

ICT2.0:2020-W172

A PATH FROM EQUILIBRIUM TO NONEQUILIBRIUM THERMODYNAMICS

Giovanni Gallavotti

Dipartimento di Fisica

Università di Roma “La Sapienza”

00183 Roma, Italia

Email: giovanni.gallavotti@roma1.infn.it

ABSTRACT

In the last decades new ideas have arisen in thermodynamics as a consequence of the advances in the applications of dynamical systems theory and results to evolution of macroscopic systems. A key question is whether it is possible to extend the domain of thermodynamics from a general theory of relations between equilibrium states to the much more general analysis of relations and properties of stationary states arising under steady forcing of mechanical systems in contact with heat reservoirs. The familiar description of equilibrium via probability distributions (like the micro-canonical ensemble) which predicts relations between observables extending the ones of classical thermodynamics has been widely extended and under general assumptions an extension of the theory of ensembles has been proposed. The new theory contains classic thermodynamics, as derived from statistical mechanics, as a special case: it relies on “chaos” and raises important questions that are still debated: is there a suitable extension of the “entropy”, are there universal relations that can be theoretically predicted?

INTRODUCTION: Equilibrium

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.

Equilibrium thermodynamics was born in the early 1800's. On the basis of the experimental results on the properties of rarefied gases (like the Boyle-Mariotte law) Carnot develops, [1], his theorem on the ideal efficiency of machines operating by extracting or damping heat in two reservoirs and converting it into work. His long, very careful and detailed, analysis is presented after a brief introduction (10 pages), starting with the above words, in which present and future usefulness and importance of the vapor machines is enthusiastically (and optimistically) outlined (1824).

The theorem shows that the most efficient machines must operate running in a reversible cycle in which the vapor (of water, air, alcohol or other gases) evolves through a sequence $\mathcal{P}$ of equilibrium states;¹

¹Reversible transformations were essential in Carnot's analysis, and he carefully insists to make clear the subtle argument that permits to avoid regarding their definition, requiring for instance
Copyright © 2020 by ASME

and this remains true even if the vapor is replaced by a liquid or a solid.

It is an example of what today we call a “universal law”, *i.e.* a law that applies to a very large class of systems isolating, among their properties, a few that they all verify in a quantitative form *without adjustable parameters*.

An immediate consequence is the possibility of defining the absolute temperature of a heat reservoir: it is simply defined in terms of the maximum efficiency of a machine operating between the reservoir of interest and a fixed reservoir to which a conventional temperature value is attributed.²

A few years later Krönig, [2], established the proportionality of the absolute temperature to the average kinetic energy and Clausius, [3], wrote the first of the works leading to *entropy*, 1850, whose existence constitutes the *second law* of thermodynamics and is implied by Carnot’s theorem.

The meaning of the word “entropy” was explained by Clausius himself [4, p.390]:

“I propose to name the quantity S the entropy of the system, after the Greek word ἡ τροπή “the transformation”, [in German, Verwandlung]. I have deliberately chosen the word entropy to be as similar as possible to the word energy: the two quantities to be named by these words are so closely related in physical significance that a certain similarity in their names appears to be appropriate.”

The notion of entropy became quickly fundamental for the theory and applications of equilibrium thermodynamics and almost identified with it: soon it was accompanied by the question of which would be its definition in terms of the atomistic representation of

matter.

And in 1866 Boltzmann, [5], proposed to link it to the purely mechanical Maupertuis’s *Action Principle*, imagining that atoms were moving on periodic orbits spanning all phase space points compatible with the mechanical conservation laws.³

This was a first attempt towards the formulation of the *ergodic hypothesis* and two years later, [8], led to the description of the statistical properties of a system in equilibrium via what are now called *microcanonical and canonical ensembles*. The consequences developed by Boltzmann in 1868 were recognized by Maxwell (1879) who commented them [9],⁴ and by Gibbs, [10].

Entropy S_A is defined for every equilibrium state A of a given system and is useful to establish the balance between heat exchanged by a system with the surrounding thermostats⁵ and work performed on the surrounding mechanical devices.

