The role of fluctuations in thermodynamics: a critical answer to Jaworski's paper 'Higher-order moments and the maximum entropy inference'

1988 J. Phys. A: Math. Gen. 21 843

(http://iopscience.iop.org/0305-4470/21/3/041)

J. Phys. A: Math. Gen. 21 (1988) 843-847. Printed in the UK

COMMENT

The role of fluctuations in thermodynamics: a critical answer to Jaworski's paper 'Higher-order moments and the maximum entropy inference'

Giovanni Paladin and Angelo Vulpiani

Dipartimento di Fisica, Università di Roma 'La Sapienza', P le A Moro 2, 00185 Roma, Italy and GNSM-CISM, Roma, Italy

Received 18 May 1987, in final form 7 September 1987

Abstract: We show that energy fluctuations, and thus the higher-order moments of energy, contain essential information and cannot be neglected even in the thermodynamical limit. In contrast with some of Jaworski's statements, we point out how all the thermodynamical properties depend in a crucial way on the fluctuations on the basis of realistic physical assumptions. Indeed, phenomenological thermodynamics allows one to conclude that the Gibbs canonical distribution is the only possible probability distribution which does not violate the second principle. It follows that the knowledge of the free energy as a function of temperature β^{-1} is equivalent to that of the probability law governing fluctuations. This probability law is therefore a characteristic of a body which can be investigated by measuring either energy moments at fixed β or the mean values of entropy and energy at varying β.

About thirty years ago, Jaynes (1957a, b) formulated a principle of maximum entropy as a criterion to pick up the probability distribution which is best suited for a macroscopic description of physical systems. Following this principle the equilibrium statistical mechanics can be interpreted as a special kind of statistical inference. In practice, the result of the maximum entropy procedure depends on the constraints that are imposed. The canonical ensemble, for example, follows by the maximisation of the Shannon entropy (Shannon 1948) fixing the mean energy, the grand canonical ensemble fixing also the mean number of particles. Moreover, the results are determined by the choice of the variables since the Shannon entropy is not invariant under a change of variables. Note that the principle of maximum entropy does not *a priori* establish what the essential information and the good set of variables are.

We do not want to discuss here the Jaynes approach in its general aspects and implications, but just limit ourselves to some considerations originated by a recent paper in this journal. Jaworsky (1987) has discussed the maximum entropy inference employing the higher-order moments of energy. It is well known that maximising entropy with respect to the canonical variables (p, q) with the constraint $\langle H\rangle = U$ one obtains the Gibbs weight $\rho_G \propto \exp(-\beta H)$. It is thus natural to contemplate the possible role of $\langle H^2\rangle$, $\langle H^3\rangle$ and so on. If one in fact knows (e.g., by experimental measurements) the first n successive moments $\langle H^i\rangle = U_i$, then the Shannon entropy is maximised by the generalised Gibbs distribution

$$\rho = \exp\left(-\sum \alpha_i H^i\right) Z_p^{-1} \tag{1}$$

where $Z_\rho(\alpha_1, \alpha_2, \ldots, \alpha_n) = \int \exp(-\Sigma\, \alpha_i H^i)\, dp\, dq$, and the parameters α_i are Lagrange multipliers such that

$$U_i = -\partial \ln(Z_\rho)/\partial \alpha_i. \tag{2}$$

The Shannon entropy of ρ is

$$\sigma = \ln Z_\rho + \sum \alpha_i U_i \tag{3}$$

which satisfies the relations:

$$\alpha_i = \partial \sigma / \partial U_i. \tag{4}$$

We here use unities for which the Boltzmann constant $K_B = 1$, and we have to stress that the Shannon entropy σ in general differs from the statistical entropy S defined as the microcanonical entropy computed at $H = U$. Let us now consider a system composed of N particles in a volume V_N and Hamiltonian H_N. In the case of finite N the generalised Gibbs state is $\rho_N = \exp(-\Sigma\, \alpha_{i,N} H^i_N) Z_\rho^{-1}$ and we shall denote the densities $u_i = \lim_{N\to\infty} U_{i,N}/N$, as well as $s = \lim_{N\to\infty} S_N/N$, while it is useful to define a set of variables β_i which do not vanish in the thermodynamic limit as

$$\alpha_{i,N} = \beta_i / N^{i-1} + \mathrm{O}(N^{-i}) \tag{5}$$

with $\beta_1 = \beta$ if $n = 1$.

Jaworski pointed out that the contribution arising from the knowledge of the higher-order energy moments U_i to statistical entropy is negligible in the thermodynamical limit $N \to \infty$, $N/V_N =$ constant, since in this limit the α_i vanish for $i \neq 1$ and the statistical entropy density $s(\rho)$ of a generalised distribution coincides with the entropy density $s(\rho_G)$ of the Gibbs distribution. For finite N one therefore has:

$$S_N(\rho) = S_N(\rho_G) + \mathrm{o}(N). \tag{6}$$

Jaworski then claims that the information contained in the moments is not essential from a thermodynamic point of view and can be neglected in a maximal entropy inference of the statistical weight because 'it would be difficult to agree that ... all thermodynamic properties depend in some essential way on energy fluctuation'. We do not believe that this is true. The main points of our remark are as follows.

