

Tsallis Entropy and Jaynes' Information Theory Formalism*

A. Plastino^{1,2} and A. R. Plastino^{1,2}

¹ National University La Plata

C.C. 727, 1900 La Plata, Argentina

² Argentine National Research Council (CONICET)

Received 07 December, 1998

The role of Tsallis' non-extensive Information Measure within an à la Jaynes Information-Theory-based formulation of Statistical Mechanics is discussed in rather detailed fashion.

I Introduction

In spite of its great success, the Statistical Mechanics paradigm based on the Boltzmann-Gibbs entropy measure seems to be inadequate to deal with many interesting physical scenarios [1, 2, 3]. Astronomical self-gravitating systems constitute an important illustrative example of these difficulties [4]. A considerable effort has been devoted by astrophysicists to develop a thermostatistical description of self-gravitating systems along the lines of standard Statistical Mechanics. The failure of those attempts was due to the nonextensivity effects associated with the long range of the gravitational interaction [4].

Ten years ago Tsallis proposed a generalization of the celebrated Boltzmann-Gibbs (BG) entropic measure [5]. The new entropy functional introduced by Tsallis [5] along with its associated generalized thermostatics [6, 7] is nowadays being hailed as the possible basis of a theoretical framework appropriate to deal with nonextensive settings [8, 9, 10]. This entropy has the form

$$S_q = \frac{1}{q-1} \left(1 - \int f(\mathbf{x})^q d\mathbf{x} \right). \quad (1)$$

where \mathbf{x} is a dimensionless state-variable, f corresponds to the probability distribution and the entropic index q is any real number. This entropy recovers the standard Boltzmann-Gibbs entropy $S = -\int f \ln f d\mathbf{x}$ in the limit $q \rightarrow 1$. S_q is *nonextensive* such that

$S_q(A+B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B)$, where A and B are two systems independent in the sense that $f(A+B) = f(A)f(B)$. It is clear that q can be seen as measuring the degree of nonextensivity.

Many relevant mathematical properties of the standard thermostatics are preserved by Tsallis' formalism or admit natural generalizations [8-14]. Tsallis' proposal was shown to be consistent both with Jaynes' Information Theory formulation of statistical mechanics [15], and with the dynamical thermostating approach to statistical ensembles [16].

The recent application of Tsallis' theory to an increasing number of physical problems is beginning to provide a picture of the kind of scenarios where the new formalism is useful. Self-gravitating systems constituted the first physical problem discussed within the nonextensive thermostatics [17]. That early application, in turn, inspired Boghosian's treatment of the two dimensional pure electron plasma, yielding the first experimental confirmation of Tsallis theory [18]. A possible solution of the solar neutrino puzzle based on Tsallis thermostatics has been advanced [19]. Some cosmological implications of Tsallis proposal have also been worked out [20]. The behaviour of dissipative low dimensional chaotic systems [21], as well as self organized critical systems [22] have been discussed in connection with the new approach. Tsallis entropy has also been advanced as the basis of a thermostatistical foundation

* This work is dedicated to the memory of Prof. E. T. Jaynes, who passed away on 30 April 1998.

of Lévy flights and distributions [23]. Tsallis nonextensive statistical formalism proved to be a useful framework for the analysis of many interesting properties of nonlinear Fokker-Planck equations [24-29]. It has been shown that Tsallis maximum entropy (MaxEnt) distributions can also arise naturally as stationary solutions of *linear* Fokker-Planck equations [30].

Tsallis bold attempt to develop a complete thermostatistical formalism on the basis of a nonlogarithmic entropy functional has raised many interesting issues related both to the mathematical structure and physical implications of general thermostatistical formalisms [31, 32]. Tsallis pioneering work has stimulated the exploration of the properties of other generalized or alternative information measures [33, 34]. On the other hand, it has been recently realized that some important features are shared by extended families of thermostatistical formalism [31, 32].

Tsallis' theory can be elegantly formulated in terms of Jaynes' Information Theory (IT) approach to Statistical Mechanics. It is our purpose here that of reviewing this type of formulation, that helps placing Tsallis' thermostatics in an adequate context.

II Basic ideas of Jaynes' IT approach

III The maximum entropy probability distribution

Information Theory (IT) [35] provides one with a powerful inference methodology in order to describe general properties of arbitrary systems on the basis of scarce information. Indeed, it purports to yield the least-biased description that can be devised on the basis of some specific data, in any possible situation. Within a Statistical Mechanics' context Jaynes [36-39] was able to employ IT ideas so as to reformulate and generalize the basic foundations of the field, in what constituted a rather spectacular advance. The essential ideas underlying Jaynes' IT-based methodology [36-39] can best be introduced with reference to the following, quite general environment. Let A_r , ($r = 1, \dots, M$) be a set of (real) random variables that characterize some system \mathcal{S} of interest. These variables adopt the (possible) values $A_r(i)$ with (properly normalized) probabilities $p(i)$,

where $i = 1, \dots, N$ enumerates the possible "states" of \mathcal{S} . Assume now that our "experimental" information concerning \mathcal{S} is limited to the set of mean values ("expectation" values)

$$\langle A_r \rangle \equiv a_r = \sum_i p(i) A_r(i), \quad r = 1, \dots, M, \quad (2)$$