A key property is that if a transformation through a “path” $\mathcal{P}$ of successive equilibrium states is *reversible* and leads from an equilibrium state A to an equilibrium state B then the entropy variation is $S_B - S_A = \int_{\mathcal{P}} \frac{dQ}{T}$, if dQ is the amount of heat that the system receives while in contact with a heat reservoir at absolute temperature T , whatever the reversible path $\mathcal{P}$ is. This allows to set the value of the entropy of a generic equilibrium state, determining it up to an additive constant.

However entropy is important also in irreversible processes $\mathcal{I}$: it was established by Clausius that in such a process $S_B - S_A \geq \int_{\mathcal{I}} \frac{dQ}{T}$, where T is the absolute temperature of the reservoir from which the system receives the amount dQ of heat.

In particular if the transformation $A \rightarrow A$ is along an irreversible path $\mathcal{I}$ it is $\oint_{\mathcal{I}} \frac{dQ}{T} \leq 0$. Also if $\mathcal{I}$ is an irreversible adiabatic path (*i.e.* a sequence of trans-

a difference in temperature which “can be considered as vanishing”, an oxymoron: “À la vérité, les choses ne peuvent pas se passer rigoureusement comme nous l’avons supposé ..., [1, p.13-14].

²*e.g.* if the reservoir is water at its triple point then in the Kelvin scale the absolute temperature is fixed to be $T_0 = 273.16^\circ\text{K}$. The temperature T_1 of another reservoir at higher temperature (say) is then given, in principle, by running a reversible machine between the two reservoirs and deriving T_1 so that the efficiency is $1 - \frac{T_0}{T_1}$.

³The idea was again proposed four years later by Clausius, [6, 7].

⁴Interestingly *without even mentioning* the *H*-theorem, discovered in the meantime by Boltzmann.

⁵*i.e.* large bodies whose temperature and volume variations are negligible during the observation time

formations with no heat exchange) it cannot lead from A to A , and necessarily leads from A to a state B with larger entropy.

If a system evolves first from a state A to B at constant temperature T_2 receiving a quantity of heat Q_2 , then evolves adiabatically from B to B' at temperature T_1 , then to A' at the temperature T_1 ejecting a quantity of heat Q_1 and then adiabatically back to A it will be

$$\frac{Q_2}{T_2} - \frac{Q_1}{T_1} \leq 0, \quad \text{which means: } \frac{Q_2 - Q_1}{Q_2} \leq \frac{T_2 - T_1}{T_2}$$

Since $Q_2 - Q_1$ must be, by energy conservation, the work performed by the system in the cycle, the inequality means that the efficiency in the transformation of heat into work cannot exceed $\frac{T_2 - T_1}{T_2}$, and it can equal it if the cycle is a reversible path: which is Carnot's theorem formulated in terms of Clausius' entropy properties.

The continuous use of the second law, hence of Carnot's theorem, led to a long debate on the ergodic hypothesis, on entropy, and on the resolution of the antimony between reversible microscopic dynamics and irreversible macroscopic evolutions, which in some respects is still going on.

Microscopic stationarity

In recent times there has been a widespread interest in extending Thermodynamics to a theory dealing with stationary states: these are states of systems on which steady non-conservative forces may act, keeping currents flowing through the systems: and currents can be of various kinds like transporting matter, heat, electric charge, ...

The stationary states are a natural generalization of the equilibrium states (which are very special cases of stationary states) and, of course, a main question is whether general, system independent, properties can be assigned to such states and be useful in studying their properties.

To understand the recent developments it is essential to keep in mind that the study of a physical system

starts from an initial datum X where X is a point in "phase space" $\mathcal{F}$, often determined by the $6N$ coordinates of the N molecules or atoms. And the point X is generated via suitable experimental devices (it has become common to say "*following a prefixed protocol*") and is a *random set of coordinates*, because of the many unavoidable uncontrolled actions influencing the devices.

