ELSEVIER

Generating statistical distributions without maximizing the entropy

A. Plastino[a,*], E.M.F. Curado[b]

[a]*Facultad de Ciencias Exactas, Universidad Nacional de La Plata, IFLP-CONICET, Argentine National Research Council, C.C. 727, 1900 La Plata, Argentina*
[b]*Centro Brasileiro de Pesquisas Fisicas (CBPF), Rua Xavier Sigaud 150, Urca, Rio de Janeiro, Brazil*

Abstract

We show here how to use pieces of thermodynamics' first law to generate probability distributions for generalized ensembles when only level-population changes are involved. Such microstate occupation modifications, if properly constrained via first law ingredients, can be associated not exclusively to heat and acquire a more general meaning.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Thermodynamics; Microscopic probability distribution; First law; Second law; MaxEnt

1. Introduction

The first law of thermodynamics is one of physics' most important statements. Together with the second law, the two constitute strong pillars of our understanding of Nature. In statistical mechanics an underlying microscopic substratum is added that is able to explain not only these laws but the whole of thermodynamics itself [1–4], one of whose basic ingredients is a microscopic probability distribution (PD) that controls the population of microstates of the system under consideration [1]. We will be concerned here *only* with changes that affect exclusively microstate-population. The way these changes are related to changes in a system's extensive quantities provides the essential content of the first law [2]. In this effort we show that the above mentioned PD establishes a link between this aspect of the first law, on the one hand, and the Maximum Entropy principle (MaxEnt), on the other, according to the scheme given below.

- Hypothesis: for
- a given a concave entropic form (or information measure) S together with
- (1) a mean internal energy U, (2) mean values $A_v \equiv \langle \mathscr{A}_v \rangle$; $(v = 1, \ldots, M)$ of M additional extensive quantities $\mathscr{A}_v$, and (3) a temperature T,

*Corresponding author.

E-mail addresses: plastino@fisica.unlp.edu.ar (A. Plastino), evaldo@cbpf.br (E.M.F. Curado).

- Thesis: then, for any system described by

  (1) a microscopic PD $\{p_i\}$, and
  (2) assuming a *reversible process via* $p_i \to p_i + \mathrm{d}p_i$,
  (3) one can verify that:

  (1) if the PD $\{p_i\}$ maximizes S this entails $\mathrm{d}U = T\,\mathrm{d}S - \sum_{v=1}^{M} \gamma_v\,\mathrm{d}A_v$, or, alternatively,
  (2) if $\mathrm{d}U = T\,\mathrm{d}S - \sum_{v=1}^{M} \gamma_v\,\mathrm{d}A_v$, this predetermines a unique PD that maximizes S.

It should be remarked that, curiously enough, this uniqueness of the PD does not demand (at this stage) concavity (or convexity) of the entropy with regard to the distribution of probabilities, a requirement that arises a posteriori, in further developing the theory of statistical mechanics [1]. The transit from (1) to (2) has been studied, for instance, in Refs. [5,6] (by no means an exhaustive list!). Succinctly, given a specific S-form,

$$\mathrm{d}U = T\,\mathrm{d}S - \sum_{v=1}^{M} \gamma_v\,\mathrm{d}A_v \;\Leftrightarrow\; \text{MaxEnt prob. distr. } \{p_i\}.$$

2. The proof

Consider 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 just by k. The sum runs over a set of quantum numbers, collectively denoted by i (characterizing levels of energy ε_i), which specify an appropriate basis in Hilbert's space and $\mathscr{P} = \{p_i\}$ is an (as yet unknown) un-normalized PD such that $\sum_i p_i = \text{constant}$, the "constant" being set eventually equal to unity (often it is preferably, for practical purposes, to postpone normalization until the pertinent computation is finished).

Let f be an arbitrary smooth function of the p_i, in such a way $p_i f(p_i)$ is a concave function. Further, consider M quantities A_v that represent mean values of extensive physical quantities $\mathscr{A}_v$. These take, for the state i, the value a_i^v with probability p_i. Also, we suppose that g is 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 $\sum_i g(p_i) = 1$. The mean energy U and the A_v are given by

$$U = \sum_i \varepsilon_i g(p_i), \quad A_v = \sum_i a_i^v g(p_i). \tag{2}$$

Assume now that the probability-set $\mathscr{P}$ changes in the fashion

$$p_i \to p_i + \mathrm{d}p_i \quad \text{with} \quad \sum_i \mathrm{d}p_i = 0 \text{ (normalization!)}, \tag{3}$$

which in turn generates corresponding changes $\mathrm{d}S$, $\mathrm{d}A_v$, and $\mathrm{d}U$ in, respectively, S, A_v, and U.

The essential point that we are introducing in this effort is that we want *to make sure that*, *in the above described circumstances*, *the following condition*, *related to the first law*, *is obeyed*:

$$\mathrm{d}U - T\,\mathrm{d}S + \sum_{v=1}^{M} \mathrm{d}A_v \lambda_v = 0, \tag{4}$$

with T the temperature. As a consequence of (4), a little algebra yields, up to first order in the $\mathrm{d}p_i$, the condition

$$\sum_i [C_i^{(1)} + C_i^{(2)}]\,\mathrm{d}p_i \equiv \sum_i K_i\,\mathrm{d}p_i = 0,$$

$$C_i^{(1)} = \left[\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right] g'(p_i), \quad C_i^{(2)} = -kT\,[f(p_i) + p_i f'(p_i)], \tag{5}$$

where the primes indicate derivative with respect to p_i. Eq. (5) should hopefully yield one and just one expression for the p_i.