The question to be answered is the following one: what can we assert concerning the (unknown) probability distribution $\{p(i)\}$? As, in general, $M < N$ (indeed, in most realistic situations we have $M \ll N$), many different distributions $\{p(i)\}$ are compatible with the information supply (2). However, IT claims that the BEST (or least-biased) one is precisely that which maximizes the Thermodynamical entropy [36-39]). We see that in order to find this purportedly "one and only" $\{p(i)\}$ we face an extremalization problem of the Lagrange sort, in which one extremalizes a given functional S , subject to a set of constraints, i.e., Eqs. (2) (our "input" or "prior" knowledge) supplemented by the normalization condition

$$\langle A_0 \rangle \equiv a_0 = \sum_i p(i) = 1. \quad (3)$$

As we deal with $M + 1$ constraints we must introduce an equal number of Lagrange multipliers, that we shall denote by $\lambda_0 - 1$, λ_r , ($r = 1, \dots, M$) and freely extremalize

$$Q = - \sum_i p(i) \ln p(i) - (\lambda_0 - 1) \langle A_0 \rangle - \sum_{r=1}^M \lambda_r \langle A_r \rangle, \quad (4)$$

that is, we set

$$\delta_{p(i)} Q = - \sum_i (\delta p(i)) \{ \ln p(i) + \lambda_0 + \sum_{r=1}^M \lambda_r A_r(i) \} = 0, \quad (5)$$

for any arbitrary $\delta p(i)$, which entails that

$$\ln p(i) = -\lambda_0 - \sum_{r=1}^M \lambda_r A_r(i), \quad (6)$$

so that, on the basis of the prior information (Eqs. (2) and (3) we infer the distribution $\{p(i)\}$ of the typical exponential (which guarantees the essential positivity requirement on the distribution) appearance

$$p(i) = \exp\{-\lambda_0 - \sum_{r=1}^M \lambda_r A_r(i)\}. \quad (7)$$

We introduce now the useful abbreviation

$$\text{Tr}f[A_r] \equiv \sum_i f[A_r(i)], \quad (8)$$

for any (analytical) function f of the $A_r(i)$. For example, we have

$$\langle A_r \rangle = \text{Tr}pA_r. \quad (9)$$

We define now the “partition” function

$$Z = \text{Tr} \exp \left[- \sum_{r=1}^M \lambda_r A_r \right], \quad (10)$$

which, on account of (3) leads immediately to

$$Z = \exp(\lambda_0), \quad (11)$$

i.e.,

$$\lambda_0 = \ln Z, \quad (12)$$

a pair of relationships that will be frequently encountered herefrom. The first derivative of (12) yields

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_s} &= \frac{\partial \ln Z}{\partial \lambda_s} \\ &= -Z \text{Tr} \{ A_s \exp[-\sum_r \lambda_r A_r] \} \\ &= -\text{Tr}[pA_s] \\ &= -\langle A_s \rangle, \end{aligned} \quad (13)$$

which, when properly interpreted, solves our variational problem, as we now proceed to show. Notice that, in (8), λ_r , $r = 1, \dots, M$ are the only variables. Consequently, Z is a function of these M Lagrange multipliers. So is then λ_0 , in view of (12)

$$\lambda_0 \equiv \lambda_0(\lambda_1, \dots, \lambda_M). \quad (14)$$

We should then regard (14) as a set of (coupled) highly non-linear equations in the M variables λ_r , with the input-information (the $\langle A_r \rangle$) on the r.h.s. When solved, this system provides us with Z and the Maximum Entropy $\{p(i)\}$.

III The main properties of the Maximum Entropy Probability (MEP) distribution

Our maximum entropy (or maximum ignorance) acquires the aspect

$$S = - \sum_i p(i) \ln p(i) = \sum_{r=0}^M \lambda_r \langle A_r \rangle, \quad (15)$$

On the basis of (12) and (15) we readily ascertain that λ_0 and S are related by means of a quite general Legendre transform [36, 37, 38, 39]

$$\lambda_0 \equiv \ln Z = S - \sum_{r=1}^M \lambda_r \langle A_r \rangle, \quad (16)$$

which clearly tells us that, as Z is a function of the Legendre multipliers, S must be a function of the mean values. This mathematical result is consistent with Shannon's interpretation of S as a missing information function, that is, S measures our ignorance once the $\langle A_r \rangle$ are given. In addition to (14) we write then

$$S = S(a_1, \dots, a_M). \quad (17)$$

We can define also a generalized “free energy” by selecting a special Lagrange multiplier, say λ_s , and writing

$$F_s = -\ln Z / \lambda_s \quad (18)$$

$$= \langle A_s \rangle - \sum_{r \neq s} \frac{-\lambda_r}{\lambda_s} \langle A_r \rangle - \frac{S}{\lambda_s} \quad (19)$$

$$= \langle A_s \rangle - \sum_{r \neq s} \mu_r \langle A_r \rangle - T_s S, \quad (20)$$

with

$$\mu_r = \frac{-\lambda_r}{\lambda_s}; \quad T_s = \lambda_s^{-1}, \quad (21)$$

positive, maximizing S is tantamount to minimizing F_s . From (15) we also obtain

$$\frac{\partial S}{\partial \langle A_r \rangle} = \lambda_r, \quad (22)$$

the relation “conjugate” to (14). See also that

$$\frac{\partial \langle A_v \rangle}{\partial \lambda_s} = - \frac{\partial^2 \ln Z}{\partial \lambda_v \partial \lambda_s} = \frac{\partial \langle A_s \rangle}{\partial \lambda_v}, \quad (23)$$

and that

$$\frac{\partial p(i)}{\partial \lambda_s} = p(i) [\langle A_s \rangle - A_s(i)], \quad (24)$$

which leads to

$$\begin{aligned} \frac{\partial \langle A_r \rangle}{\partial \lambda_s} &= \frac{\partial [\sum_i p(i) A_r(i)]}{\partial \lambda_s} \\ &= \langle A_r \rangle \langle A_s \rangle - \langle A_r A_s \rangle \\ &= -\langle (A_r - \langle A_r \rangle)(A_s - \langle A_s \rangle) \rangle \end{aligned} \quad (25)$$

