to a few concrete results, among which the *fluctuation theorem*, [16, 20], which deals with a symmetry property of the finite time averages of the phase space contraction σ in Eq.(1). This could be of applicative interest provided a physical interpretation of the phase space contraction is proposed: and research in this direction leads again to a discussion of the notion of *entropy* and to the natural question of whether a definition of entropy can be extended to nonequilibrium stationary states in analogy with the corresponding definition for equilibrium states (which are a very special case of stationary states).

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

If an initial distribution is given, arbitarily assigning a probability $p(c)$ with each microcell c, it will evolve (in the sense of convergence of the average values of smooth observables) towards the SRB distribution if the chaotic hypothesis is accepted. So that $-\sum_c p(c) \log p(c)$ will tend to a maximum equal to the logarithm of the number $\mathcal{N}_h$ of microcells (no matter which value of h is chosen): in fact the one-cycle property of the evolution of the microcells implies that any distribution evolves towards the distribution giving equal weight to each microcell.

This indicates that it will be possible to have a kind of H-theorem, *i.e.* the existence of a Lyapunov function for the evolution towards the stationary state. However the number $\mathcal{N}_h$ can be estimated, [30], and its logarithm seems to depend on the precision h. The dependence is, *remarkably*, an additive constant solely depending on the rather arbitrary value of h *only in the case of equilibrium states*. For nonequilibrium states it depends nontrivially also on other parameters. This seems to indicate that it will not be possible to define entropy of a stationary state unambiguously, although a kind of H-theorem could still hold.

Arguments in this direction can be found in the literature, [31, 32], and the controversial aspects of this matter will not be touched, [33].

We shall here discuss, having in mind the possibility of experimental tests, more in detail an attempt to find a physical interpretation for the divergence, Eq.(5), of the equations of motion.

Several simulations have led to propose identification between the phase space contraction with "entropy creation rate". This is a simpler notion than entropy itself, at least in systems in which there is a clear separation between "system" and surrounding "thermostats": the

amount of heat ceded to the thermostats and their temperature should be clearly identifiable in such cases. The entropy *created* during the stationary processes could be identified with the increase of entropy of the thermostats. Which is well defined being the sum of the ratios of the heat received by the various thermostats to their temperatures, because the thermostats should be regarded in thermodynamic equilibrium and therefore there is no ambiguity in defining their entropy variations.

The question will be investigated conveniently in a simple, but quite general, model of thermostatted system: it will appear that although there is a relation between entropy creation rate and phase space contraction, still the two notions are quite different. *Nevertheless* the difference can be expressed as a *total time derivative* of a suitable observable and therefore its has no influence, [34], or a controlled influence, [35], on the fluctuation theorem making it, in principle, experimentally testable in experiments in which heat transferred to thermostats is measurable.

The system consists in $N \equiv N_0$ particles in a container $\mathcal{C}_0$ and of N_a particles in n containers $\mathcal{C}_a$ which play the role of *thermostats*: their positions will be denoted $\mathbf{X}_a$, $a = 0, 1, \ldots, n$, and $\mathbf{X} \stackrel{def}{=} (\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_n)$. Interactions will be described by a potential energy

$$W(\mathbf{X}) = \sum_{a=0}^{n} U_a(\mathbf{X}_a) + \sum_{a=1}^{n} W_a(\mathbf{X}_0, \mathbf{X}_a) \tag{9}$$

i.e. thermostats particles only interact indirectly, via the system. All masses will be $m = 1$, for simplicity.

Particles in $\mathcal{C}_0$ will also be subject to external, possibly nonconservative, forces $\mathbf{F}(\mathbf{X}_0, \boldsymbol{\Phi})$ depending on a few strength parameters $\boldsymbol{\Phi} = (E_1, E_2, \ldots)$. It is convenient to imagine that the force due to the confining potential determining the geometrical of the region $\mathcal{C}_0$ is included in $\mathbf{F}$, so that one of the parameters is the volume $V = |\mathcal{C}_0|$. See Fig.2 below.

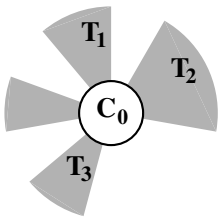

Fig.2 The reservoirs occupy finite regions outside C_0, *e.g.* sectors $C_a \subset R^3$, $a = 1, 2 \ldots$. Their particles are constrained to have a *total* kinetic energy K_a constant, by suitable forces $\boldsymbol{\vartheta}_a$, so that the reservoirs "temperatures" T_a, are well defined, by $K_a = \sum_{j=1}^{N_a} \frac{1}{2} (\dot{\mathbf{X}}_j^a)^2 \stackrel{def}{=} \frac{3}{2} N_a k_B T_a \stackrel{def}{=} \frac{3}{2} N_a \beta_a^{-1}$.