2.1. Equality of the K_i in Eq. (5)

We proceed to show now that all the K_i are equal. As the dp_i are linked by the relation

$$\sum_i K_i \, dp_i = 0, \tag{6}$$

we can write, if we are dealing with N micro-states,

$$dp_N = -\sum_{i=1}^{N-1} dp_i. \tag{7}$$

Substituting (7) in (6) we obtain

$$\sum_{i=1}^{N-1} (K_i - K_N)\, dp_i = 0. \tag{8}$$

Now, since the $N-1$ "population-variations" dp_i are independent, this entails that each term in (8) should vanish by itself, which implies that

$$K_i = K_N \quad \text{for all } i = 1, \ldots, N-1. \tag{9}$$

Thus, $K_i =$ constant (in the sense of being independent of i) $= K$ for all i.

2.2. The role of K

Interestingly enough, we do not need to give a specific value to K for our present purposes, although it will become clear below that it is related to the probabilities-normalization constant. We only need to ascertain the K-role, in the following sense. We have, on account of (9),

$$K = C_i^{(1)} + C_i^{(2)} \quad \text{(for any } i),$$

$$C_i^{(1)} = \left[\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right] g'(p_i),$$

$$C_i^{(2)} = -kT[f(p_i) + p_i f'(p_i)], \tag{10}$$

so that, if we redefine things in the fashion

$$T_i^{(1)} = f(p_i) + p_i f'(p_i),$$

$$T_i^{(2)} = -\beta\left[\left(\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right) g'(p_i) - K\right] \quad (\beta \equiv 1/kT), \tag{11}$$

we can recast (10) as

$$T_i^{(1)} + T_i^{(2)} = 0 \quad \text{(for any } i), \tag{12}$$

a relation whose importance will be presently become manifest.

2.3. The MaxEnt route revisited

Assume now that you wish to extremize S subject to the constraints of fixed valued for (i) U and (ii) the M values A_v. This is achieved via Lagrange multipliers (1) β and (2) $M\gamma_v$. We also need a normalization Lagrange multiplier ξ.

$$\delta_{\{p_i\}}\left[S - \beta U - \sum_{v=1}^{M} \gamma_v A_v - \xi \sum_i p_i\right] = 0, \tag{13}$$

leading to, with $\gamma_v = \beta\lambda_v$, to

$$0 = \delta_{p_m} \sum_i \left(p_i f(p_i) - \left[\sum_i \beta g(p_i)\left(\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right) + \xi p_i\right]\right), \tag{14}$$

so that

$$0 = f(p_i) + p_i f'(p_i) - \left[\beta g'(p_i)\left(\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right) + \xi\right] \Rightarrow \quad \text{if } \xi \equiv \beta K,$$

$$0 = f(p_i) + p_i f'(p_i) - \beta\left[g'(p_i)\left(\sum_{v=1}^{M} \lambda_v a_i^v + \varepsilon_i\right) + K\right] \Rightarrow 0 = T_i^{(1)} + T_i^{(2)}. \tag{15}$$

Clearly, (12) and the last equality of (15) are *one and the same equation*! The equivalence stated in the Abstract is thus proven.

3. Discussion

We have here endeavored to show that appropriate manipulation of some ingredients of the first law of thermodynamics can be used to generate the equilibrium microscopic probability distribution (PD) that describes a system within the framework of a generalized ensemble [1], and that such an approach is an alternative to the MaxEnt-one. We were exclusively concerned with changes that affect exclusively microstate-population and, more specifically, with the way these modifications are related to internal variation of the system's extensive quantities [2]. We started with (1) a given concave entropic form (or an information measure (IM)) S, (2) a mean internal energy U, and M mean values $A_v \equiv \langle \mathscr{A}_v \rangle$ ($v = 1, \dots, M$) of M extensive quantities $\mathscr{A}_v$, (3) a temperature T, and demonstrated that, for any system described by a microscopic probability distribution (PD) $\{p_i\}$, assuming a reversible process via $p_i \to p_i + \mathrm{d}p_i$ that is forced to verify the relation $\mathrm{d}U = T\,\mathrm{d}S - \sum_{v=1}^{M} \gamma_v \,\mathrm{d}A_v$, we got an equation that yields a unique PD that maximizes S. By way of contrast, MaxEnt starts from S and, extremizing it with appropriate constraints, allow one to find the system's PD. In other words, $\mathrm{d}U = T\,\mathrm{d}S - \sum_{v=1}^{M} \gamma_v \,\mathrm{d}A_v \Leftrightarrow$ MaxEnt prob. distr. $\{p_i\}$. An alternative route to microscopic PDs, with some first law flavor, has thus been found in the present communication.

References

[1] R.K. Pathria, Statistical Mechanics, Pergamon Press, Exeter, 1993;
    H.B. Callen, Thermodynamics, Wiley, New York, 1960;
    E.A. Desloge, Thermal Physics, Holt, Rhinehart and Winston, New York, 1968.
[2] F. Reif, Statistical and Thermal Physics, McGraw-Hill, New York, 1965.
[3] J.J. Sakurai, Modern Quantum Mechanics, Benjamin, Menlo Park, CA, 1985.
[4] E.T. Jaynes, Phys. Rev. 106 (1957) 620;
    E.T. Jaynes, Phys. Rev. 108 (1957) 171;
    E.T. Jaynes, in: R.D. Rosenkrantz (Ed.), Papers on Probability, Statistics and Statistical Physics, Reidel, Dordrecht, Boston, 1987;
    A. Katz, Principles of Statistical Mechanics, The Information Theory Approach, Freeman and Co., San Francisco, 1967;
    A. Plastino, A.R. Plastino, Braz. J. Phys. 29 (1999) 50.
[5] R.S. Mendes, Physica A 242 (1997) 299.
[6] E.M.F. Curado, Braz. J. Phys. 29 (1999) 36.