A Shannon-Tsallis transformation

E. Rufeil Fiori[1]

Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba Ciudad Universitaria, 5000 Córdoba, Argentina

A. Plastino[2,3]

[2]*National University La Plata & CONICET IFLP-CCT, C.C. 727 - 1900 La Plata, Argentina*
[3] *Universitat de les Illes Balears and IFISC-CSIC, 07122 Palma de Mallorca, Spain*

Abstract

We determine a general link between two different solutions of the MaxEnt variational problem, namely, the ones that correspond to using either Shannon's or Tsallis' entropies in the concomitant variational problem. It is shown that the two variations lead to equivalent solutions that take different appearances but contain the same information. These solutions are linked by our transformation.

Key words: Shannon entropy, Tsallis entropy, MaxEnt.
PACS: 89.70.Cf, 05.90.+m, 89.75.Da, 89.75.-k

1 Introduction

Nonextensive statistical mechanics (NEXT) [1,2,3], a generalization of the orthodox Boltzmann-Gibbs (BG) one, is actively investigated and applied in many areas of scientific endeavor. NEXT is based on a nonadditive (though extensive [4]) entropic information measure, that is characterized by a real index q (with q = 1 recovering the standard BG entropy). It has been used with regards to variegated systems such as cold atoms in dissipative optical lattices [5], dusty plasmas [6], trapped ions [7], spinglasses [8], turbulence in the heliosheath [9], self-organized criticality [10], high-energy experiments at LHC/CMS/CERN [11] and RHIC/PHENIX/Brookhaven [12], low-

---

dimensional dissipative maps [13], finance [14], galaxies [15], Fokker-Planck equation's applications [16], etc.

A typical NEXT feature is that it can be expressed by recourse to generalizations à la q of standard mathematical concepts [17]. Included are, for instance, the logarithm [q-logarithm] and exponential functions (usually denoted as $e_q(x)$, with $e_{q=1}(x) = e^x$), addition and multiplication, Fourier transform (FT) and the Central Limit Theorem (CLT) [18]. The q-Fourier transform F_q exhibits the nice property of transforming q-Gaussians into q-Gaussians [18]. Recently, plane waves, and the representation of the Dirac delta into plane waves have been also generalized [19,20,21,22].

Our central interest here resides in the q-exponential function, regarded as the MaxEnt variational solution [3] if the pertinent information measure is Tsallis' one. We will show that there is a transform procedure that converts any Shannon-MaxEnt solution [23,24] into a q-exponential, without modification of the associated Lagrange multipliers, that carry with them all the physics of the problem at hand. Why? Because of the Legendre transform properties of the MaxEnt solutions [see for instance [23,24,25,26,27,28,29,30] and references therein].

Accordingly, we are here proving that the physics of a given problem can be discussed in equivalent fashion by recourse to either Shannon's measure or Tsallis' one, indistinctly.

2   The central idea

We wish to connect orthodox exponentials with q-exponentials. Let us consider the Shannon-MaxEnt solution for a constraint given by the average value of the variable in question, that we call u:

$$p(u)du = \exp\left[-\mu - \lambda u\right]du = e^{-\mu}\,e^{-\lambda u}du,$$
$$\langle u \rangle = K,$$
$$\int p(u)du = 1, \tag{1}$$

where the two Lagrange multipliers μ and λ correspond, respectively, to normalization and conservation of the $u-$mean value $< u >= K$.

Consider now a second variable x such that $dx/du = g(x)$, with $g(x)$ a function we will wish to determine below. Assume that in the second variable we express the Tsallis-MaxEnt solution, with the same constraints, but employing the above mentioned q-exponential functions $(e_q(x) = [1 + (1-q)x]^{1/(1-q)}$. The

support of this function is sometimes finite, depending on the $q-$value. See more details in, for instance, [2]).

$$p(x)dx = C\,e_q(-\lambda\,x)\,dx;\ \ C = \text{normalization const.} \tag{2}$$

We want
$$p(x)dx = p(u)du.$$

This entails:
$$C\,e_q(-\lambda\,x)\,dx = p(x)dx = e^{-\mu}\,e^{-\lambda u(x)}(dx/g(x)), \tag{3}$$

that is, given that $du = g(x)^{-1}\,dx$,
$$C\,g(x)\,e_q(-\lambda\,x) = e^{-\mu}\,\exp{[-\lambda\int\,dx\,g(x)^{-1}]}. \tag{4}$$

