

# Peculiarities of some classical variational treatments using the maximum entropy principle

A. Plastino<sup>a,c,d</sup> and M.C.Rocca<sup>a,b,c</sup>

<sup>a</sup> Departamento de Física, Universidad Nacional de La Plata,

<sup>b</sup> Departamento de Matemática, Universidad Nacional de La Plata,

<sup>c</sup> Consejo Nacional de Investigaciones Científicas y Tecnológicas, (IFLP-CCT-CONICET)-C. C. 727, 1900 La Plata - Argentina.

<sup>d</sup> SThAR - EPFL, Lausanne, Switzerland.

Received 5 June 2018; accepted 2 August 2018

We study some peculiarities of the classical variational treatment that applies Jaynes' maximum entropy principle. The associated variational treatment is usually called MaxEnt. We deal with it in connection with thermodynamics' reciprocity relations. Two points of view are adopted: (A) One of them is purely abstract, concerned solely with ascertaining compliance of the variational solutions with the reciprocity relations in which one does not need here to have explicit values for the Lagrange multipliers. The other, (B) is a straightforward variation process in which one explicitly obtains the specific values of these multipliers. We focus on the so called  $q$ -entropy because it illustrates a situation in which the above two approaches yield different results. We detect an information loss in extracting the explicit form of the normalization-associated Lagrange multipliers.

**Keywords:** Tsallis-entropy; MaxEnt; Variational treatments; Reciprocity relations.

PACS: 05.20.-y; 05.70.Ce; 05.90.+m

## 1. Introduction

The most popular approach to develop the main Statistical Mechanics formalisms proceeds today via application of Jaynes' maximum entropy principle, usually abbreviated as MaxEnt [1]. We apply it here in connection with generalized entropies of the Tsallis type, that have become a very important sub-field of statistical mechanics, with several thousand papers and applications on many scientific disciplines [2, 3]. Since the number of references is at least ten times the order just mentioned, we will mainly direct the reader to [2, 3] and references therein. In this effort we focus attention on reciprocity relations and re-consider some issues concerning generalized entropies that, we believe, lack yet adequate understanding, even if they have been on the discussion table for many years in variegated publications. In particular, we want to shed light on some issues regarding the canonical ensemble. More specifically:

1. The way to explicitly obtain the normalization Lagrange multiplier  $\lambda_N$  in the Tsallis' variational problem with *linear constraints*.
2. A two-way approach to reciprocity relations.
3. Differences between them that entail information loss.

## 2. Preliminary matters

### 2.1. Notation

- $\lambda_U$  is the energy  $U$  multiplier, which is related to the system's temperature  $T$ ,

- $\lambda_N$  is the normalization multiplier.

In statistical mechanics, these multipliers are always endowed with meaningful physical information [4].

We use the  $q$ -functions [2]

$$e_q(x) = [1 + (1 - q)x]^{1/(1-q)};$$

$$e_q(x) = \exp(x) \text{ for } q = 1; \quad (1)$$

$$\ln_q(x) = \frac{x^{(1-q)} - 1}{1 - q};$$

$$\ln_q(x) = \ln(x) \text{ for } q = 1. \quad (2)$$

### 2.2. Reciprocity relations and thermodynamics

It is well known that the Legendre transform (LT) constitutes an operation that converts a real function  $f$  with real variable  $x$  into another  $f_T$ , of another variable  $y$ , keeping constant the information content of  $f$ . The derivative of  $f$  becomes the argument of  $f_T$ .

$$f_T(y) = xy - f(x); \quad y = f'(x) \Rightarrow \text{reciprocity.} \quad (3)$$

The LT becomes its own inverse. Famously, one appeals to it to pass from Lagrangians to Hamiltonians in classical mechanics.