The equations of motion will be, assuming unit mass,

$$\ddot{\mathbf{X}}_{0i} = -\partial_i U_0(\mathbf{X}_0) - \sum_a \partial_i U_a(\mathbf{X}_0, \mathbf{X}_i) + \mathbf{F}_i \ddot{\mathbf{X}}_{ai} =$$
$$= -\partial_i U_a(\mathbf{X}_a) - \partial_i U_a(\mathbf{X}_0, \mathbf{X}_i) - \alpha_a \dot{\mathbf{X}}_a \tag{10}$$

where the last term $-\alpha_a \dot{\mathbf{X}}_a$ is a phenomenological force that implies that thermostats particles keep constant kinetic energies: α_a can, for instance, be defined by

$$-\alpha_a \stackrel{def}{=} \frac{L_a - \dot{U}_a}{3 N_a k_B T_a} \tag{11}$$

where $L_a = -\partial_{\mathbf{X}_a} W_a(\mathbf{X}_0, \mathbf{X}_a) \cdot \dot{\mathbf{X}}_a$ is the work done per unit time by the forces that the particles in $\mathcal{C}_0$ exert on the particles in $\mathcal{C}_a$; here k_B denotes Boltzmann's constant.

The exact form of the forces that have to be added in order to insure the kinetic energies constancy should not really matter, within wide limits. But this is a property that is not obvious and which is much debated. The above thermostatting forces choice is dictated by Gauss' *least effort* principle for the constraints $K_a = const$, see [16, Appendix 9.4]: this is a criterion that has been adopted in several simulations, [36]. Independently of Gauss' principle it is immediate to check that if α_a is defined by Eq.(11) then the kinetic energies K_a are strictly constants of motion.

The work L_a in Eq.(11) will be interpreted as *heat* $\dot{Q}_a$ ceded, per unit time, by the particles in $\mathcal{C}_0$ to the a-th thermostat (because the "temperature" of $\mathcal{C}_a$ remains constant, hence the thermostats can be regarded in thermal equilibrium). The *entropy creation rate* due to heat exchanges between the system and the thermostats can, therefore, be naturally defined by

$$\sigma^0(\dot{\mathbf{X}}, \mathbf{X}) \stackrel{def}{=} \sum_{a=1}^{N_a} \frac{\dot{Q}_a}{k_B T_a} \tag{12}$$

It should be stressed that here *no entropy notion* is introduced for the stationary state: only variation of the thermostats entropy is considered and it should not be regarded as a new quantity because the thermostats should be considered in equilibrium at a fixed temperature.

The question is whether there is any relation between σ_0 and the phase space contraction σ of Eq.(5) for the equations of motion in Eq.(10). The latter can be immediately computed and is (neglecting $O(\max_{a>0} N_a^{-1})$)

$$\sigma^\Gamma(\dot{\mathbf{X}}, \mathbf{X}) = \sum_{a>0} \frac{3N_a - 1}{3N_a} \frac{\dot{Q}_a - \dot{U}_a}{k_B T_a} = \sum_{a>0} \frac{\dot{Q}_a}{k_B T_a} - \dot{U} \tag{13}$$

where $U = \sum_{a>0} \frac{3N_a - 1}{3N_a} \frac{U_a}{k_B T_a}$. Hence in this example, physically interesting, in which the thermostats are "external" to the system volume (unlike to what happens in the common examples in which they act inside the volume of the system), the phase space contraction is not

the entropy creation rate, [34]. *However it differs from the entropy creation rate by a total derivative.*

This is relevant because the definition Eq.(12) *has meaning independently of the equations of motions* and can, therefore, be suitable for experimental tests, [34, 35, 37].

Finally I mention that the identification, up to a total time derivative, of phase space contraction with entropy creation rate can be shown to cover the entropy creation rate in systems whose evolution can be approximated by macroscopic continua equations, like fluids described by Navier Stokes equations. In the sense that, again, phase space contraction of the interacting particles systems that underly the macroscopic equations is related, in the stationary states, to the entropy creation rate independently defined in classical nonequilibrium Thermodynamics, [38]: and it differs from it by a total time derivative, [37].

V. 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 and modernity of his thought. It hands us a legacy that is often obscured by his other major contribution represented by the Boltzmann equation and its very acute interpretation ([39], see also the short note [40] on BURBURY's comments).

It is a method of investigation which is still very fruitful and used in various forms when "cut–offs" or "regularizations" are employed in the most diverse fields. In my view it has been and still is important in the recent developments in the theory 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, Onsager-Machlup fluctuations theory, see [16, 23, 41, 42]) is, hopefully, only an example.

In the Appendix I quote a few paragraphs from BOLTZMANN which illustrate his strong feelings about discrete views versus continuum viewpoints.

Appendix

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*", [43, p.227];

"*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*, [43, p.227];

"*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.*", [43, p.55].

for instance.

Source of the talk at the Boltzmann Symposium in München, 11-13 October, 2006

[1] J. Gibbs, *Elementary principles in statistical mechanics (reprint)* (Schribner, Cambridge (USA), 1902).

[2] S. Brush, *History of modern physical sciences: The kinetic theory of gases* (Imperial College Press, London, 2003).

