Equivalence between maximum entropy principle and enforcing $dU = TdS$

A. Plastino

La Plata National University and Argentina's National Research Council (CONICET), C. C. 727, 1900 La Plata, Argentina

E. M. F. Curado

Centro Brasileiro de Pesquisas Fisicas (CBPF), Rua Xavier Sigaud 150, Urca, Rio de Janeiro, Brasil

(Received 26 April 2005; published 26 October 2005)

We study here a hitherto unexplored microscopic connection between the well-known thermodynamical relation $dU = TdS$ and Jaynes' maximum entropy principle (MaxEnt) for determining probability distributions for the canonical ensemble.

DOI: 10.1103/PhysRevE.72.047103 PACS number(s): 02.50.−r, 03.67.−a, 05.30.−d

I. INTRODUCTION

The first law of thermodynamics is one of physics' most important statements [1]. In statistical mechanics, an underlying *microscopic substratum* is added that is able to explain thermodynamics itself [2–4]. On this substratum, a probability distribution (PD), that controls the population of microstates, is a basic ingredient of statistical mechanics [2–4]. Changes that exclusively affect the microstate-population give rise to "heat" [2–4]. How these changes are related to energy-changes provides an essential content of the first law [2].

In this effort we will work exclusively within a *microscopic*, information theory context. As the starting point of the microscopic considerations that allow one to arrive at the pertinent PD that describes the system of interest, we ask ourselves whether there is an *alternative route* to the maximum entropy approach (MaxEnt) of Jaynes' that would use the $dU = TdS$ component of the *first law* [3] (where we have used the Clausius expression for a reversible process $\delta Q = TdS$, related to the second law), instead of maximizing, à la Jaynes, S with constraints (the energy in the case of the canonical ensemble, addressed here). The work part of the first law is not included in this canonical ensemble derivation due to its subtlety, but we hope to be able to include it in a future work. The idea is: if only the system's level population varies, say from p_i to $p_i + dp_i$, can the $dU = TdS$ relation be determined by itself a unique probability distribution p_i? As point of fact, we will show that: ♣ given a concave information measure S, a mean energy U and a temperature T, and ♣ for any system described by a microscopic probability distribution (PD) $\{p_i\}$, ♣ assuming a reversible heat transfer process via $p_i \to p_i + dp_i$:

(1) If the PD $\{p_i\}$ maximizes S this entails $dU = TdS$, and, alternatively,

(2) If $dU = TdS$, this predetermines a unique PD that maximizes S.

Symbolically, given a specific entropic form or information measure S, the relation $dU = TdS \Leftrightarrow$ MaxEnt in the sense that both sides of the arrow uniquely fix *the same* PD $\{p_i\}$. The transit from (**1**) to (**2**) has been studied, for instance, in [5,6] (by no means an exhaustive list!).

II. OUR CENTRAL ARGUMENT

We shall start the present considerations by assuming that one deals with a rather *general* information measure of the form

$$S = k \sum_i p_i f(p_i), \tag{1}$$

where, for simplicity's sake, Boltzmann's constant is denoted now just by k. The sum runs over a set of quantum numbers, collectively denoted by i (characterizing levels of energy ϵ_i), that specify an appropriate basis in Hilbert space and $\mathcal{P} = \{p_i\}$ is an (as yet unknown) unnormalized probability distribution such that

$$\sum_i p_i = \text{const}, \tag{2}$$

the "constant" being set eventually equal to unity. Often it is preferable, for practical purposes, to postpone normalization until the pertinent computation is finished. Finally, f is an arbitrary smooth function of the p_i, satisfying the condition that $pf(p)$ is concave. Further, we assume that mean values of quantities A that take the value A_i with probability p_i are evaluated according to

$$\langle A \rangle = \sum_i A_i g(p_i), \tag{3}$$

with g another arbitrary smooth, monotonic function of the p_i such that $g(0) = 0$ and $g(1) = 1$. We do not need to require the condition $\Sigma_i g(p_i) = 1$. In particular, the mean energy U is given by

$$U = \sum_i \epsilon_i g(p_i). \tag{4}$$

Assume now that the set $\mathcal{P}$ changes in the fashion

$$p_i \to p_i + dp_i, \quad \text{with} \sum_i dp_i = 0 [\text{cf. } (2)], \tag{5}$$

which in turn generates corresponding changes dS and dU in, respectively, S and U. We are talking just about level-population changes, i.e., heat. We want then to make sure that the heat part, $dU = TdS$ of thermodynamics' first law, where we have used the Clausius relation $\delta Q = TdS$, is obeyed, so that we impose the condition that, in the above described circumstances,

$$dU - TdS = 0, \tag{6}$$

with T the temperature. As a consequence of (6), a little algebra yields, up to first order in the dp_i, the condition

$$\sum_i (\epsilon_i g'(p_i) - kT[f(p_i) + p_i f'(p_i)])dp_i \equiv \sum K_i dp_i = 0, \tag{7}$$

