Non-Extensive Entropy from Incomplete Knowledge of Shannon Entropy?

F. Sattin

Consorzio RFX, Associazione EURATOM-Enea per la fusione, Corso Stati Uniti 4, 35127 Padova – Italy

Received August 17, 2004; accepted November 4, 2004

PACS numbers: 05.70.Ln, 05.40.−a, 47.27.−i

Abstract

In this paper we give an interpretation of Tsallis’ nonextensive statistical mechanics based upon the information-theoretic point of view of Luzzi et al. [cond-mat/0306217; cond-mat/0306247; cond-mat/0307125], suggesting Tsallis entropy to be not a fundamental concept but a derived one, stemming from an incomplete knowledge of the system, not taking properly into account its interaction with the environment. This interpretation seems to avoid some problems occurring with the original interpretation of Tsallis statistics.

Nonextensive statistical mechanics (NESM) introduced by Tsallis [1] has been raising considerable interest for its ability in interpreting several experimental results (for a review of the applications of Tsallis formalism, see the updated website: http://tsallis.cat.cbpf.br/biblio.htm).

The standard way of introducing NESM is through the definition of a generalized entropy functional:

\[ S_q = \frac{1}{q - 1} \sum_i p_i^q \]  

(1)

from which a whole generalized thermodynamics can be recovered (see, e.g., [2]).

With the usual procedure of constrained entropy maximization, one can deduce from (1) the equilibrium Probability Distribution Function (PDF) for any given system. Indeed, it is possible to show that the set of probabilities \( p_i \) maximizing \( S_q \) within the canonical ensemble are

\[ p_i = \frac{1}{Z_q} \left[ 1 - \beta (1 - q) \xi_i \right]^{(1/q) - 1} \]  

(2)

with \( \xi_i \) denoting the energy of the \( i \)-th microstate and the Lagrange multiplier \( \beta \) that plays the role of generalized inverse temperature. This expression enjoys so widespread consideration that Tsallis’ NESM has become often synonymous for power-law PDFs. Actually, there is a huge and ever-increasing number of systems where PDFs of some quantity are measured and found to be consistent with power laws (2).

It is important to point out that Eq. (1) is a sufficient condition for (2) to hold, but by no means a necessary one. There are other different definitions for the entropy which lead to the same extremizing \( p_i \)-distribution (e.g., the Rényi entropy). Even more important, PDFs of the type (2) can be recovered without any explicit recourse to a definition of entropy whatsoever: they can be seen as the equilibrium or quasi-equilibrium distributions for systems interacting with thermal baths subject to fluctuations. This fact was first empirically found for a particular system by Wilk and Włodarczyk [3], then Beck [4] provided a theoretical interpretation (see also Sattin and Salasnich [5]), pointing out that, in principle, fluctuations of the thermal bath could lead to several classes of PDFs, of which Eq. (2) represents only one particular case. Finally, Beck and Cohen [6] enframed all these findings within the unifying concept of superstatistics.

A matter of debate is, therefore, the role that must be assigned to the generalized entropy \( S_q \) (Eq. (1)); does it have a physical meaning just like standard Shannon entropy has in thermostatistics, or is it just a useful compact expression to derive non-exponential PDFs? Indeed, some authors [7] have pointed out that \( \xi_i \) should have a meaning just within an information-theoretic approach, as an informational entropy, in situations where the observer cannot have access to a complete characterization of the system (i.e., the states of the system cannot be properly characterized, for example systems with fractal structure, or with long range interactions).

In this work we will provide some considerations about this question, following essentially the path of Luzzi et al. [7]. We do not pretend to give a definite answer here; in particular we must admit the severe constraint that our calculations hold only for the nonextensive exponent close to unity. This covers many of the cases treated in literature, but there are several exceptions. However, we will provide the reader with some comments and analytical calculations supporting Luzzi et al.’s viewpoint.