The probability of selecting the initial datum X in $\mathcal{F}$ is essentially always unknown, but it is *assumed to have a density*: which means that there is a function $\rho(X)$ such that $\rho(X)dX$ is the probability that, repeating the experiment many times, the datum X falls in the volume element dX . Such a probability distribution is commonly called "absolutely continuous" (with respect to the volume).

Here attention will be devoted to data X generated randomly with "absolutely continuous" distribution. It will be seen that the assumption of existence, behind the protocol of the observation, of an unknown but absolutely continuous probability distribution for the initial data generation is *far from obvious* and in the end it turns out to be a *far reaching physical law*, that could be called law of "initial chaos".

Given the initial data X , the system will evolve describing in time a trajectory denoted $t \rightarrow S_t X = X(t)$. In theoretical studies the evolution $X(t)$ will be supposed to be defined by a solution $X(t)$ of a differential equations $\dot{X} = F(X)$.

The interaction with the thermostats is influential and cannot be ignored. This means that the evolution equations must involve the interaction with the surroundings: which is a difficulty because it leads, eventually, to consider infinite systems.⁶

The same difficulty is met even in equilibrium when it is imagined that the temperature is fixed via the contact with a single thermostat.⁷

⁶i.e. interaction with the surroundings which interact also with their own surroundings, which interact with their own surroundings,

⁷In this respect Boltzmann in his 1868 paper shewed that in a very large system in a microcanonical equilibrium a finite volume subsystem behaves as if in contact with a reservoir with a

The difficulty is bypassed, in theoretical studies, by using dynamical models involving finitely many particles obeying equations of motion in which are introduced phenomenological forces (typically proportional to currents, circulating in the system, via “transport coefficients”). Such forces model dissipative effects controlling the actions of the non conservative forces or act by imposing suitable constraints for the same purpose. This gives rise to models which involve forces that are believed to be *physically equivalent* to more physical infinite thermostats but contain *finitely many particles*, [11, 12].

Given the equations of motion, an initial datum X will evolve into $X(t)$, reaching a *stationary state*.

Stationarity is in the sense that the observables will fluctuate in time but will have well defined time independent “*statistics*”, *i.e.* distribution of fluctuations around well defined time averages; and for quite a few observables even their values will be essentially constant in time (no need to average them) if the system is large (*i.e.* it consists of a large number of molecules).

In particular, from a microscopic viewpoint, the phase point representing at time t the microscopic state $S_t X$ of the system started in the configuration X , will spend a fraction of the total observation time ϑ visiting, in the limit $\vartheta \rightarrow \infty$, any chosen open set⁸ Δ with a well defined frequency $P_0(\Delta)$.

Therefore imagining phase space divided into small regions $\{\Delta_i\}_{i=1,\dots,\mathcal{N}}$ the frequencies of visit $P_0(\Delta_i)$ will determine the average values of all observables O whose variation in each of the Δ_i is negligible, so that their values can be denoted $O(\Delta_i)$: their time average will be $\langle O \rangle = \sum_i O(\Delta_i) P_0(\Delta_i)$.

Of course the size of the sets Δ_i must be adapted to the observables that are studied: therefore the choice of the Δ_i can be regarded as a choice of a “coarse graining” of phase space.

The frequencies $P_0(\Delta_i)$ will arise from a probability distribution P_0 on $\mathcal{F}$ *invariant* with respect to the evolution: in the sense that if X evolves in time τ into

$S_\tau X$ then *any* (reasonable, *e.g.* open) set Δ evolves into some $S_\tau \Delta$ and the two sets are visited with the *same frequency* $P_0(\Delta) = P_0(S_\tau \Delta)$.

A fundamental theorem, by Sinai in its simplest form, states that if the motion of the system is “*chaotic*”, as is the case for virtually all systems ultimately constituted by atoms, then the probability is 1 that a protocol, of the above mentioned kind for obtaining the initial datum X , generates a motion $S_t X$ which visits the sets Δ in phase space with well defined frequency $P_0(\Delta)$, [13–17].