We can also write

$$\begin{aligned} \frac{\partial S}{\partial \lambda_s} &= \sum_r \lambda_r \frac{\partial \langle A_r \rangle}{\partial \lambda_s} \\ &= - \sum_r \lambda_r [\langle A_r A_s \rangle - \langle A_r \rangle \langle A_s \rangle] \end{aligned} \quad (26)$$

which, in the particular case of having $M = 1$ specializes to

$$\frac{\partial S}{\partial \lambda_1} = -\lambda_1 [\langle A_1^2 \rangle - (\langle A_1 \rangle)^2], \quad (27)$$

so that the derivative of the entropy with respect to λ_1 gives immediately the dispersion.

III The modified Kinchin axiomatics

One of the salient contributions of Information Theory (IT) is that of yielding a recipe for ascertaining in precise and unambiguous terms the amount of information (*the information measure*) that an observer possesses concerning a given phenomenon when only a probability distribution (PD) is known. The informational content of a normalized probability distribution $P(i)$, ($i = 1, \dots, N$), where the subindex i runs over the states of the system one is trying to study, is given by Shannon's information measure (IM) [35]

$$S = - \sum_{i=1}^N P(i) \ln[P(i)], \quad (28)$$

where the choice of the logarithm basis is used to fix the informational units. If the basis is 2 then S is measured in *bits*.

In a more formal vein one is led to consider Kinchin Axioms [40] as providing the conceptual foundations of Information Theory. Consider a system Σ composed of two subsystems (Σ^1, Σ^2). Let P_i^m be a PD associated to subsystem Σ^m ($m = 1, 2$). The PD corresponding to the total system is labelled by two subindexes i, j , one for each of the subsystems. In general, the two subsystems will be correlated, so that one needs the conditional probability $Q(j|i)$ of finding Σ^2 in state j when one is sure that the state of Σ^1 is that labelled by i , and a concomitant *conditional* IM, expressed in terms of $Q(j|i)$. Kinchin axioms read

1) For a system Σ described by a PD $P(i)$, $i = 1, \dots, N$, the IM is a function only of the $P(i)$

$$S\{P\} = S(P(1), \dots, P(N)). \quad (29)$$

2) For such a system $S\{P\} \leq S\{\text{uniform PD}\}$, where the uniform PD is, of course, $P_{\text{uniform}} = \frac{1}{N}$, for all i .

3) Suppose that, instead of dealing with N states we confront $N+1$ ones, with the proviso that $P(N+1) = 0$. Then S does not change.

4) Let Σ be composed of two subsystems, as explained above. Then

$$S(\Sigma) = S(\Sigma^1) + \sum_i P_i^1 S\{Q|i\}. \quad (30)$$

These four axioms lead in univocal fashion to Shannon's IM (28). To most people, the first three axioms appear self-evident. However, the last one does not seem to enjoy the same status. One may think that a more *natural* phrasing of the fourth axiom would be

$$4') \quad S\{P\} = S\{P^1\} + S\{P^2\}, \quad (31)$$

but the ensuing, modified set of axioms leads not to just one but to *two* IM's. One of them is Shannon's. The other reads

$$S_{R,\beta} = (1 - \beta)^{-1} \ln \left(\sum_{i=1}^N P_i^\beta \right); \quad \beta \in \Re, \quad (32)$$

which is known as Rény's IM and has found extensive applications in connection with fractals and Cantor sets [41].

Jaynes [36] has shown that if one chooses Boltzmann's constant as the informational unit and identifies Shannon's IM with the Thermodynamical entropy, then the whole of Statistical Mechanics can be elegantly reformulated, without any reference to the notion of *ensemble*, by extremalization of Shannon's S , subject to the constraints posed by the *a priori* information one may possess concerning the system of interest (the *Maximum Entropy Principle (MEP)*) [37, 38, 39]. Rény's IM [42] cannot be regarded as a physical entropy, as it does not have a definite concavity when expressed as a function of the pertinent $P(i)$.

Can we think of still an alternative version of the fourth postulate that will yield an IM of definite concavity different from Shannon's one? If we advance the following axiom [7]

$$4'')$$

$$S\{P\} = S\{P^1\} + S\{P^2\} + (1 - q)S\{P^1\}S\{P^2\}; \quad q \in \mathfrak{R}, \quad (33)$$

one is led to Tsallis's entropy [1, 2, 3, 7]

$$S_q = (q - 1)^{-1} \sum_{i=1}^N [P(i) - P(i)^q], \quad (34)$$

and is related to Rény's IM $S_q(R)$ in the following fashion

$$S_q(R) = (1 - q)^{-1} \ln \left(1 + (1 - q) S_q \right). \quad (35)$$

IV Tsallis' generalized statistical mechanics

Within a classical Gibbsian context, Tsallis [1, 2, 3, 7] showed that his entropy leads to a Generalized Statistical Mechanics (GSM). Consider a system Σ with M possible microscopic configurations and let $\{p_i\}$ stand for the probability of finding the system in the configuration i . As stated above, the associated Tsallis' IM, to be regarded herefrom as a physical entropy, reads

$$S_q = \sum_{i=1}^M (p_i - p_i^q) (q - 1)^{-1}, \quad (36)$$

with q a real parameter (we have a different statistics for every possible q -value) and

$$\sum_i p_i = 1. \quad (37)$$

In order to study the limit $q \rightarrow 1$ we write

$$S_q = \sum_i p_i \left[\{1 - e^{(q-1)} \ln p_i\} (q - 1)^{-1} \right], \quad (38)$$

and find that for $q \rightarrow 1$

$$S_1 \equiv \lim_{q \rightarrow 1} S_q = -k \sum_i p_i \ln p_i, \quad (39)$$

i.e., for $q = 1$ Tsallis' entropy coincides with the Gibbs-Shannon one.