[3] L. Boltzmann, *Über die mechanische Bedeutung des zweiten Haupsatzes der Wärmetheorie*, vol. 1, p.9 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[4] L. Boltzmann, *Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen*, vol. 1, p.237 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[5] L. Boltzmann, *Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten*, vol. 1, p.49 of *Wissenschaftliche Abhandlungen, ed. F. Hasen-
öhrl* (Chelsea, New York, 1968).

[6] L. Boltzmann, *Entropie und Wahrscheinlichkeit*, vol. 286 of *Ostwalds Klassiker der Exacten Wissenschaften, Ed.*

D. Flamm (Verlag Harri Deutsch, ISBN 3-8171-3286-7, Frankfurt am Main, 2000).

[7] L. Boltzmann, *Einige allgemeine sätze über Wärmegleichgewicht*, vol. 1, p.259 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[8] P. Ehrenfest and T. Ehrenfest, *The conceptual foundations of the statistical approach in Mechanics* (Dover, New York, 1990).

[9] S. Brush, *The kind of motion that we call heat, (I, II)* (North Holland, Amsterdam, 1976).

[10] L. Boltzmann, *Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respektive den Sätzen über das Wärmegleichgewicht*, vol. 2, p.164 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[11] L. Boltzmann, *Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht des lebendigen Kraft*, vol. 1, p.288 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[12] L. Boltzmann, *Bemerkungen über einige Probleme der mechanischen Wärmetheorie*, vol. 2, p.112 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[13] H. Helmholtz, *Prinzipien der Statistik monocyklischer Systeme*, vol. III of *Wissenschaftliche Abhandlungen* (Barth, Leipzig, 1895).

[14] H. Helmholtz, *Studien zur Statistik monocyklischer Systeme*, vol. III of *Wissenschaftliche Abhandlungen* (Barth, Leipzig, 1895).

[15] L. Boltzmann, *Über die Eigenshaften monozyklischer und anderer damit verwandter Systeme*, vol. 3, p.122 of *Wissenschaftliche Abhandlungen* (Chelsea, New-York, 1968).

[16] G. Gallavotti, *Statistical Mechanics. A short treatise* (Springer Verlag, Berlin, 2000).

[17] D. Ruelle, Journal of Statistical Physics **85**, 1 (1996).

[18] D. Levesque and L. Verlet, Journal of Statistical Physics **72**, 519 (1993).

[19] G. Gallavotti, Journal of Statistical Physics **78**, 1571 (1995).

[20] G. Gallavotti and E. G. D. Cohen, Physical Review Letters **74**, 2694 (1995).

[21] G. Gallavotti and E. G. D. Cohen, Journal of Statistical Physics **80**, 931 (1995).

[22] D. Ruelle, Annals of the New York Academy of Sciences **357**, 1 (1980).

[23] G. Gallavotti, F. Bonetto, and G. Gentile, *Aspects of the ergodic, qualitative and statistical theory of motion* (Springer Verlag, Berlin, 2004).

[24] D. Ruelle, Journal of Statistical Physics **95**, 393 (1999).

[25] G. Gallavotti, Journal of Statistical Physics **84**, 899 (1996).

[26] Y. G. Sinai, Functional analysis and Applications **2**, 70 (1968).

[27] Y. G. Sinai, Russian Mathematical Surveys **27**, 21 (1972).

[28] D. Ruelle, *Ergodic theory*, vol. Suppl X of *The Boltzmann equation, ed. E.G.D Cohen, W. Thirring, Acta Physica Austriaca* (Springer, New York, 1973).

[29] D. Ruelle, American Journal of Mathematics **98**, 619 (1976).

[30] G. Gallavotti, Communication in Mathematical Physics **224**, 107 (2001).

[31] G. Gallavotti, Chaos **14**, 680 (2004).

[32] G. Gallavotti, cond-mat/0606477 (2006).

[33] S. Goldstein and J. L. Lebowitz, Physica D **193**, 53 (2004).

[34] G. Gallavotti, Chaos **16**, 023130 (+7) (2006).

[35] F. Bonetto, G. Gallavotti, A. Giuliani, and F. Zamponi, Journal of Statistical Mechanics p. P05009 (2006).

[36] D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Fluids* (Academic Press, New-York, 1990).

[37] G. Gallavotti, Journal of Statistical Plysics and cond-mat/0608444 (2006).

[38] S. de Groot and P. Mazur, *Non equilibrium thermodynamics* (Dover, Mineola, NY, 1984).

[39] L. Boltzmann, *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, vol. 1, p.316 of *Wissenschaftliche Abhandlungen, ed. F. Hasenöhrl* (Chelsea, New York, 1968).

[40] L. Boltzmann, *On the minimum theorem in the theory of gases*, vol. 3, p.546 of *Wissenschaftliche Abhandlungen* (Chelsea, New-York, 1968).

[41] J. Lebowitz and H. Spohn, Journal of Statistical Physics **95**, 333 (1999).

[42] G. Gallavotti, cond-mat/0402676 (2004).

[43] L. Boltzmann, *Theoretical Physics and philosophical writings, ed. B. Mc Guinness* (Reidel, Dordrecht, 1974).

REVTEX