where the primes indicate derivatives with respect to p_i. Equations (5) and (7) should hopefully yield one, "and just one," expression for the p_i. We proceed to show now that all the K_i are equal. Indeed, select just two of the dp's$\neq 0$, say, dp_i *and* dp_j with the remaining $dp_k=0$ for $k\neq j$ and $k\neq i$, which entails, according to Eq. (5), $dp_i=-dp_j$. In these circumstances, for Eq. (7) to hold we necessarily have $K_i=K_j$. But, since i and j have been arbitrarily chosen, *a posteriori* we find $K_i=K$ for all i. The value of K will be determined by the normalization condition (K is, in fact, related to the partition function) on the ensuing probability distribution, to be determined by the relation

$$K = \epsilon_i g'(p_i) - kT[f(p_i) + p_i f'(p_i)] \Rightarrow [f(p_i) + p_i f'(p_i)] - \beta[\epsilon_i g'(p_i) - K] = 0,$$

$$\beta \equiv 1/kT. \tag{8}$$

Alternatively, assume now that you wish to extremize S subject to the constraint of a fixed U, which is achieved via a Lagrange multiplier β,

$$\delta_{\{p_i\}}\left[S - \beta U - \xi \sum_i p_i\right] = 0, \quad \text{i.e.,}$$

$$\delta_{p_m} \sum_i [p_i f(p_i) - \beta g(p_i)\epsilon_i - \xi p_i] = 0, \quad \text{entailing}$$

$$f(p_i) + p_i f'(p_i) - \beta g'(p_i)\epsilon_i - \xi = 0,$$

that, after setting $\xi = -\beta K$, becomes

$$f(p_i) + p_i f'(p_i) - \beta[g'(p_i)\epsilon_i - K] = 0. \tag{9}$$

Since (8) and (9) *are the same equation*, the equivalence between Eqs. (8) and (9) implies that we can simultaneously traverse the directions $(\mathbf{1}) \Leftrightarrow (\mathbf{2})$ mentioned in the Introduction. *The equivalence stated in the Abstract is thus proven*. Let us look now at some examples of pedagogical value that exhibit our ideas at work. For simplicity's sake, we will adjust normalization of the p_i at the end of the calculation.

A. Shannon's entropy

Here we take

$$f(p_i) = -\ln(p_i), \quad \text{and } g(p_i) = p_i. \tag{10}$$

In these circumstances, Eq. (8) becomes

$$-\epsilon_i = kT[\ln(p_i) + 1] - K, \tag{11}$$

which immediately yields [remember (5)]

$$p_i = \exp(-1 + \beta K - \beta\epsilon_i) \tag{12}$$

that after normalization yields the canonical Boltzmann-Gibbs distribution (BD). We conclude that this PD is the only one that guarantees obeyance to the first law for Shannon's entropy. *A posteriori*, one ascertains that the BD maximizes entropy as well, with U as a constraint.

B. Tsallis measure with linear constraints

We have now, for any real number q [7–9],

$$f(p_i) = \frac{(1 - p_i^{q-1})}{q - 1}, \quad \text{and } g(p_i) = p_i, \tag{13}$$

so that $f'(p_i) = -p_i^{q-2}$ and Eq. (8) becomes, with $\beta=(1/kT)$,

$$qp_i^{q-1} = 1 + (q-1)\beta K - (q-1)\beta\epsilon_i, \tag{14}$$

which after normalization yields a distribution often referred to as the Tsallis' one [8]

$$p_i = Z_q^{-1}[1 - (q-1)\beta'\epsilon_i]^{1/(q-1)},$$

$$Z_q = \sum_i [1 - (q-1)\beta'\epsilon_i]^{1/(q-1)}, \tag{15}$$

where $\beta' \equiv \beta/(1+(q-1)\beta K)$.

C. Tsallis measure with nonlinear constraints, normalized

After using non-normalized constraints [10], this is the standard treatment nowadays [7]. It was proposed in [11]. One has

$$g(p_i) = \frac{p_i^q}{w_q}; \quad w_q = \sum_i p_i^q; \quad U_q = \sum_i g(p_i)\epsilon_i, \tag{16}$$

which entails

$$g'(p_i) = \frac{qp_i^{q-1}}{w_q}\left[1 - \frac{p_i^q}{w_q}\right]. \tag{17}$$

This is to be inserted into (8) and one finds

$$(1-q)\beta\epsilon_i g'(p_i) = qp_i^{q-1} - 1 + (1-q)\beta K \tag{18}$$

i.e.,

$$qp_i^{q-1} = \left[1 - (1-q)\beta K + \frac{(1-q)\beta q p_i^{q-1}\epsilon_i}{w_q}\left(1 - \frac{p_i^q}{w_q}\right)\right], \tag{19}$$

or

$$\frac{1 - (1-q)\beta K}{qp_i^{q-1}} = \left[1 - (1-q)\beta\frac{\epsilon_i}{w_q}\left(1 - \frac{p_i^q}{w_q}\right)\right]. \tag{20}$$