LT' reciprocity relations constitute thermodynamics' essential formal ingredient [5]. For two functions  $I$  (for instance, information measure) and  $\tilde{I}$ , one has

$$I(A_1, \dots, A_M) = \tilde{I} + \sum_{k=1}^M \lambda_k A_k, \quad (4)$$

with the  $A_i$  extensive variables and the  $\lambda_i$  independent intensive ones. Obviously, the Legendre transform main goal is that of changing the identity of our relevant independent variables. For  $\tilde{I}$  we have

$$\tilde{I}(\lambda_1, \dots, \lambda_M) = I - \sum_{k=1}^M \lambda_k \langle A_k \rangle. \tag{5}$$

One further has [5]

$$\begin{aligned} \frac{\partial \tilde{I}}{\partial \lambda_k} &= -\langle A_k \rangle; & \frac{\partial I}{\partial \langle A_k \rangle} &= \lambda_k; \\ \frac{\partial I}{\partial \lambda_i} &= \sum_k^M \lambda_k \frac{\partial \langle A_k \rangle}{\partial \lambda_i}, \end{aligned} \tag{6}$$

the reciprocity relations, the last one being the so-called Euler theorem. In this paper we pay special attention to the specific reciprocity relation

$$\frac{\partial I}{\partial \langle A_k \rangle} = \lambda_k. \tag{7}$$

Why? Because in Jaynes' philosophy [1]  $I$  is the information amount, to be maximized subject to known constraints  $\langle A_k \rangle$ . The associated Lagrange multipliers are to be obtained by solving the partial differential equations (7), that here after will be called the determining relations (DR).

**2.3. Boltzmann-Gibbs (BG) MaxEnt variational problem**

It is useful to recapitulate it. We work on  $\mathcal{R}^N$ . The volume element is called  $dV$ . One has

$$S_{BG} = - \int dV f(p) = - \int dV p \ln p, \tag{8}$$

with the variational problem leading to

$$f'(p) - \lambda_N - \lambda_U U = 0, \tag{9}$$

and

$$f' = - \ln p - 1, \tag{10}$$

We define now  $g(\xi)$  as the inverse of  $f'(\xi)$  such that  $g[f'(\xi)] = \xi$  and here

$$g(\nu) = \exp[-(\nu + 1)],, \tag{11}$$

and thus

$$\xi_{ME} = g(\lambda_N + \lambda_U U), \tag{12}$$

with

$$\begin{aligned} \xi_{ME} &= \exp[-(\lambda_N + 1 + \lambda_U U)] \\ &= \exp[-(\lambda_N + 1)] \exp[-\lambda_U U], \end{aligned} \tag{13}$$

so that one can easily extract, via normalization, a partition function  $Z$

$$\begin{aligned} \int dV \exp[-(\lambda_N + 1)] \exp[-\lambda_U U] &= 1 \\ \int dV \exp[-\lambda_U U] &= Z = \exp[(\lambda_N + 1)] \\ \xi_{ME} &= \frac{\exp[-\lambda_U U]}{\int dV \exp[-\lambda_U U]} \\ \ln Z &= \lambda_N + 1, \end{aligned} \tag{14}$$

and one obtains explicitly the relation between  $Z$  and  $\lambda_N$ . It goes without saying that the reciprocity relations are satisfied [1], in particular the determining relation of the precedent subsection. Moreover, it is found [1] that

$$\lambda_U = 1/kT, \tag{15}$$

with  $k$  Boltzmann's constant.

**3. Normalization of the Lagrange multiplier in Tsallis' MaxEnt with linear constraints**

**3.1. Variational problem**

This subject was fully treated for i) trace-form entropies and ii)  $M$  observables as constraints, in [6]. We regard it instructive the explicit Tsallis-application of such discussion, which, as far as we know, has not been given before in the detailed fashion available here.

Our probability density functions (PDFs) are designed with Greek letters like  $\xi$ .  $\xi_{ME}$  stands for the MaxEnt PDF.

We have two identical ways of defining the q- entropy

$$S_1 = \int dV f(\xi), \tag{16}$$

with

$$f(\xi) = \frac{\xi - \xi^q}{q - 1}, \tag{17}$$

and

$$S_2 = \frac{1}{q - 1} \left[ 1 - \int dV \xi^q \right], \tag{18}$$

that however, leads to different variational problems, as we shall immediately see. Our a priori knowledge is the mean energy  $\langle U \rangle$  (canonical ensemble). The MaxEnt variational problem for  $S_2$  becomes

$$\begin{aligned} f'(\xi) - \lambda_N - \lambda_U U &= -q\xi^{q-1}/(q - 1) \\ - \lambda_N - \lambda_U U &= 0, \end{aligned} \tag{19}$$