We will try, hence, to give Eq. (1) a different, information-theoretic interpretation: Tsallis entropy is considered as a partial information entropy, arising from incomplete information over the whole system under consideration. The starting point is the empirical observation of the existence of PDFs of the kind (2) (or, more generally, non-Gaussian), together with the explanation of their existence as due to the interaction of the system \( \Sigma \) with a fluctuating thermal bath \( B \). From this, expressions of the form (1) can be recovered, but they do not express any fundamental property of the system. Instead, they are simply approximate expressions arising when we have an incomplete knowledge of the whole system \( \Sigma + B \). We remind that this fact is perhaps implicit in Beck’s superstatistics formalism, and was recently verified through numerical experiments by Potgieter and Costa [8]. These authors showed that PDFs (2) necessarily arise when our system (endowed with ordinary Maxwell-Boltzmann statistics) interacts with a thermal bath made of a finite number of degrees of freedom, hence fluctuating in its macroscopic quantities (e.g., temperature).

In the following we will use the unnormalized form of Tsallis statistics. Nowadays, there is evidence in support of a normalized form, where mean quantities (observables) must be divided by \( \sum_i p_i \). However, we do not expect that the use of one form rather than another may affect the present results.

Let us consider the (Shannon) entropy of the total system \( \Sigma + B \):

\[ S = - \sum_{ij} h_i(j) \ln[h_i(j)] \]  

(3)
where $b_i$ is the probability for the bath $B$ to be in the $i$-th state, and $\rho(x, j)$ is the conditional probability for the system $\Sigma$ to be in the state $j$ if $B$ is in the state $i$. If $\Sigma + B$ is an isolated system, its equilibrium distribution can be found by extremizing (3) with the normalization conditions $\sum_i b_i \rho(x, j) = 1$, that is

$$\delta F = \delta \left( -\sum_i b_i \rho(x, j) \right) = 0 \Rightarrow b_i \rho(x, j) = \text{const.} \quad (4)$$

Note that this does not imply $x = \text{const}$, that is $\Sigma + B$ follows microcanonical statistical mechanics, not $\Sigma$.

Let us now suppose that there is a large (almost but not exactly unitary) probability for the bath to be just in one single state, say $I$; thus write

$$b_i \sim (1 - \delta_{i,1})(1 - 2\delta_{i,1}). \quad (5)$$

In the expression above $(1 - 2\delta_{i,1})$ is just a small parameter, useful for perturbative expansions. Its meaning is that of an effective width of the fluctuations around the most probable state $I$. If we insert Eq. (5) into (3) we can write

$$S \sim -\sum_i \rho(x, j) \ln[\rho(x, j)] \sim (1 - 2\delta_{i,1}) \sum_{\text{dis}} d_i \rho(x, j) \ln[\rho(x, j)] \quad (6)$$

and the functional $F$ defined in (4) is rewritten,

$$F \sim -\sum_i \rho(x, j) \ln[\rho(x, j)] \sim (1 - 2\delta_{i,1}) \sum_{\text{dis}} d_i \rho(x, j) \ln[\rho(x, j)] \times \ln[d_i \rho(x, j)] \quad (7)$$

The first term in the r.h.s of (6) would be the ordinary entropy for equilibrium distribution can be found by extremizing (3) with the normalization condition $\sum_i b_i \rho(x, j) = 1$, that is

$$\delta F = \delta \left( -\sum_i b_i \rho(x, j) \right) = 0 \Rightarrow b_i \rho(x, j) = \text{const.} \quad (4)$$

Eqs. (7) and (11) can be identified provided the following relations hold:

$$\sum_j \rho(x, j) \ln[\rho(x, j)] = \sum_j \rho(x, j) \ln[\rho(x, j)],$$

$$\sum_j \rho(x, j) \ln[\rho(x, j)] = \sum_j \rho(x, j) \ln[\rho(x, j)].$$

Note that this does not imply $x = \text{const}$, that is $\Sigma + B$ follow microcanonical statistical mechanics, not $\Sigma$.