Remember now that
$$\frac{d\,e_q(x)}{dx} = e_q(x)^q, \tag{5}$$

so that, taking the logarithm of Eq. (4) we find
$$\ln C + \ln g(x) + \ln{[e_q(-\lambda\,x)]} = -\mu - \lambda\int dx\,g(x)^{-1}. \tag{6}$$

Now, we derive w.r.t. x Eq. (6) and have
$$\frac{g'(x)}{g(x)} - \lambda\,e_q(-\lambda\,x)^{q-1} = -\lambda\,g(x)^{-1}, \tag{7}$$

which leads to a differential equation for our desired $g(x)$:
$$g'(x) - \lambda\,e_q(-\lambda\,x)^{q-1}\,g(x) + \lambda = 0. \tag{8}$$

Solving this equation we establish the link we are looking for.

3   The differential equation

For simplicity we now set
$$P(x) = -\lambda\,e_q(-\lambda\,x)^{q-1};\ \ Q(x) = -\lambda, \tag{9}$$

and recast (8) as

$$g'(x) + P(x) g(x) = Q(x). \tag{10}$$

Introduce now the integrating factor I

$$I = \exp\left[\int dt\, P(t)\right], \tag{11}$$

and multiply (10) by it

$$g'(x)I + P(x)I\, g(x) = Q(x)I. \tag{12}$$

Note that

$$\frac{d}{dx}[g(x)I] = g'(x)I + P(x)I\, g(x), \tag{13}$$

so that (11) becomes

$$\frac{d}{dx}[g(x)I] = Q(x)I. \tag{14}$$

Integrating this we have now

$$g(x)I = Q(x)\int I dx + c. \tag{15}$$

Finally, we can formally express our "solution" function $g(x)$ as

$$g(x) = \frac{\int Q(x)I dx + c}{I}. \tag{16}$$

Thus, for the differential equation

$$g'(x) - \lambda e_q(-\lambda x)^{(q-1)}g(x) + \lambda = 0, \tag{17}$$

we obtain the solution

$$g(x) = \exp\left(\lambda \int^x e_q(-\lambda x')^{(q-1)}dx'\right),$$

$$\left[-\lambda \int^x \exp\left(-\lambda \int^{x'} e_q(-\lambda x'')^{(q-1)}dx''\right)dx' + c\right], \tag{18}$$

4

where c is an integration constant. Now, using

$$\lambda \int^x e_q(-\lambda x')^{(q-1)} \mathrm{d}x' = \ln\left(e_q(-\lambda x)^{(-1)}\right),$$ (19)

and

$$-\lambda \int^x \exp\left(-\lambda \int^{x'} e_q(-\lambda x'')^{(q-1)} \mathrm{d}x''\right) \mathrm{d}x' = \frac{e_q(-\lambda x)^{(2-q)}}{2-q},$$ (20)

we finally arrive at

$$g(x) = e_q(-\lambda x)^{(-1)}\left[\frac{e_q(-\lambda x)^{(2-q)}}{2-q} + c\right],$$ (21)

which provides us with the Tsallis-Shannon Jacobian

$$J(x) = \frac{1}{g(x)}.$$ (22)

Consider now the instance $q \to 1$. One obviously ought to have $g(x) = 1$. We face

$$g(x) \to e^{\lambda x}\left[e^{-\lambda x} + c\right]$$ (23)
$$= 1 + c\,e^{\lambda x},$$ (24)

which entails $c = 0$ and

$$g(x) = \frac{e_q(-\lambda x)^{(1-q)}}{2-q} = \frac{1 - (1-q)\lambda x}{2-q}.$$ (25)

3.1 Expansion near q=1

Let us now take $q = 1 - \epsilon$ in the solution $g(x) = g(x; q)$ (Eq.(25)),

$$g(x; q = 1 - \epsilon) = \frac{1 - \epsilon\lambda x}{1 + \epsilon}.$$ (26)