From its definition $S_q \geq 0$. S_q vanishes (for all q) in the case $M = 1$, and, for $M > 1$, $q > 0$, whenever one of the p_i equals unity and the remaining ones, of course, vanish. A global, absolute maximum of S_q (for all q) obtains, according to the modified Kinchin's axioms, in the case of *equiprobability*, when all $p_i = 1/M$. In such an instance we have

$$S_q = k[M^{1-q} - 1](1 - q)^{-1}, \quad (40)$$

that, in the limit $q \rightarrow 1$ leads to the celebrated Boltzmann expression

$$S_1 = k \ln M. \quad (41)$$

Tsallis' entropy exhibits a series of notable properties that reinforce the idea that S_q is indeed a physical quantity. We list some of them below.

IV.1 Concavity

Let us consider two PD's $\{p_i^a\}$ and $\{p_i^b\}$, where i labels the members of a set of M microstates. For a real λ such that $0 < \lambda < 1$ we define an "intermediate" distribution $\{p_i\}$ by recourse to

$$p_i \equiv \lambda p_i^a + (1 - \lambda) p_i^b. \quad (42)$$

One easily verifies that

$$\begin{aligned} q > 0 &\rightarrow S_q[p] \geq \lambda S_q[p^a] + (1 - \lambda) S_q[p^b] \\ q < 0 &\rightarrow S_q[p] \leq \lambda S_q[p^a] + (1 - \lambda) S_q[p^b] \end{aligned} \quad (43)$$

The functional $S_q[p_i]$ is *concave* for $q > 0$ and *convex* for $q < 0$ (S_q being constant ($= M - 1$) for $q = 0$).

IV .2 Pseudo-additivity

Consider two independent systems A and B characterized by possessing M_a and M_b microstates, respectively and assume that the corresponding PD's are

$$\begin{aligned} A &\longrightarrow \{p_1^A, p_2^A, \dots, p_{M_a}^A\} \\ B &\longrightarrow \{p_1^B, p_2^B, \dots, p_{M_b}^B\} \end{aligned} \quad (44)$$

The total, composite system $A \cup B$ (of microstates given by all possible pairs of A - and de B -microstates) is described by the PD

$$p_{ij}^{A \cup B} = p_i^A p_j^B, \quad (45)$$

and one easily finds that its associated entropy is

$$S_q^{A \cup B} = S_q^A + S_q^B + (1 - q) S_q^A S_q^B. \quad (46)$$

As a consequence we have

$$\begin{aligned} S_q^{A \cup B} &< S_q^A + S_q^B, \quad (q > 1) \\ S_q^{A \cup B} &> S_q^A + S_q^B, \quad (q < 1) \\ S_q^{A \cup B} &= \underbrace{S_q^A + S_q^B}_{\text{extensivity}}, \quad (q = 1), \end{aligned} \quad (47)$$

so that, except for $q = 1$, Tsallis' entropy is a *non-extensive* quantity, this being its main difference vis-a-vis the orthodox one.

IV .3 Canonical ensemble

Tsallis found that by extremalization of S_q under the constraints posed by both normalization and assumed knowledge of the internal energy, that is

$$\sum_{i=1}^M p_i = 1, \quad (48)$$

$$\sum_{i=1}^M p_i \varepsilon_i = U, \quad (49)$$

one obtains the generalized canonical distribution [15]

$$p_i^* = \frac{1}{Z_q} [1 - \beta(q - 1)\varepsilon_i]^{1/(q-1)}, \quad (50)$$

where

$$Z_q^* = [1 - \beta(q - 1)\varepsilon_i]^{1/(q-1)}, \quad (51)$$

is the generalized partition function.

However, more interesting results obtain if one introduces as constraint the *generalized internal energy* [2]

$$U_q = \sum_{i=1}^M p_i^q \varepsilon_i, \quad (52)$$

which leads to the PD

$$p_i = \frac{1}{Z_q} [1 - \beta(1 - q)\varepsilon_i]^{1/(1-q)}, \quad (53)$$

with

$$Z_q = [1 - \beta(1 - q)\varepsilon_i]^{1/(1-q)}. \quad (54)$$

Curado and Tsallis [6] found that the whole mathematical (Legendre-transform based) structure of thermodynamics becomes in this fashion *invariant* under a change of the q -value (from unity to any other real number). Indeed, one finds, for example, relations of the form

$$-\frac{\partial}{\partial \beta} [Z_q^{1-q} - 1](1 - q)^{-1} = U_q, \quad (55)$$

$$\frac{\partial}{\partial U_q} \left(\frac{S_q}{k} \right) = \beta. \quad (56)$$

$$\left(\frac{Z_q^{1-q} - 1}{1 - q} \right) + \beta U_q = S_q, \quad (57)$$

identical to their well-known $q = 1$ -counterparts if one replaces $\ln Z$ by

$$\frac{Z_q^{1-q} - 1}{1 - q}. \quad (58)$$

One immediately realizes that

$$(S_q) [U_q], \quad \text{and} \quad \left(\frac{Z_q^{1-q} - 1}{1 - q} \right) [\beta], \quad (59)$$

are related by a Legendre transform.