Let us now imagine a researcher wishing to determine the microcanonical-ensemble functional $F_{\Sigma}$, within the present framework, the question itself is overwhemingly probable. In conclusion, the appearance of non-orthodox entropy appears not to be related to some deficiencies of the principles of physics, but to the fact that we are trying to get correct results using the wrong functional ($F_{\Sigma} - \text{Eq. (9)}$ – instead of $F - \text{Eq. (7)}$).

We have shown that it is possible to identify the non-extensivity parameter $q$ with $x$. But the latter (or, better, $1 - 2\delta_{i,1}$) is simply a measure of the fluctuations of the heat bath around its most probable state. Therefore $q$ by no means an intrinsic property of $\Sigma$ and it is completely determined by $B$. If no fluctuations are possible, $x = 1$ and the system $\Sigma$ is endowed with Shannon entropy. This can happen either because $\Sigma + B$ do not interact at all, so each of them is described within the microcanonical ensemble; or because $B$ has a huge number of degrees of freedom (thermodynamic limit) so that its most probable state is overwhelmingly probable.

We point out some consequences of this interpretation. Let us consider two systems $\Sigma_1$ and $\Sigma_2$, modelled through non-ordinary statistics with different non-extensivity exponents (i.e. $q_1 \neq q_2$ and both $\neq 1$). One question sometimes found in literature is: which is the statistics of the combined system? There does not appear to be a simple way of writing an expression of the kind (1) from two subsystems which, individually, are described by two different non-extensivity exponents. Indeed, there are several papers by Wang and other authors [9] where this issue is debated, without reaching a satisfactory answer. We may mention also the paper [10], where some steps in this direction were made, but without conclusive results (strictly speaking, this paper states that zeroth-law of thermodynamics applies to quasi-equilibrium systems with long range interactions. This is not the same as saying that it holds within nonextensive statistical mechanics, although some connections may undoubtedly be seen). We see instead that now, within the present framework, the question itself becomes ill-posed. In fact, either $\Sigma_1$ or $\Sigma_2$ are in contact with the same heat bath, and in this case they are forced to have the same $q$, or they are in contact with two different heat baths $B_1$, $B_2$. But in this latter case the two independent systems to be considered are $(\Sigma_1 + B_1)$, $(\Sigma_2 + B_2)$. These composite systems follow ordinary extensive statistics, hence ordinary addition of independent entropies applies: It is not possible to have two systems interacting with different $q$ indices.

$\Box$ Physica Scripta 2005
A related issue is: let us suppose to put into thermal contact the two heat baths $B_1$, $B_2$, isolated from the outside. We have therefore the whole system $S = B_1 + B_2$, and $B_1$, $B_2$ could be alternatively regarded as the heat bath or the system studied. Let us suppose that a system $S$ be endowed with $q_1$, $q_2$, statistics if it were separately interacting with $B_1$, $B_2$. Then, there appears an apparent contradiction within our scheme: in fact, if we chose $B_1$ as heat bath, and $S = B_2$, then $B_2$ should follow $q_1$-statistics, while, in the opposite case it is $B_1$ to follow $q_1$-statistics. But is this a real contradiction? Not at all. In fact, it is necessary to consider the distinction between thermal bath and system: from an operational point of view, they are distinct by the fact that the former acts on the latter, but no feedback is possible; the state— and the statistics— of $B$ are not appreciably influenced by the presence of $S$. We must therefore ask how $B_1$, $B_2$ relate to each other. If, quite contrary to the above reasoning, one of them were not of macroscopical size with respect to the other, we are still in the situation of one heat bath ($B_1$) and a system ($B_2$). Previous equations are not symmetric with respect to the interchange bath-system: no conclusion about the statistics of $B_1$ can be drawn, indeed, its statistics is given as known. Therefore, we simply expect the dynamics of $B_2$ to be distorted to the extent that now it follows $q_1$-statistics. Let us suppose, instead, that $B_1$, $B_2$ are of comparable magnitude and each able to appreciably affect the dynamics of the other. In this case there is no reason to pretend that the unperturbed statistics $q_1$, $q_2$, still hold. Instead, the whole system will follow an “intermediate” statistics. We can still use the heat bath-system picture, but now each of the two systems will be endowed with the same statistics. The intermediate statistics will not likely even be Tsallis-like. This result is important since it is exactly what has been found in the investigation [11], prior to this study, which is therefore an independent confirmation. Also, a different proof has been found in the investigation [11], prior to this study, which provide a system ruled by Tsallis statistics, with the “energy” $E$ as independent variable:

$$ p(E) = \frac{1}{Z} \left( 1 + \beta q(q - 1)E^{q-1} \right) $$

(A1)

We identify three regimes, depending upon the value of $E$: I) first, the “low-energy” regime, $\beta q(q - 1)E^{q-1} \leq 1$. This region is scarcely interesting from our point of view: in this range, regardless of its precise dependence on $E$, $p(E)$ can be expanded into a Taylor series truncated to first order, hence no real information about the true shape of the PDF can be obtained. II) Then, there is the usual “intermediate” range, $\beta q(q - 1)E^{q-1} \leq 1$, where the precise analytical form (exponential, stretched exponential, power-law, etc...) for $p(E)$ is looked for. III) Any physical system is, however, finite, hence there must be an extreme-value (very high $E$) region, where $p(E)$ falls sharply to zero. We will stay away also from the region affected by these “finite-size” effects, and consider only the “intermediate” region ($E$).

Under this hypothesis, we can discard the unity term in the denominator of (A1) and write

$$ p(E)dE \approx \frac{1}{Z} \left( 1 + \beta q(q - 1)E^{q-1} \right) dE $$

$$ = \frac{1}{Z} \beta q(q - 1) \exp \left[ -X \left( \frac{1}{q - 1} \right) \right] dX $$

$$ \equiv C q \exp \left[ -BX \right] dX $$

(A2)

where we have set $\beta E = \exp(X)$, and $B = 1/(q - 1) - 1$. 　

Eq. (A2) admits a straightforward interpretation: at least in a restricted energy region (but it is the most important region for this kind of studies), the probability distribution of a system ruled by Tsallis statistics can be mapped into that of a system ruled by ordinary Maxwell-Gibbs statistics, in thermal equilibrium with its
surroundings, with “energy” $X$ and “temperature” $1/B$. Notice, however, that writing $p(E)$ in the form (A2) is not necessary to our purposes, just slightly simplifies the following computations.

Let us now move to the next step, and consider two systems $S_1, S_2$ parameterized with the nonextensivity indices $q_1, q_2$ respectively. In order to give an intuitive picture of the problem, we may think of each system as a box containing a large number of particles, respectively $N_1$ and $N_2$. Each particle is given a value: its “energy” $E$. Apart from the individual $E$ label, particles are identical. The statistical distribution of $E$ follows Tsallis statistics, with the corresponding value for $q$, in each box. In other words, the probability of randomly picking up a particle labelled with energy $E$ from box $i$ ($i = 1, 2$) is given by Eq. (A1) with $q = q_i, q_2$.

The attempt of describing the combined system $(S_1 + S_2)$ in terms of a single distribution is tantamount to pouring all the particles into a single box. The resulting combined statistics will come out by randomly picking a large number of particles out of this box. Let us ask how much the probability is of extracting from the urn a particle labelled with energy $E$. The answer is, trivially:

$$p(S_1 + S_2, E) = \frac{N_1}{N_1 + N_2} p(S_1, E) + \frac{N_2}{N_1 + N_2} p(S_2, E). \quad (A3)$$

In this expression, $N_i/(N_1 + N_2)$ is the probability of choosing a particle out of those previously in the box $i$, and $p(S_i, E)$ is the probability that this particle had energy $E$. Since we are interested in the high-energy tail of the distribution, we can replace $p(S_i, E)$ by its form (A2), which yields:

$$p(S_1 + S_2, X) = \exp(-B_1 X - \mu_1) + \exp(-B_2 X - \mu_2) \quad (A4)$$

where $\mu_i = -\ln(N_i)/(N_1 + N_2)$ is a sort of chemical potential.