Thus, a first-order expansion in ϵ gives:

5

$$g(x; q = 1 - \epsilon) = 1 - (1 + \lambda x)\,\epsilon + O(\epsilon^2). \tag{27}$$

3.2 The case $q = 2$

Now, let us look at the $q = 2 - \epsilon$ scenario for our solution $g(x) = g(x; q)$;

$$g(x; q = 2 - \epsilon) = \frac{1 + (1 - \epsilon)\lambda x}{\epsilon}. \tag{28}$$

That, in the limit of $\epsilon \to 0$ we have

$$g(x; q = 2 - \epsilon) \asymp \frac{(1 + \lambda x)}{\epsilon}, \tag{29}$$
$$\to +\infty \;\; \text{as} \;\; \epsilon \to 0^+ \;\; (\text{i.e.}, q \to 2^-) \;\; \text{and} \tag{30}$$
$$\to -\infty \;\; \text{as} \;\; \epsilon \to 0^- \;\; (\text{i.e.}, q \to 2^+). \tag{31}$$

There is a divergence in the first term, in the form of $1/\epsilon$.

We are then in a position to state that, for $\lambda x > -1$, and given the definition of the q-exponential (see [2]), our "inverse-Jacobian" function $g(x)$ is positive for $q < 2$ and negative for $q > 2$, while diverging at $q = 2$. We conclude that the transform we are studying is not valid only in the isolated case $q = 2$ and changes sign there.

4 Arbitrary constraint

We generalize now the preceding considerations to the case of a generalized constraint $<\ h(x)\ >$, with $h \in \mathcal{L}_2$. The concomitant Shannon MaxEnt solution is [24]

$$p(u)du = e^{-\mu}\, e^{-\lambda h(u)} du, \tag{32}$$
$$\langle h(u) \rangle = K, \tag{33}$$
$$\int p(u)du = 1. \tag{34}$$

while the Tsallis-MaxEnt solution with the same constraint becomes

$$p(x)dx = C\, e_q(-\lambda\, h(x))\, dx; \;\; C = \text{normalization const.} \tag{35}$$

6

Assume that $u(x)$ exists and call, as before, $dx/du = g(x)$. We have, as our cornerstone the relation

$$p(x)dx = p(u)du,$$ (36)

entailing

$$C\,e_q(-\lambda\,h(x))\,dx = e^{-\mu}\,e^{-\lambda h(u(x))}(dx/g(x)),$$ (37)

so that, taking the logarithm to this equation we find

$$\ln C + \ln[g(x)] + \ln[e_q(-\lambda\,h(x))] = -\mu - \lambda\,h(u(x)).$$ (38)

Taking derivatives w.r.t. x yields

$$\frac{g'(x)}{g(x)} - \lambda\,e_q(-\lambda\,h(x))^{q-1}h'(x) = -\lambda\,\frac{1}{g(x)}h'(x),$$ (39)

which leads to a differential equation for our desired transformation function $g(x)$

$$g'(x) - \lambda\,e_q(-\lambda\,h(x))^{q-1}\,h'(x)g(x) + \lambda h'(x) = 0,$$ (40)

quite similar in shape to Eq. (17), being thus solved in similar fashion. Using Eq. (16) we encounter

$$\int P = \int e_q(-\lambda h(x))^{q-1}(-\lambda h'(x))\mathrm{d}x,$$ (41)

$$= \ln(e_q(-\lambda h(x))),$$ (42)

$$e^{-\int P} = e_q(-\lambda h(x))^{-1},$$ (43)

$$\int Qe^{\int P} = \int e_q(-\lambda h(x))(-\lambda h'(x))\mathrm{d}x,$$ (44)

$$= \frac{e_q(-\lambda h(x))^{2-q}}{2-q},$$ (45)

that leads to

$$g(x) = e_q(-\lambda h(x))^{-1}\left[\frac{e_q(-\lambda h(x))^{2-q}}{2-q} + c\right].$$ (46)

It is clear that for $q = 1$ we have $g(x) = 1$ if c=0. Thus,

$$g(x) = \frac{1 - (1-q)\lambda h(x)}{2-q}.$$ (47)

Also, the regime-change at $q = 2$ discussed above does not change.