with

$$f'(\xi) = -q\xi^{q-1}/(q - 1). \tag{20}$$

Instead, for  $S_1$  one has

$$\frac{1 - q\xi^{q-1}}{q - 1} - \lambda_U U - \lambda_N = 0, \tag{21}$$

$$f'(\xi) = \frac{1 - q\xi^{q-1}}{q - 1}. \tag{22}$$

We define now  $g(\xi)$  as the inverse of  $f'(\xi)$  such that  $g[f'(\xi)] = \xi$ . One has for the  $S_1$  instance

$$g(\nu) = q^{1-q}[1 - (q - 1)\nu]^{1/(q-1)} = q^{1-q}e_{(2-q)}(\nu). \tag{23}$$

From (21) it is obvious that

$$\xi_{ME} = g(\lambda_N + \lambda_U U), \tag{24}$$

or

$$\xi_{ME} = g(\lambda_N + \lambda_U U) = q^{1-q}e_{(2-q)}(\lambda_N + \lambda_U U), \tag{25}$$

so that you cannot extract  $\lambda_N$  from that expression. One does not obtain explicitly the relation between  $Z$  and  $\lambda_N$  as in (14) for BG.

A similar result is obtained if we consider  $S_2$  instead of  $S_1$ , i.e.,

$$\begin{aligned} \xi_{ME} &= g(\lambda_N + \lambda_U U) \\ &= [1 - (q - 1)(\lambda_N + \lambda_U U)]^{1/(q-1)}. \end{aligned} \tag{26}$$

Note that the  $g$  for  $S_1$  and the  $g$  for  $S_2$  are slightly different.

This extraction problem affects only  $S_q$ 's MaxEnt treatment with *linear* constraint and can be avoided by recourse to nonlinear constraints, as in Refs. [7, 8].

Our interest resides precisely in discussing the peculiarities of the MaxEnt variational treatment that are emerging with regards to the present problem (and the reason for dealing with linear constraints here). The information contained in our variational problem can be managed in two different manners. The following Subsection deals with the first of these ways, that we may call the abstract one. Sections 4 and 5 consider an alternative route, that we may call the explicit one. The two ways yield different results, as we shall see, originating what we regard as an information management problem.

### 3.2. Reciprocity relations

These pitfalls notwithstanding, the reciprocity relations hold. This is so because they depend only on the formal existence of the function  $g(\nu)$ . We specify now for this case the general treatment of [6]. We do not need, in the subsequent manipulations, to distinguish between  $S_1$  and  $S_2$ .

Because of  $\xi$ -normalization it is clear that, in general, both for  $S_1$  and for  $S_2$ , one has

$$\begin{aligned} \frac{\partial}{\partial \lambda_U} \int dV \xi &= 0 \\ &= \int dV g'(\lambda_N + \lambda_U U) \left[ \frac{\partial \lambda_N}{\partial \lambda_U} + U \right] = 0, \end{aligned} \tag{27}$$

a relation that we presently will employ. Also, we have, for  $\partial \langle U \rangle / (\partial \lambda_U)$

$$\frac{\partial \langle U \rangle}{\partial \lambda_U} = \int dV U g'(\lambda_N + \lambda_U U) \left[ \frac{\partial \lambda_N}{\partial \lambda_U} + U \right]. \tag{28}$$

Next we consider  $\partial S / \partial \lambda_U$  and write

$$\frac{\partial S}{\partial \lambda_U} = \frac{\partial}{\partial \lambda_U} \int dV f[g(\lambda_N + \lambda_U U)] \tag{29}$$

$$\begin{aligned} &= \int dV f'[g(\lambda_N + \lambda_U U)] \\ &\quad \times g'(\lambda_N + \lambda_U U) \left[ \frac{\partial \lambda_N}{\partial \lambda_U} + U \right]. \end{aligned} \tag{30}$$

Remembering that  $f'g(\nu) = \nu$  and (27) we simplify this to

$$\frac{\partial S}{\partial \lambda_U} = \lambda_U \int dV U g'(\lambda_N + \lambda_U U) \left[ \frac{\partial \lambda_N}{\partial \lambda_U} + U \right], \tag{31}$$

leading to

$$\frac{\partial S}{\partial \lambda_U} = \lambda_U \frac{\partial \langle U \rangle}{\partial \lambda_U}, \tag{32}$$

the Euler relation. Now we have

$$\frac{\partial S}{\partial \langle U \rangle} = \frac{\partial S}{\partial \lambda_U} \frac{\partial \lambda_U}{\partial \langle U \rangle} = \lambda_U \frac{\partial \langle U \rangle}{\partial \lambda_U} \frac{\partial \lambda_U}{\partial \langle U \rangle} = \lambda_U, \tag{33}$$

the first reciprocity relation.