The l.h.s. of (A4) must be, in turn, a Tsallis distribution, hence, we get the final result

$$\exp(-B_1 X - \mu_{1,\infty}) = \exp(-B_1 X - \mu_1) + \exp(-B_2 X - \mu_2). \quad (A5)$$

It is obvious that Eq. (A5) cannot be fulfilled for arbitrary $X$, given generic $B_1, \mu_1$, the exception being, as expected, the equal-$q$ case, for which $B_1 = B_2, \mu_1 = \mu_2 = 1/\ln(2), \mu_{1,\infty}$. This concludes our proof.

Other interesting consequences can be drawn from Eq. (A4). In order to fix ideas, let us set $B_1 < B_2$. As long as $\mu_1 > \mu_2$, does not hold, the former exponential on the r.h.s of (A4) will prevail over the latter, hence the asymptotic behaviour of the combined distribution will essentially match $q_2$-statistics. On the other hand, widely differing values between $\mu_1$ and $\mu_2$ can locally override this trend and give birth to an extremely complicated behaviour. Since different $\mu$’s can be made roughly corresponding to hugely different numbers of particles (although there is also a $q$-dependent contribution from the partition function, $C_q$), this statement is equivalent to saying anything but that the overall statistics of the system is driven by the larger in size of the two subsystems.

Some final comments are: a paper [13] has recently appeared, in which the possibility of defining a unique effective temperature for non-equilibrium critical systems is criticized. This systems are exactly between those thought amenable to study using nonextensive statistics. It is also remarkable that the impossibility of defining a common temperature, in that work, is directly related to the existence of fluctuations away from the Gaussian limit.

Secondly, within Beck’s superstatistics interpretation of NESM [14], previous results do appear fairly reasonable: since, there, the entropic index of a system is a measure of the fluctuations of some generalized temperature, it seems difficult to relate within one and the same formulation fluctuations from two completely uncorrelated systems.

References
1. Tsallis, C., J. Stat. Phys. 82, 479 (1998).
2. Tsallis, C. and Sona, A. M. C., Phys. Rev. E67, 026106 (2003).
3. Wilk, G. and Włodarczyk, Z., Phys. Rev. Lett. 84, 2770 (2000).
4. Beck, C., Phys. Rev. Lett. 87, 100601 (2001).
5. Sattin, F. and Salasnich, L., Phys. Rev. E 65, 035106 (2002).
6. Beck, C. and Cohen, E. G. D., Physica A 322, 267 (2003).
7. Luzzi, R., Vasconcellos, ´A. R. and Galvão Ramos, J., cond-mat/0306217.
8. Vasconcellos, ´A. R., Galvão Ramos, J. and Luzzi, R., cond-mat/0306247.
9. Luzzi, R., Vasconcellos, ´A. R. and Galvão Ramos, J., cond-mat, 0307325.
10. Rapisarda, F. Q. and Costa, U. M. S., Physica A 321, 482 (2003); ibid 342, 165 (2004).
11. Wang, Q. A., cond-mat/0311058; Wang, Q. A. et al., Chaos, Solitons & Fractals, to appear [cond-mat/0310390]; Wang, Q. A. et al., Europhys. Lett. 68, 606 (2004); Suuki, K. and Hotta, M., cond-mat/9912454; Suuki, K. and Hotta, M., cond-mat/0004158.
12. Moyano, L. G., Baldwin, F. and Tsallis, C., cond-mat/0305091.
13. Sattin, F. J., Phys. A: Math. Gen. 36, 1583 (2003).
14. Tsallis, C., in “Noneextensive statistical mechanics and its applications” (Abe, S. and Okamoto, Y., eds) (Springer 2003).
15. Calabrese, P. and Gambassi, A., J. Stat. Mech.: Theor. Exp. P07013 (2004).