Theorem: there is a unique probability distribution P_0 (called “SRB distribution”) determining the frequencies $P_0(\Delta)$ and therefore the statistical properties of the motion. Furthermore:

- (a) *the frequencies $P_0(\Delta)$ are independent on the protocol generating X , provided it generates random X ’s with an absolutely continuous distribution,*
- (b) *and given the protocol the frequencies are X -independent with probability 1,*
- (c) *and if there is one invariant distribution which is absolutely continuous then it must be P_0 .*

The requirement, aside from the mentioned law of initial chaos, is that the equations of motion generate *chaotic motions in a mathematically precise sense* that will be briefly called “*chaotic hypothesis*”, or CH, essentially meaning that

Chaotic Hypothesis: any datum X evolves into $S_t X \equiv X(t)$ for $t > 0$,

- (a) *never stopping (*i.e.* $\min |\dot{X}| > 0$) and,*
- (b) *after a transient time t_0 , an observer following $S_t X = X(t)$ and oriented as $\dot{X}(t)$ will see $X(t)$ as a hyperbolic fixed point, *i.e.* as a saddle point, and*
- (c) *data $X \in \rho A$ evolve so that $S_t X$ covers densely a smooth surface ρA (“attracting set”).*

The theorem applies to any system with this property, whether isolated (as in studying equilibrium) or in contact with external thermostats and under action of non conservative forces (as in general stationary states).⁹

temperature fixed and is described by a canonical ensemble.

⁸Hence any closed, or just “measurable”, set in $\mathcal{F}$.

⁹The requirement that the attracting set is a smooth surface can be weakened replacing “smooth” with “closed”: in the first case the motion on ρA is a “Anosov system” while in the second
Copyright © 2020 by ASME

The result eliminates an essential difficulty. Leaving aside the need to adapt the cells Δ_i to the observables selected for analysis, if the equations of motion generate chaotic motions¹⁰ then there are, as a mathematical theorem, many (*actually infinitely many*) stationary probability distributions $P \neq P_0$ that are invariant: *i.e.* in the cases covered by the above theorem there are data which visit the sets Δ with frequencies $P(\Delta)$ different from $P_0(\Delta)$, *i.e.* different from the SRB distribution.

Hence, without applying the theorem, requiring stationarity does not provide a recipe to determine the statistical properties of the motion.¹¹

The theorem allows to select unambiguously *which is the distribution* that determines the statistics P_0 observed in a given case, hence which is the *physically important statistics* associated with the equations of motion: it is valid both in equilibrium cases and in nonequilibrium ones.

Of course this implies accepting as a *physical law* that any protocol generating the system initial configurations X is “absolutely continuous” and, furthermore, the evolution of such X ’s is chaotic (leaving aside evident exceptions): this law (proposed by Ruelle, [18, 19], together with its main implications), cannot be proved and it has to be accepted as a law of Physics.

The (idealized)¹² case of isolated systems, *i.e.* the equilibrium cases for systems obeying Hamiltonian equations of motion with a fixed energy, have the property that one, P_0 , among the many stationary distributions compatible with the constraints (usually just with a fixed value of the energy), is such that $P_0(\Delta)$ can be expressed simply as an integral over phase space $P_0(\Delta) \equiv \int_{\Delta} \rho_0(X) dX$ of a density function

it satisfies “Axiom A”, [18].

¹⁰As it is the case apart from very few remarkable exceptions, like the arrays of elastic oscillators with interactions linear in the relative distances.

¹¹In equilibrium the ergodic hypothesis, for instance, has to be added to select the microcanonical ensemble distributions as the ones describing the statistics of the motions.

¹²Because it is only possible to contemplate a protocol that generates initial data constrained to have a prefixed energy

$\rho_0(X)$, which exists as a consequence of Liouville’s theorem.¹³

Hence P_0 , under the chaotic hypothesis, determines the statistical properties of the motions X . It can be said that isolated systems not only conserve the total energy but also admit an *invariant* absolutely continuous way of measuring the phase space elements of the energy surface (here invariant means that $P_0(\Delta) = P_0(S_t \Delta)$ for all t, Δ). This also shows that the CH implies, for isolated systems, the ergodic hypothesis.