(i) The actual dependence of energy U and statistical entropy S on the temperature T is a trivial consequence of the definition of microcanonical entropy (see (7) below) and does not depend on the detailed form of ρ. This is a simple extension of Jaworski's results.

(ii) The parameters β_i are shown to be uniquely determined assuming the validity of usual thermodynamics. It is in fact possible to show that the only distribution which does not violate the second law of thermodynamics is the Gibbs distribution, and so β_i should be identically zero for $i \neq 1$.

We want to stress that point (ii) disagrees with Jaworski's approach on the basis of physical grounds (essentially the second principle) even if all his results are mathematically correct. In the conclusions of this comment we shall briefly discuss the possibility of introducing a generalised canonical distribution $\rho \neq \rho_G$ satisfying a 'generalised second principle' but only in some particular systems (Grad 1952).

Let us now consider a large number of different samples ('realisations') of the same system with N particles. A sample can of course be regarded just as a 'small' part of a larger system. Each sample is then characterised by its energy per particle, say x, and we can group all the realisations with energy per particle between x and $x + dx$

in the same set $\Omega(x)$. The number H_x of realisations belonging to $\Omega(x)$ is assumed to increase exponentially with N as

$$H_x \propto \exp(s(x)N) \tag{7}$$

where $s(x)$ is the microcanonical entropy per particle corresponding to the energy x.

We can now compute the partition function of the Gibbs distribution integrating over x:

$$Z_N(\beta) = \int H_x \exp(-\beta x N)\,\mathrm{d}x = \int \exp[N(s(x) - \beta x)]\,\mathrm{d}x. \tag{8}$$

In the limit $N \to \infty$ only one set gives a relevant contribution to Z_N. By means of the steepest descent method one in fact obtains

$$Z_N(\beta) = \exp(-\beta f(\beta) N)$$

with

$$\beta f(\beta) = \min_x[\beta x - s(x)] = \beta u - s(u) \tag{9}$$

where the minimal condition in the Legendre transformation (9) implies

$$\left.\frac{\mathrm{d}s}{\mathrm{d}x}\right|_{x=u} = \beta(u). \tag{10}$$

The probability of finding a sample with $x \neq u$ vanishes for large N as

$$P(x) = Z^{-1} \exp(-\beta x N)\, H_x \propto \exp[-\beta N(\psi_\beta(x) - f(\beta))] \tag{11}$$

where $\psi = \beta x - s(x) \geqslant f$ and f is the free energy per particle of the system. We have derived these well known results in order to point out that the hierarchy of sets $\Omega(x)$ can be investigated at varying β, even if they give a negligible contribution to the state function u and s. Indeed, one easily recognises that each temperature selects a particular energy $x = u(\beta)$ which minimises $\psi_\beta(x)$. If one knows *a priori* the statistical weight, it follows that the shape of $P(x)$, or equivalently the values of the moments $\langle x^i \rangle = u_i$, are fully determined by the knowledge of the free energy as function of β, i.e. of $S(\beta)$ and $U(\beta)$. We can repeat all the above arguments with a generical statistical weight $\rho \propto \exp(-\Gamma(u, x)N)$, where in the Jaworski paper the particular form $\Gamma = \Sigma\, \beta_i x^i$ is considered. The result is straightforward as the saddle point estimate of the partition function is given by the x value which minimises $\Gamma(u, x) - s(x)$ and the temperature $T = \beta^{-1}$ is defined by the standard relation (10). Since $s(x)$ is given by (7), $S(T)$ and $U(T)$ as functions of T do not depend on the particular form of Γ. Moreover, since $\langle H^i \rangle = \langle H \rangle^i (1 + \mathrm{o}(N^i))$, one sees that $\sigma(\rho) = \sigma(\rho_{\mathrm{G}}) - \mathrm{o}(N)$.

It is, however, clear that different forms of Γ lead to different fluctuations. Jaworski therefore argues that the actual Γ can be decided only on experimental grounds by measurement of the energy moments even if, in his opinion, the question is rather useless since Γ does not affect the fundamental thermodynamical relations.

Let us note that fluctuations can nevertheless be observed in many interesting circumstances and are a relevant property of a body. This implies that there is only one statistical weight which can reproduce real thermodynamics. Moreover, Szilard proved long ago in his dissertation thesis that the second law of thermodynamics gives information about fluctuations of microscopic parameters and not only on their mean values (Szilard (1925), for a similar approach to foundations of thermodynamics see Mandelbrot (1962, 1964)).