4.1 Special case $h(x) = x^2$

For the special case of a variance constraint we have

$$g'(x) - \lambda \, e_q(-\lambda \, x^2)^{q-1} \, 2xg(x) = -\lambda 2x, \tag{48}$$

whose solution is

$$g(x) = \frac{1 - (1 - q)\lambda x^2}{2 - q}. \tag{49}$$

Let us take now $q = 1 - \epsilon$ in the solution $g(x) = g(x; q)$. Then, an expansion near $\epsilon = 0$ yields

$$g(x; q = 1 - \epsilon) = 1 - \left(1 + \lambda x^2\right)\epsilon + O(\epsilon^2). \tag{50}$$

5 Generalization to M constraints

Now, we generalize to the case of M constraints of the form:

$$\langle h_i(u) \rangle = K_i; \quad i = 1, \ldots, M. \tag{51}$$

Lets use vector notation and call $\mathbf{h} = (h_1, h_2, \ldots, h_M)$ and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_M)$ its associated Lagrange multipliers. Then we have for the Shannon MaxEnt solution

$$p(u)du = e^{-\mu} \, e^{-\lambda \cdot \mathbf{h}(u)} du, \tag{52}$$

while the Tsallis-MaxEnt solution with the same constraints becomes

$$p(x)dx = C \, e_q(-\lambda \cdot \mathbf{h}(x)) \, dx; \quad C = \text{normalization const.} \tag{53}$$

Assume that $u(x)$ exists and call, as before, $dx/du = g(x)$. Then, following the steps of the previous section and solving the corresponding differential equation, we obtain

$$g(x) = e_q(-\lambda \cdot \mathbf{h}(x))^{-1} \left[\frac{e_q(-\lambda \cdot \mathbf{h}(x))^{2-q}}{2 - q} + c \right], \tag{54}$$

where $c = 0$ in order to get $g(x) = 1$ in the limit $q \to 1$. Thus,

$$g(x) = \frac{1 - (1-q)\lambda \cdot \mathbf{h}(x)}{2 - q}. \tag{55}$$

6 Conclusions

We have shown here that, from a MaxEnt practitioner view-point, one can indistinctly employ Shannon's logarithmic entropy S or Tsallis' power-law one S_q (for any q except $q = 2$). The physics described is the same. To choose between S and S_q is just a matter of convenience in the sense of getting simpler expressions in one case than in the other.

The link between the two concomitant probability distributions $P_{Shannon}(x)$ and $P_{Tsallis}(x)$ is given by the Jacobian $J = 1/g$, where g is the simple function

$$g(x) = \frac{1 - (1-q)\lambda h(x)}{2 - q}, \tag{56}$$

with λ the pertinent Lagrange multiplier, and $h(x) \in \mathcal{L}_2$ an arbitrary function whose mean value $< h >$ constitutes MaxEnt's informational input.

7 Appendix: The four different Tsallis' treatments

A savvy Tsallis practitioner may wonder what happens with the four different ways of computing $q-$mean values that one finds in Tsallis' literature (see [31] and references therein). In addition to the normal expectation values we have employed above, one also encounters, for a quantity $A(x)$ averaging ways that, themselves, depend upon q (see below). This transforms our cornerstone-equality $p_{Shannon}(u)du = p_{Tsallis}(x)dx$ into something much more complicated. However, there is a way out, following the discoveries reported in [31].

Bernoulli published in the *Ars Conjectandi* the first formal attempt to deal with probabilities already in 1713 and Laplace further formalized the subject in his *Théorie analytique des Probabilités* of 1820. In the intervening centuries Probability Theory (PT) has grown into a rich, powerful, and extremely useful branch of Mathematics. Contemporary Physics heavily relies on PT for a large part of its basic structure, Statistical Mechanics [32,33,34], of course, being a most conspicuous example. One of PT basic definitions is that of the mean