In the literature several checks and applications of the new ideas can be found: among them I quote one the early consequence of the CH which gives a universally valid extension to systems stationary under small forcing *non zero* and *time reversible* of the Onsager reciprocal relations, [20]. Further extensions have been studied for irreversibly evolving systems: I just select a study about the Navier-Stokes equation in 2 dimensions, [21].

CITING REFERENCES

REFERENCES

- [1] Carnot, S., 1824. *Réflexions 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.
- [2] Krönig, A., 1856. “Grundzüge einer Theorie der Gase”. *Annalen der Physik und Chemie*, **XCIX**, pp. 315–322.
- [3] Clausius, R., 1851. “On the motive power of heat, and on the laws which can be deduced from it for the theory of heat”. *Philosophical Magazine*, **2**, pp. 1–102.
- [4] Clausius, R., 1865. “Über einige für Anwendung bequeme formen der Hauptgleichungen der mechanischen Wärmetheorie”. *Annalen der Physik und Chemie*, **125**, pp. 353–401.

¹³Which states that any Hamiltonian system admits an invariant distribution with density over the surface of fixed energy (which is usually the only constraint).

- [5] Boltzmann, L., 1866. “Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie”. *Wiener Berichte*, **53**, (W.A.,#2), pp. 195–220, (9–33).
- [6] Clausius, R., 1871. “Ueber die Zurückführung des zweiten Hauptsatzes der mechanischen Wärmetheorie und allgemeine mechanische Prinzipien”. *Annalen der Physik*, **142**, pp. 433–461.
- [7] Clausius, R., 1872. “Bemerkungen zu der priorität reclamation des hrn. boltzmann”. *Annalen der Physik*, **144**, pp. 265–280.
- [8] Boltzmann, L., 1868. “Studien über das gleichgewicht der lebendigen kraft zwischen bewegten materiellen punkten”. *Wiener Berichte*, **58**, (W.A.,#5), pp. 517–560, (49–96).
- [9] Maxwell, J. C., 1879. “On Boltzmann’s theorem on the average distribution of energy in a system of material points”. In: *The Scientific Papers of J.C. Maxwell*, Cambridge University Press, Ed. W.D. Niven, Vol.2, pp. 713–741.
- [10] Gibbs, J., 1902. *Elementary principles in statistical mechanics*. Schribner, Cambridge.
- [11] Evans, D. J., and Morriss, G. P., 1990. *Statistical Mechanics of Nonequilibrium Fluids*. Academic Press, New-York.
- [12] Gallavotti, G., 2014. *Nonequilibrium and irreversibility*. Theoretical and Mathematical Physics. Springer-Verlag.
- [13] Sinai, Y. G., 1968. “Markov partitions and C-diffeomorphisms”. *Functional Analysis and its Applications*, **2**(1), pp. 64–89.
- [14] Sinai, Y. G., 1968. “Construction of Markov partitions”. *Functional Analysis and its Applications*, **2**(3), pp. 70–80.
- [15] Sinai, Y., 1977. *Lectures in ergodic theory*. Lecture notes in Mathematics. Princeton University Press, Princeton.
- [16] Bowen, R., 1975. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Vol. 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-Heidelberg.
- [17] Bowen, R., and Ruelle, D., 1975. “The ergodic theory of axiom A flows”. *Inventiones Mathematicae*, **29**, pp. 181–205.
- [18] Ruelle, D., 1989. *Chaotic motions and strange attractors*. Accademia Nazionale dei Lincei, Cambridge University Press, Cambridge.
- [19] Ruelle, D., 1989. *Elements of differentiable dynamics and bifurcation theory*. Academic Press, New-York.
- [20] Gallavotti, G., 1996. “Chaotic hypothesis: Onsager reciprocity and fluctuation–dissipation theorem”. *Journal of Statistical Physics*, **84**, pp. 899–926.
- [21] Gallavotti, G., 2020. “Ensembles, Turbulence and Fluctuation Theorem”. *European Physics Journal, E*, **43:37**.