If we sum both members of the above equation over the running index i ($i=1,\ldots,\mathcal{N}$, where $\mathcal{N}$ is the number of nondegenerate energy levels i) we get

$$\sum_i \left\{ \frac{1-(1-q)\beta K}{q p_i^{q-1}} - \left[1 - \frac{(1-q)\beta}{w_q}\left(\epsilon_i - \frac{U_q}{\mathcal{N}} \right) \right] \right\} = 0, \tag{21}$$

which can only vanish if each i term vanishes by itself. Tsallis *et al.* [11] argue at this point that one is, of course, free to shift the energy scale so as to add a fixed amount $W = U_q[(1-\mathcal{N})/\mathcal{N}]$ to each ϵ_i. Since the origin of the energy spectrum can always be freely chosen, one can legitimately assume then the *uniform* energy-shift

$$\epsilon_i \mapsto \varepsilon_i; \quad \varepsilon_i = \epsilon_i + U_q \frac{1-\mathcal{N}}{\mathcal{N}}.$$

This prompts one to write the pertinent, properly normalized probability distribution in the Tsallis-Mendes-Plastino (TMP) fashion [11]

$$p_i = Z_q^{-1} \left[1 - \frac{(1-q)\beta}{w_q}(\epsilon_i - U_q) \right]^{1/(1-q)},$$

$$Z_q = \sum_i \left[1 - \frac{(1-q)\beta}{w_q}(\epsilon_i - U_q) \right]^{1/(1-q)}. \tag{22}$$

D. Exponential entropic form

This is given in [6,12] and also used in [13]. One has

$$f(p_i) = \frac{1-\exp(-bp_i)}{p_i} - S_0, \tag{23}$$

where b is a positive constant and $S_0 = 1 - \exp(-b)$, together with

$$g(p_i) = \frac{1-e^{-bp_i}}{S_0} \Rightarrow g'(p_i) = \frac{be^{-bp_i}}{S_0}, \tag{24}$$

which, inserted into (8), after a little algebra, leads to

$$p_i = \frac{1}{b}\left[\ln \frac{b}{S_0 - \beta K} + \ln\left(1 - \frac{\beta\epsilon_i}{S_0} \right) \right], \tag{25}$$

which, after normalization, gives the correct answer [12].

III. CONCLUSIONS

We have endeavored to show in this communication that, from a microscopic perspective, in a process where just $p_i \rightarrow p_i + dp_i$, MaxEnt and $dU = TdS$ co-imply themselves in reciprocal fashion for the canonical ensemble, that is, (1) assuming entropy is maximum (with constraints) one immediately derives $dU = TdS$, and (2) if you assume the validity of $dU = TdS$ and an information measure, this predetermines a probability distribution that maximizes entropy with the internal energy as a constraint. Our demonstration is of quite a general nature. Thus, the entropic form invoked in the two items above is not restricted to be the Boltzmann-Gibbs-Shannon (BGS) one, but to any suitable entropic form. Currently, much interesting work has been performed using alternative forms [7]. Restriction on just how far the BGS form can be extended are discussed, for instance, in [14]. The first item is known (see, for instance, [5,6]), but, as far as we know, the present is the first instance in which the second item has received detailed discussion.

[1] H. B. Callen, *Thermodynamics* (Wiley, New York, 1960); E. A. Desloge, *Thermal Physics* (Holt, Rhinehart, and Winston, New York, 1968).

[2] R. K. Pathria, *Statistical Mechanics* (Pergamon, Exeter, 1993); F. Reif, *Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).

[3] B. H. Lavenda, *Statistical Physics* (Wiley, New York, 1991); B. H. Lavenda, *Thermodynamics of Extremes* (Albion, West Sussex, 1995).

[4] E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957); *Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz (Reidel, Dordrecht, Boston, 1987); A. Katz, *Principles of Statistical Mechanics, The Information Theory Approach* (Freeman and Co., San Francisco, 1967).

[5] R. S. Mendes, Physica A **242**, 299 (1997).

[6] E. M. F. Curado, Braz. J. Phys. **29**, 36 (1999).

[7] *Nonextensive Entropy: Interdisciplinary Applications*, edited by M. Gell-Mann and C. Tsallis (Oxford University Press, Oxford, 2004).

[8] C. Tsallis, Braz. J. Phys. **29**, 1 (1999); A. Plastino and A. R. Plastino, *ibid.* **29**, 50 (1999).

[9] A. R. Plastino and A. Plastino, Phys. Lett. A **177**, 177 (1993).

[10] E. M. F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991).

[11] C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).

[12] E. M. F. Curado and F. D. Nobre, Physica A **335**, 94 (2004).

[13] N. Canosa and R. Rossignoli, Phys. Rev. Lett. **88**, 170401 (2002).

[14] A. R. Plastino and A. Plastino, Phys. Lett. A **226**, 257 (1997).