9

value of an observable $\mathcal{A}$ (a measurable quantity). Let A stand for the linear operator or dynamical variable associated with $\mathcal{A}$. Then,

$$\langle A \rangle = \int dx \, p(x) \, A(x). \tag{57}$$

This was the averging procedure that Tsallis used in his first, pioneering 1988 paper [1], and the one discussed in the preceding Sections. It is well known that, in some specific cases, it becomes necessary to use "weighted" mean values, of the form

$$\langle A \rangle = \int dx \, f[p(x)] \, A(x), \tag{58}$$

with f an analytical function of p. This happens, for instance, when there is a set of states characterized by a distribution with a recognizable maximum and a large tail that contains low but finite probabilities. One faces then the need of making a pragmatical (usually of experimental origin) decision regarding f [31]. In the first stage of NEXT-development, its pioneer practitioners made the pragmatic choice of using "weighted" mean values, of rather unfamiliar appearance for many physicists. Why? The reasons were of theoretical origin. It was at the time believed that, using the familiar linear, unbiased mean values, one was unable to get rid of the Lagrange multiplier associated to probability-normalization. Since the Tsallis' formalism yields, in the limit $q \to 1$, the orthodox Jaynes-Shannon treatment, the natural choice was to construct weighted expectation values (EVs) using the index q,

$$\langle A \rangle_q = \int dx \, p(x)^q A(x), \tag{59}$$

the so-called Curado-Tsallis unbiased mean values (MV) [35]. As shown in [3], employing the Curado-Tsallis (CT) mean values allowed one to obtain an *analytical* expression for the partition function out of the concomitant MaxEnt process [3]. This EV choice leads to a non extensive formalism endowed with interesting features: i) the above mentioned property of its partition function Z, ii) a numerical treatment that is relatively simple, and iii) proper results in the limit $q \to 1$. It has, unfortunately, the drawback of exhibiting un-normalized mean values, i.e., $\langle\langle 1 \rangle\rangle_q \neq 1$. The latter problem was circumvented in the subsequent work of Tsallis-Mendes-Plastino (TMP) [36], that "normalized" the CT treatment by employing mean values of the form

$$\langle A \rangle_q = \int dx \, \frac{p(x)^q}{\mathcal{X}_q} \, A(x); \quad \mathcal{X}_q = \int dx \, p(x)^q. \tag{60}$$

Most NEXT works employ the TMP procedure. However, the concomitant treatment is not at all simple. Numerical complications often ensue, which has

encouraged the development of different, alternative approach called the OLM one [37], that preserves the main TMP-idea (the $\mathcal{X}_q$ normalization sum) but is numerically simpler. Now, despite appearances, the four Tsallis' treatments are equivalent, as shown in [31]. By equivalence we mean that if one knows the probability treatment P_i; $i = 1, 2, 3, 4$ obtained by anyone of the four treatments, there is a unique, automatic way to write down P_j; $j \neq i$. More precisely, any P_i is a q-exponential, and they all possess the same information-amount if the pair q, β is appropriately "translated" from a version to the other [31]. Indeed,

$$
\begin{aligned}
P_i &= Z^{-1} \exp_{q^*}(-\beta^* x), \\
Z &= \sum_i \exp_{q^*}(-\beta^* x).
\end{aligned}
\tag{61}
$$

Then, as made explicit in [31], given any of the four possible P_i's q^*, β^*, one can get the q, β values appropriate for the q-exponential of any $j \neq i$. Such "dictionary" allows one to translate the results obtained in the preceding Sections to any other of the three remaining averaging procedures.

References

[1] C. Tsallis, J. Stat. Phys. **52** (1988) 479.

[2] M. Gell-Mann, C. Tsallis (Eds.), *Nonextensive Entropy. Interdisciplinary Applications* (Oxford University Press, New York, 2004); C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Springer, New York, 2009).

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

[4] C. Tsallis, M. Gell-Mann, Y. Sato, Proc. Natl. Acad. Sci. USA **102** (2005) 15377; F. Caruso, C. Tsallis, Phys. Rev. E **78** (2008) 021102.

[5] P. Douglas, S. Bergamini, F. Renzoni, Phys. Rev. Lett. **96** (2006) 110601; G.B. Bagci, U. Tirnakli, Chaos **19** (2009) 033113.

[6] B. Liu, J. Goree, Phys. Rev. Lett. **100** (2008) 055003.