V Generalized Entropies and Information Theory

Plastino and Plastino [15] have generalized the work of ref. [6] by 1) embedding it within a purely quantal (Hilbert space) context and 2) using Jaynes' approach to SM, which allows one to deal with (the IT equivalent of) *any* ensemble, accomodating both equilibrium and off-equilibrium situations on an equal footing.

The first step in that direction is, of course, the construction of a statistical operator (or density operator, or density matrix) $\hat{\rho}$ able to account for *all* the available information, on the one hand, and that maximizes Tsallis' entropy, on the other one. The available (a priori or prior) information can (in general) be casted in the form of a set of expectation values (EV). Here we need, however, *generalized* EV's, in the spirit of Curado and Tsallis [6] (see above the definition of a generalized internal energy). We assume prior knowledge of M EV's, corresponding to M operators \hat{O}_i ,

$$\langle \hat{O}_i \rangle_q = Tr(\hat{\rho}^q \hat{O}_i), \quad i = 1, \dots, M, \quad (60)$$

where, we insist, generalized EV's are being employed, according to the definition

$$\langle \hat{O}_i \rangle_q = \langle \hat{\rho}^{q-1} \hat{O}_i \rangle, \quad (61)$$

with an ordinary EV on the r.h.s. Of course, normalization entails

$$Tr \hat{\rho} = 1. \quad (62)$$

After a bit of algebra, recourse to that Lagrange multipliers method provides us with the normalized density operator that reproduces the M known (generalized) EV's (Cf Eq. (60)) and maximizes Tsallis' entropy. One finds [7]

$$\hat{\rho} = Z^{-1} \left(1 - (1-q) \sum_i \lambda_i \hat{O}_i \right)^{1/(1-q)}, \quad (63)$$

where Z is the partition function

$$Z = Tr \left(1 - (1-q) \sum_i \lambda_i \hat{O}_i \right)^{1/(1-q)}, \quad (64)$$

and we have M Lagrange multipliers λ_i that guarantee compliance with the M EV-related constraints. However, a small difficulty remains. The density operator is a definite positive one and, as it stands, this is not guaranteed by Eq. (63). Consequently, we must require that the operator

$$\hat{A} = 1 - (1-q) \sum_i \lambda_i \hat{O}_i, \quad (65)$$

appearing between the parenthesis in (63) be a positive definite one. This entails that the eigenvalues of \hat{A} must be non-negative quantities. An ad-hoc requirement (to be justified below) is then to be introduced at this point. An heuristic cut-off is needed. Instead of (63) we write

$$\hat{\rho} = Z^{-1} \left[\hat{A} \Theta(\hat{A}) \right]^{1/(1-q)}, \quad (66)$$

with Z given by

$$Z = Tr \left[\hat{A} \Theta(\hat{A}) \right]^{1/(1-q)}, \quad (67)$$

and $\Theta(x)$ the step (Heaviside's) function

$$\Theta(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x > 0 \end{cases}. \quad (68)$$

Equations (66-67) should be interpreted as follows. Let $|i\rangle$ and α_i , respectively, the eigenvectors and eigenvalues of the operator (66), so that (spectral decomposition)

$$\hat{A} = \sum_i \alpha_i |i\rangle \langle i|. \quad (69)$$

In this special basis $\hat{\rho}$ adopts the appearance

$$\hat{\rho} = Z^{-1} \sum_i f(\alpha_i) |i\rangle \langle i|, \quad (70)$$

with $f(x)$ defined according to

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x^{1/(1-q)} & \text{for } x > 0. \end{cases}. \quad (71)$$

Using now the shorthand notation

$$\hat{B}' = \hat{B} \Theta(\hat{B}), \quad (72)$$

and

$$\hat{B} = \hat{I} - (1-q) \sum_i \lambda_i \hat{O}_i, \quad (73)$$

the generalized entropy S_q is given by

$$\begin{aligned}
S_q &= \frac{1}{(q-1)} \text{Tr} \left(\hat{\rho}^q \left[\hat{\rho}^{1-q} - \hat{I} \right] \right) \\
&= \frac{1}{(q-1)} \text{Tr} \left(\hat{\rho}^q \left[Z^{q-1} \hat{B}' - \hat{I} \right] \right) \\
&= \frac{Z^{q-1}}{(q-1)} \text{Tr} \left(\hat{\rho}^q \hat{B}' \right) - \frac{\text{Tr}(\hat{\rho}^q)}{(q-1)}. \quad (74)
\end{aligned}$$

$$\hat{\rho}^q \left(\hat{I} - (1-q) \sum_i \lambda_i \hat{O}_i \right) \Theta \left(\hat{I} - (1-q) \sum_i \lambda_i \hat{O}_i \right) = \hat{\rho}^q \left(\hat{I} - (1-q) \sum_i \lambda_i \hat{O}_i \right), \quad (75)$$

where, of course, all negative eigenvalues (of \hat{A}) have been conveniently dropped (our cut-off). With a clear conscience we can now write

$$\begin{aligned}
S_q &= \frac{Z^{q-1} - 1}{q-1} \text{Tr}(\hat{\rho}^q) + Z^{q-1} \sum_i \lambda_i \text{Tr}(\hat{\rho}^q \hat{O}_i) \\
&= \frac{Z^{q-1} - 1}{q-1} \text{Tr}(\hat{\rho}^q) + Z^{q-1} \sum_i \lambda_i \langle \hat{O}_i \rangle_q. \quad (76)
\end{aligned}$$