Introduce now the Jaynes' parameter  $\tilde{S}$  (the Legendre transform of  $S$ )

$$\tilde{S}(\lambda_U) = S(\langle U \rangle) - \lambda_U \langle U \rangle(\lambda_U). \tag{34}$$

It is clear that

$$\frac{\partial \tilde{S}}{\partial \lambda_U} = \frac{\partial S}{\partial \langle U \rangle} \frac{\partial \langle U \rangle}{\partial \lambda_U} - \lambda_U \frac{\partial \langle U \rangle}{\partial \lambda_U} - \langle U \rangle = -\langle U \rangle, \tag{35}$$

the second reciprocity relation. Note that we do not need to explicitly ascertain specific values to  $\lambda_U$  and  $\lambda_N$  in order to establish the reciprocity relations. In particular, we do not need to solve the determining equation of Subsec. 2.2.

### 3.3. Solving for the Lagrange multipliers

Usually, this is a very difficult numerical problem. A practical alternative is to numerically solve, once we have  $\xi_{ME}$ , the  $M + 1$  set of equations that read, using the notation of Subsec. 2.2,

$$\int dV \xi_{ME}(\lambda_0, \lambda_1, \dots, \lambda_M) A_k = \langle A_k \rangle, \tag{36}$$

with  $k = 0, 1, \dots, M$ , ( $A_0 = 1$ ). This gives the  $M + 1$  Lagrange multipliers in terms of the assumedly known  $M + 1$  quantities  $\langle A_k \rangle$  [in particular, if there is an energy multiplier (called  $\lambda_U$  above) it is set equal to  $1/kT$ ]. See Ref. [9] for details. Can one bypass this difficult process? This is what we will try to do next.

**4. Explicit Lagrange multipliers in Tsallis' original variational problem for  $S_1$**

Return now to (21). We saw that we can not immediately derive from it a value for  $\lambda_N$  [2]. A heuristic solution, is to set

$$\lambda_N = -\frac{q}{q-1} Z_T^{q-1} + \frac{1}{q-1} = \frac{1}{q-1} [1 - qZ_T^{q-1}] \quad (37)$$

with  $Z_T$  unknown at this stage, and rename  $\lambda_U$  as follows

$$\lambda_U = qZ_T^{1-q}\beta. \quad (38)$$

where  $\beta = (1/k_B T)$ ,  $T$  the temperature. The variational problem is now

$$\begin{aligned} \frac{1 - q\xi^{q-1}}{(q-1)} &= -\frac{q}{q-1} Z_T^{1-q} \\ &+ \frac{1}{q-1} + Z_T^{1-q} q\beta U = 0 \end{aligned} \quad (39)$$

$$\xi^{q-1} = Z_T^{1-q} [1 - (q-1)\beta U], \quad (40)$$

so that

$$\xi_{ME} = Z_T^{-1} [1 - (q-1)\beta U]^{1/(q-1)}, \quad (41)$$

where  $\beta$  is NOT the variational multiplier  $\lambda_U$ , as stated above. Further,

$$Z_T = \int dV [1 - (q-1)\beta U]^{1/(q-1)}. \quad (42)$$

We have now

$$\xi^q \xi^{1-q} = \xi; \quad \xi^q \ln_q(\xi) = \frac{\xi - \xi^q}{1 - q}, \quad (43)$$

and then

$$S_q = - \int dV \xi^q \ln_q(\xi) = \int dV \xi [1 - \xi^{q-1}] / (q-1) \quad (44)$$

$$= \int dV \xi [1 - (1/Z_T)^{q-1} [1 - (q-1)\beta U]] / (q-1) \quad (45)$$

$$= \int dV \xi [ [1 - (1/Z_T)^{q-1}] / (q-1) + (1/Z_T)^{q-1} \beta U ] \quad (46)$$

$$= \int dV \xi [ \ln_q(Z_T) + Z_T^{1-q} \beta U ]. \quad (47)$$

$$S_q = \ln_q(Z_T) + Z_T^{1-q} \beta \langle U \rangle, \quad (48)$$

so that

$$\frac{\partial S_q}{\partial \langle U \rangle} = \beta Z_T^{1-q} = \lambda_U / q, \quad (49)$$

the quasi-reciprocity relation with a denominator  $q$ . Giving  $\lambda_N$  an explicit form has resulted in a wrong reciprocity relation. This fact could be interpreted as an information loss, as a result of not solving the determining equation of Subsec. 2.2.