[7] R.G. DeVoe, Phys. Rev. Lett. **102** (2009) 063001.

[8] R.M. Pickup, R. Cywinski, C. Pappas, B. Farago, P. Fouquet, Phys. Rev. Lett. **102** (2009) 097202.

[9] L.F. Burlaga, N.F. Ness, Astrophys. J. **703** (2009) 311.

[10] F. Caruso, A. Pluchino, V. Latora, S. Vinciguerra, A. Rapisarda, Phys. Rev. E **75** (2007) 055101(R); B. Bakar, U. Tirnakli, Phys. Rev. E **79** (2009) 040103(R); A. Celikoglu, U. Tirnakli, S.M.D. Queiros, Phys. Rev. E **82** (2010) 021124.

[11] V. Khachatryan, et al., CMS Collaboration, J. High Energy Phys. **1002** (2010) 041; V. Khachatryan, et al., CMS Collaboration, Phys. Rev. Lett. **105** (2010) 022002.

[12] Adare, et al., PHENIX Collaboration, Phys. Rev. D **83** (2011) 052004; M. Shao, L. Yi, Z.B. Tang, H.F. Chen, C. Li, Z.B. Xu, J. Phys. G **37** (8) (2010) 085104.

[13] M.L. Lyra, C. Tsallis, Phys. Rev. Lett. **80** (1998) 53; E.P. Borges, C. Tsallis, G.F.J. Ananos, P.M.C. de Oliveira, Phys. Rev. Lett. **89** (2002) 254103; G.F.J. Ananos, C. Tsallis, Phys. Rev. Lett. **93** (2004) 020601; U. Tirnakli, C. Beck, C. Tsallis, Phys. Rev. E **75** (2007) 040106(R); U. Tirnakli, C. Tsallis, C. Beck, Phys. Rev. E **79** (2009) 056209.

[14] L. Borland, Phys. Rev. Lett. **89** (2002) 098701.

[15] A.R. Plastino, A. Plastino, Phys. Lett A **174** (1993) 834.

[16] A.R. Plastino, A. Plastino, Physica A **222** (1995) 347.

[17] E.P. Borges, Physica A **340** (2004) 95.

[18] S. Umarov, C. Tsallis, S. Steinberg, Milan J. Math. **76** (2008) 307; S. Umarov, C. Tsallis, M. Gell-Mann, S. Steinberg, J. Math. Phys. **51** (2010) 033502.

[19] M. Jauregui, C. Tsallis, J. Math. Phys. **51** (2010) 063304.

[20] A. Chevreuil, A. Plastino, C. Vignat, J. Math. Phys. **51** (2010) 093502.

[21] A. Plastino and M.C. Rocca, J. Math. Phys **52** (2011) 103503.

[22] H.J. Hilhorst, J. Stat. Mech. (2010) P10023.

[23] E. T. Jaynes, Phys. Rev. **106** (1957) 620.

[24] A. Katz, *Principles of Statistical Mechanics: The Information Theory Approach* (Freeman and Co., San Francisco, 1967).

[25] A. Plastino, F. Olivares, S. Flego, M. Casas, Entropy **13** (2011) 184.

[26] S.P. Flego, A. Plastino, and A.R. Plastino, Physica A **390** (2011) 2276.

[27] S.P. Flego, A. Plastino, A.R. Plastino, J. Math. Phys. **52** (2011) 082103.

[28] S.P. Flego, A. Plastino, A.R. Plastino, Ann. Phys. **326** (2011) 2533.

[29] S. P. Flego, A. Plastino, A. R. Plastino, ENTROPY **13** (2011) 2049-2058.

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

[31] G. L. Ferri, S. Martinez, A. Plastino, J. Stat. Mech. (2005) P04009.

[32] R.K. Pathria, *Statistical Mechanics* (Pergamon Press, Exeter, 1993).

[33] F. Reif, *Statistical and thermal physics* (McGraw-Hill, NY, 1965).

[34] J. J. Sakurai, *Modern quantum mechanics* (Benjamin, Menlo Park, Ca., 1985).

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

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

[37] S. Martínez, F. Nicolás, F. Pennini, A. Plastino, Physica A **286** (2000) 489.