Now, from the very definition of the generalized entropy S_q we have,

$$\text{Tr}(\hat{\rho}^q) = 1 + (1-q)S_q, \quad (77)$$

so that (76) and (77) yield the *Generalized Euler's Theorem* [15]

$$\begin{aligned}
\frac{\partial \lambda_J}{\partial \lambda_i} &= Z^{-q} \frac{\partial Z}{\partial \lambda_i} \\
&= \frac{Z^{-q}}{1-q} \text{Tr} \left\{ \left[\hat{B}' \right]^{q/(1-q)} \frac{\partial}{\partial \lambda_i} \left(1 - (1-q) \sum_i \lambda_i \hat{O}_i \right) \right\} \\
&= -Z^{-q} \text{Tr} \left\{ \left[\hat{B}' \right]^{q/(1-q)} \hat{O}_i \right\}. \quad (80)
\end{aligned}$$

which leads to

$$\begin{aligned}
\frac{\partial \lambda_J}{\partial \lambda_i} &= -\text{Tr}(\hat{\rho}^q \hat{O}_i) \\
&= -\langle \hat{O}_i \rangle_q, \quad (81)
\end{aligned}$$

that, together with Euler's theorem, tell us that

Obviously, the operators $\hat{\rho}$ and \hat{A} commute. Thus, their product can be expressed in the common basis that diagonalizes them. Assuming this has been done, a bit of contemplative reflection should convince one that

$$S_q = \lambda_J + \sum_i \lambda_i \langle \hat{O}_i \rangle_q, \quad (78)$$

where the *Jaynes parameter* λ_J is given by

$$\lambda_J = \frac{Z(1-q) - 1}{1-q}. \quad (79)$$

The parameter (79) plays, within this generalized context, the role of the logarithm of the partition function in the orthodox SM.

Generalized EV's $\langle \hat{O}_i \rangle_q$, Jaynes parameter λ_J , and the Lagrange multipliers λ_i obey certain strictures that constitute the heart of a thermodynamical description. Partial derivation of λ_J with respect to the λ_i ($i = 1, \dots, M$) yields (consider here that our primed operator is $\hat{B}' = 1 - (1-q) \sum_i \lambda_i \hat{O}_i$)

$$\begin{aligned} \frac{\partial S}{\partial \langle \hat{O}_i \rangle_q} &= \sum_{j=1}^M \left\{ \frac{\partial \lambda_j}{\partial \lambda_j} \frac{\partial \lambda_j}{\partial \langle \hat{O}_i \rangle_q} + \frac{\partial \lambda_j}{\partial \langle \hat{O}_i \rangle_q} \langle \hat{O}_j \rangle_q \right\} + \lambda_i \\ &= \sum_{j=1}^M \left\{ -\langle \hat{O}_j \rangle_q \frac{\partial \lambda_j}{\partial \langle \hat{O}_i \rangle_q} + \frac{\partial \lambda_j}{\partial \langle \hat{O}_i \rangle_q} \langle \hat{O}_j \rangle_q \right\} + \lambda_i, \end{aligned} \quad (82)$$

and allows for the *very* important result [15]

$$\frac{\partial S}{\partial \langle \hat{O}_i \rangle_q} = \lambda_i. \quad (83)$$

Equations (81) and (83) constitute the basic IT relations in order to build up quantum SM à la Jaynes. In deriving them we reach the result that the whole of quantum SM is invariant under a change of q (from unity to any other real number).

The generalized EV's can be shown to obey an Ehrenfest's theorem [15]. Consider a density operator $\hat{D}(t)$ (not necessarily of the maximum entropy form) that evolves (in time) according to Von Neumann's equation

$$\frac{d\hat{D}}{dt} = \frac{1}{i\hbar} [\hat{H}, \hat{D}], \quad (84)$$

where \hat{H} is the system's hamiltonian. Let $|\phi_i(t)\rangle$ and α_i be, respectively, the eigenvectors and eigenvalues of \hat{D} . According to (84) the latter do not depend upon the time while the $|\phi_i(t)\rangle$ are solutions of Schroedinger's equation

$$i\hbar \frac{d}{dt} |\phi_i(t)\rangle = \hat{H} |\phi_i(t)\rangle. \quad (85)$$

From the time-independent nature of the α_i one gathers that if

$$\hat{D} = \sum_i \alpha_i |\phi_i(t)\rangle \langle \phi_i(t)|, \quad (86)$$

is a solution of Von Neumann's equation, another such solution is given by

$$\hat{D}^q = \sum_i \alpha_i^q |\phi_i(t)\rangle \langle \phi_i(t)|. \quad (87)$$

Thus, if \hat{D}^q fulfills Von Neumann's strictures, the generalized EV's $\langle \hat{O} \rangle_q$ will necessarily verify Ehrenfest's theorem [15]

$$\frac{d}{dt} \langle \hat{O} \rangle_q = \frac{1}{i\hbar} \langle [\hat{O}, \hat{H}] \rangle_q. \quad (88)$$

VI Justifying Tsallis' formalism

Plastino and Plastino [43] have justified the GSM discussed above with reference to an argument similar to that employed by Gibbs himself in deriving his canonical ensemble. The idea is to go back to Gibbs' *microcanonical* ensemble (GME).