For example, the Gibbs distribution assures that the energy fluctuation variance is related to the specific heat by the standard relation

$$\langle (x-u)^2 \rangle = c_V T^2/N \tag{12}$$

with $c_V = du/dT$. Szilard (1925) proved that only the variance given by formula (12) is compatible with the second law, while for $n=2$ in (1) we have for such an abnormal statistics $\rho \neq \rho_G$, $N\langle (x-u)^2 \rangle < c_V T^2$. If one had a variance which differs from (13), such as the variance of the generalised distribution of Jaworski, then it would be possible to change the average internal energy density of a set of systems described by ρ after thermal contacts with another set at the same temperature described by ρ_G; see Szilard (1925) for details. He in fact noted that even if for a particular cyclic process one can have a finite probability of getting an entropy decrease, the classical postulate of thermodynamics assures one that this is not possible when the process is repeated on a large number of realisations. In other terms, the global entropy cannot decrease and it is therefore not possible to use fluctuations for constructing a perpetuum mobile of the second kind. This assumption which is at the basis of phenomenological thermodynamics is violated for all the statistical distributions except the canonical Gibbs distribution. On the contrary if the energy statistics is given by ρ_G we cannot change it without any compensation. Szilard (1925) has shown that the work necessary for bringing a body from ρ_G to an abnormal statistics ρ is in general given by

$$W \sim T(\sigma(\rho_G) - \sigma(\rho)) \tag{13}$$

where $\sigma(\rho) = -\langle \ln(\rho) \rangle_\rho$ and $\sigma(\rho_G) = S = \max_\rho \sigma(\rho)$. This is the price paid to a 'Maxwell devil' in order to change fluctuation statistics. Let us recall that $\sigma(\rho_G) - \sigma(\rho) = o(N) > 0$. Nevertheless this small difference is sufficient to violate the second law. Indeed if $\rho \neq \rho_G$ were an equilibrium state, then one could bring with a small work W, ρ_G to ρ and so extract an arbitrarily large amount of work MW, performing a number M of cycles with reservoirs with statistics ρ_G at the same temperature as defined by (10).

This indicates the relevance of fluctuations since it points out that two bodies are in thermal equilibrium only if their distributions coincide. Moreover, the Gibbs distribution is the only admissible distribution, assuming that an equilibrium state, at fixed volume and number of particles, is fully characterised by the temperature.

The idea of 'generalised thermodynamics' [Jaworski 1981, 1987] thus is in disagreement with the usual formulation of the second law. Indeed, in the generalised thermodynamics $\delta Q/T$ is not an exact differential and a 'generalised' second law has to be formulated with the new exact differential $\Sigma\, \alpha_i \delta Q_i$ where δQ_i are forms of heat related to U_i and $\delta Q_1 = \delta Q$ [Ingarden and Kossakowski 1965, Jaworski 1981]. It seems to use that in this approach two bodies are in equilibrium if the two sets of all their 'temperatures' β_i coincide, and that the equality of their temperatures T is not sufficient. In this sense the β_i should be relevant, even if they do not affect the state functions S and U.

It is worth stressing that if one assumes the possibility of deriving the statistics of an open system (i.e. the canonical distribution, as usually assumed) from the ergodic hypothesis on a closed system described by the microcanonical measure, then it follows that the statistics of an open system (considered as a small part of the closed one) has to depend only on one parameter. On the contrary, the generalised canonical distribution $\rho \neq \rho_G$ involves different constraints which should imply, in the closed systems, conservation laws beyond the energy conservation. It therefore follows that such a distribution cannot be obtained from the microcanonical measure, which implies one conservation law. The statistical mechanics of particular systems with more than one conservation law was actually constructed by Grad (1952). In these cases the probability distribution has the form $\exp(-\Sigma_i g_i(q, p)/T_i)$ where g_i are the conserved quantities of a closed system and T_i are the corresponding 'temperatures'. This form is similar to (1) but here f_i are given by the dynamics and the whole set of T_i is relevant, since two states are in equilibrium only if the two sets of 'temperatures' coincide (Grad 1952).

Let us also remark that, following Mandelbrot (1962, 1964), it is possible to derive all the phenomenological thermodynamics in terms of mathematical statistics, assuming that the equilibrium state is described by the only probability distribution with a single scalar sufficient statistics. In terms of the maximum entropy formalism, one has to conclude that there is only a relevant parameter, the mean energy, and that fluctuations are not negligible details since they are fully determined by it.

We thank B B Mandelbrot and L Peliti for useful remarks.

References

Grad H 1952 *Commun. Pure Appl. Math.* **5** 455
Ingarden R S and Kossakowski A 1965 *Acta Phys. Pol.* **28** 499
Jaworski W 1981 *Acta Phys. Pol.* A **60** 645
—— 1987 *J. Phys. A: Math. Gen.* **20** 915
Jaynes E T 1957a *Phys. Rev.* **106** 620
—— 1957b *Phys. Rev.* **108** 171
Mandelbrot B B 1962 *Ann. Math. Stat.* **33** 1021
—— 1964 *J. Math. Phys.* **5** 164
Shannon E 1948 *Bell Syst. Tech. J.* **27** 379, 623
Szilard L 1925 *Z. Phys.* **32** 753