Consider a system \mathcal{S} with energy levels denoted by ε_i , weakly interacting with a thermal bath B and assume one describes the "total" system $T = \mathcal{S} + B$ by recourse to the GME when its total energy E lies in the interval

$$E_0 - \Delta < E < E_0 + \Delta, \quad (89)$$

with

$$\Delta \ll E_0. \quad (90)$$

As usual, the energy spectrum of the bath B is regarded as being of a quasi-continuous character. Plastino and Plastino [43] traverse a new road, however, in assuming that B is a *finite* system, of finite energy E_b .

As the total system T is microcanonically described, the probability p_i of finding \mathcal{S} in a state $||i\rangle$ of energy ε_i is proportional to the total number ν of T -configurations compatible with such a situation. In view of the quasicontinuous character of the B -energy spectrum, ν will be given by

$$\nu = 2\Delta\eta(E_0 - \varepsilon_i), \quad (91)$$

where $\eta(E)$ represents the number of states (per unit energy interval) of B in a neighbourhood of E . Thus,

$$\frac{p_i}{p_j} = \frac{\eta(E_0 - \varepsilon_i)}{\eta(E_0 - \varepsilon_j)}. \quad (92)$$

Let us assume that the number of states $M(E)$ of B with energy smaller (or equal) than E grows as a power α of E . Such a growth-law is often encountered. As examples we mention

- a) A set of N independent harmonic oscillators ($\alpha = N$),
- b) A set of N free (nonrelativistic) particles confined in a D -dimensional box ($\alpha = DN/2$),
- c) A set of N plane, rigid rotators ($\alpha = N/2$).

With this last assumption we find

$$\eta(E) \propto E^{\alpha-1}, \quad (93)$$

because $\eta(E)$ is essentially the derivative of $M(E)$ with respect to E . Thus

$$\begin{aligned} \frac{p_i}{p_j} &= \frac{(E_0 - \varepsilon_i)^{\alpha-1}}{(E_0 - \varepsilon_j)^{\alpha-1}} \\ &= \frac{(1 - \varepsilon_i/E_0)^{\alpha-1}}{(1 - \varepsilon_j/E_0)^{\alpha-1}}, \end{aligned} \quad (94)$$

so that, after multiplication by the convenient normalization factor Z^{-1} , with

$$Z = \sum_j \left(1 - \frac{\varepsilon_j}{E_0}\right)^{\alpha-1}, \quad (95)$$

we arrive at

$$p_i = Z^{-1} \left(1 - \frac{\varepsilon_i}{E_0}\right)^{\alpha-1}. \quad (96)$$

Setting

$$q = \frac{(\alpha - 2)}{(\alpha - 1)}, \quad (97)$$

and

$$\beta = \frac{(\alpha - 1)}{E_0}, \quad (98)$$

we obtain Tsallis' canonical distribution

$$p_i = Z_q^{-1} [1 - \beta(1 - q)\varepsilon_i]^{1/(1-q)}, \quad (99)$$

with

$$Z_q = \sum_i [1 - \beta(1 - q)\varepsilon_i]^{1/(1-q)}, \quad (100)$$

q being, of course, Tsallis' characteristic parameter. In the limit $q \rightarrow 1$ one recovers Gibbs' conventional expressions

$$p_i = \frac{1}{Z_1} e^{-\beta\varepsilon_i}, \quad (101)$$

$$Z_1 = \sum_i e^{-\beta\varepsilon_i}. \quad (102)$$

The physical meaning of the $q \rightarrow 1$ limit deserves special attention. So as to fix ideas let us consider that our thermal bath consists of N independent Harmonic oscillators ($\alpha = N$). Eqs. (97)-(98) give

$$q = \frac{(N - 2)}{(N - 1)}, \quad (103)$$

and

$$\beta = \frac{(N - 1)}{E_0}. \quad (104)$$

The limit $q \rightarrow 1$ corresponds to that situation characterized by $N \rightarrow \infty$ and $E_0 \rightarrow \infty$, the process proceeding in such a fashion as to keep constant the energy per oscillator $W = E_0/N$. Consequently, *Tsallis' generalized canonical distribution* describes a systema in thermal contact with a *finite reservoir*. Stricto sensu, infinte baths do not exist in nature, so that, in some sense, Tsallis' distribution can be regarded as the *natural* one, Gibbs' being, instead, a convenient mathematical "idealization".

The interpretation given in [43] to Tsallis GSM allows one to conclude:

1.- The values adopted by Tsallis' parameter q are determined by the nature of the appropriate thermal bath

2.- The cut-off ad-hoc condition needed so as to determine Tsallis' statistical operator appears here in a natural fashion. The probability p_i associated to the microstate $\|i\rangle$ vanishes whenever

$$\beta(1 - q)\varepsilon_i \geq 1, \quad (105)$$

which is equivalent to the condition

$$\varepsilon_i \geq E_0. \quad (106)$$

Obviously, (106) implies $p_i = 0$: the energy of the system S cannot exceed E_0 , that of the total system $T = \mathcal{S} + B$.

References

- [1] C. Tsallis, *Physica A* **221**, 277 (1995).
- [2] C. Tsallis, *Chaos, Solitons, and Fractals* **6**, 539 (1995), and references therein; an updated bibliography can be found in <http://tsallis.cat.cbpf.br/biblio.htm>
- [3] C. Tsallis, *Physics World* **10**, 42 (July 1997).
- [4] A.R. Plastino and A. Plastino, in *Condensed Matter Theories*, Volume **11**, E. Ludeña (Ed.), Nova Science Publishers, New York, USA, p. 341 (1996).
- [5] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).
- [6] E.M.F. Curado and C. Tsallis, *J. Phys.* **A24**, L69 (1991); Corrigenda: **24**, 3187 (1991) and **25**, 1019 (1992).
- [7] A.R. Plastino and A. Plastino, in *Condensed Matter Theories*, Volume **11**, E. Ludeña (Ed.), Nova Science Publishers, New York, USA, p. 327 (1996).
- [8] D.H. Zanette and P.A. Alemany, *Phys. Rev. Lett.* **75**, 366 (1995); M.O. Caceres and C.E. Budde, *ibid* **77**, 2589 (1996); D.H. Zanette and P.A. Alemany, *ibid* **77**, 2590 (1996).
- [9] A.K. Rajagopal, *Phys.Rev.Lett.* **76**, 3469 (1996).
- [10] E.K. Lenzi, L.C. Malacarne and R.S. Mendes, *Phys. Rev. Lett.* **80**, 218 (1998).
- [11] A. R. Plastino, A. Plastino and C. Tsallis, *J. Phys. A*, **27**, 5707, (1994).
- [12] G. A. Raggio, *J. Math. Phys.* **36**, 4785 (1995).
- [13] A. M. Meson and F. Vericat, *J. Math. Phys.* **37**, 4480 (1996).
- [14] R. J. V. dos Santos, *J. Math. Phys.* **38**, 4104 (1997).
- [15] A. R. Plastino and A. Plastino, *Phys. Lett. A* **177**, 177 (1993).
- [16] A.R. Plastino and C. Anteneodo, *Ann. Phys.* **255**, 250 (1997).
- [17] A. R. Plastino and A. Plastino, *Phys. Lett. A* **174**, 384 (1993); J.J. Aly, in "N-Body Problems and Gravitational Dynamics", Proc. of the Meeting held at Aussois-France (21-25 March 1993), eds. F. Combes and E. Athanassoula (Publications de l'Observatoire de Paris, 1993).
- [18] B. M. R. Boghosian, *Phys. Rev. E* **53**, 4754 (1995); C. Anteneodo and C. Tsallis, *J. Mol. Liq.* **71**, 255 (1997).
- [19] G. Kaniadakis, A. Lavagno and P. Quarati, *Phys. Lett. B* **369**, 308 (1996); P. Quarati, A. Carbone, G. Gervino, G. Kaniadakis, A. Lavagno and E. Miraldi, *Nucl. Phys. A* **621**, 345c (1996).
- [20] V.H. Hamity and D.E. Barraco, *Phys. Rev. Lett.* **76**, 4664 (1996); D.F. Torres, H. Vucetich and A. Plastino, *Phys. Rev. Lett.* **79** 1588 (1997).
- [21] C. Tsallis, A.R. Plastino and W.-M. Zheng, *Chaos, Solitons and Fractals* **8**, 885 (1997); M. L. Lyra and C. Tsallis, *Phys. Rev. Lett.* **80**, 53 (1998).
- [22] F.A. Tamarit, S.A. Cannas and C. Tsallis, *European Phys. J. B* (1998), in press; A.R.R. Papa and C. Tsallis, *Imitation Games: Power Law Sensitivity to Initial Conditions and Nonextensivity*, *Phys. Rev. E*, (1998), in press.
- [23] P. A. Alemany and D. H. Zanette, *Phys. Rev. E* **49**, 956 (1994); C. Tsallis, S.V.F. Levy, A.M.C. Souza and R. Maynard, *Phys. Rev. Lett.* **75**, 3589 (1995); Erratum: **77**, 5442 (1996).
- [24] A. R. Plastino and A. Plastino, *Physica A* **222**, 347 (1995).
- [25] C. Tsallis and D. J. Bukman, *Phys. Rev. E* **54**, R2197 (1996).
- [26] L. Borland, *Phys. Rev. E* **57**, 6634 (1998).
- [27] D.A. Stariolo, *Phys. Rev. E* **55**, 4806 (1997).
- [28] A. Compte and D. Jou, *J. Phys. A* **29**, 4321 (1996).
- [29] A. Compte, D. Jou, and Y. Katayama, *J. Phys. A* **30**, 1023 (1997).
- [30] L. Borland, *Ito-Langevin Equation within Generalized Thermostatistics*, *Phys. Lett A* (1998), in press.
- [31] A. Plastino and A.R. Plastino, *Phys. Lett. A* **226**, 257 (1997).
- [32] R.S. Mendes, *Physica A* **242**, 299 (1997).
- [33] Andrés R.R. Papa, *J. Phys. A: Math. Gen.* **31**, 5271 (1998).
- [34] S. Abe, *Phys. Lett. A* **224**, 326 (1997).
- [35] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana (1948)
- [36] Jaynes E.T., *Phys. Rev.* **106**, 620 (1957); **118**, 171 (1961).
- [37] E. T. Jaynes in *Papers on probability, statistics and statistical physics*, edited by R. D. Rosenkrantz (Reidel, Dordrecht, Holland, 1983).
- [38] Brillouin, L., *Science and Information Theory*, Academic Press, New York (1956).
- [39] W. T. Grandy, Jr., and P. W. Milonni, *Physics and probability: Essays in honor of E. T. Jaynes* (Cambridge University Press, Cambridge, England, 1993).
- [40] A. I. Kinchin, *Mathematical Foundations of Information Theory*, Dover Publ., New York (1957).
- [41] C. Beck and F. Schögl, *Thermodynamics of chaotic systems* (Cambridge University Press, Cambridge, England, 1993).
- [42] A. Rényi, *Probability theory* (North-Holland, Amsterdam, 1970).
- [43] A. R. Plastino and A. Plastino, *Phys. Lett. A* **193**, 251 (1994).