**5. Explicit Lagrange multipliers for  $S_2$**

Return now to (20). Let us insist on solving the variational problem, though, without solving the determining equation of Subsec. 2.2. An heuristic solution is to set [see [10]]

$$\lambda_N = -\frac{q}{q-1} Z_T^{1-q} \quad (50)$$

$$Z_T \text{ yet unknown} \quad (51)$$

and rename  $\lambda_U$  as follows

$$\lambda_U = qZ_T^{1-q}\gamma \quad \gamma = \frac{\lambda_U}{qZ_T^{1-q}}. \quad (52)$$

Now  $\gamma$  can be set equal to  $1/kT$ . The variational problem is here

$$-q\xi^{q-1}/(q-1) + \frac{q}{q-1} Z_T^{1-q} - qZ_T^{1-q}\gamma U = 0 \quad (53)$$

$$\xi^{q-1} = Z_T^{1-q} [1 - (q-1)\gamma U], \quad (54)$$

so that

$$\begin{aligned} \xi_{ME} &= Z_T^{-1} e_{2-q}(-\lambda_U U) = Z_T^{-1} [1 - (q-1)\gamma U]^{1/(q-1)}; \\ \xi_{ME}^{q-1} &= Z_T^{1-q} [1 - (q-1)\gamma U]^{1/(q-1)} \end{aligned} \quad (55)$$

where  $\beta$  is not the variational  $\lambda_U$ . Further,

$$Z_T = \int dV [1 - (q-1)\gamma U]^{1/(q-1)}. \quad (56)$$

Thus, we have

$$\xi^q \xi^{1-q} = \xi; \quad \xi^q \ln_q(\xi) = \frac{\xi - \xi^q}{1 - q}, \quad (57)$$

and then

$$S_q = - \int dV \xi^q \ln_q(\xi) = \int dV \xi [1 - \xi^{q-1}] / (q-1) \quad (58)$$

$$= \int dV \xi [1 - (1/Z_T)^{q-1} [1 - (q-1)\gamma U]] / (q-1) \quad (59)$$

$$= \int dV \xi [ [1 - (1/Z_T)^{q-1}] / (q-1) + (1/Z_T)^{q-1} \gamma U ] \quad (60)$$

$$= \int dV \xi [ \ln_q(Z_T) + Z_T^{1-q} \gamma U ]. \quad (61)$$

$$S_q = \ln_q(Z_T) + Z_T^{1-q} \beta \langle U \rangle = \ln_q(Z_T) + \lambda_U \langle U \rangle / q, \quad (62)$$

so that one is led to a slightly modified reciprocity relation

$$\frac{\partial S_q}{\partial \langle U \rangle} = \gamma Z_T^{1-q} = \lambda_U / q, \quad (63)$$

identical in form to that for  $S_2$ , but wrong as well. Once more, giving  $\lambda_N$  an explicit form has resulted in an incorrect reciprocity relation. We detect then what could be read as an information loss, as a result of not solving the determining equation of Subsec. 2.2.

## 6. Conclusions

We have revisited Tsallis' original treatment of the q-entropy [2] with focus on the reciprocity relations. They are valid, of course, as has been known for many years already. However, some peculiarities of the treatment have been detected here.

- There are two forms of casting the q-entropy, that we called  $S_1$  and  $S_2$ . They are identical, but the associated MaxEnt variational treatments do differ for each of them if one wants explicit values for the Lagrange multipliers.
- The Lagrange normalization multiplier  $\lambda_U$  can not be obtained in explicit fashion, as it is well known [2]. We have found here ways to overcome this obstacle and obtained two different versions of  $\lambda_U$ , associated to  $S_1$  and  $S_2$ , respectively.
- The ways to overcome the obstacle encountered here cannot be used in physical applications, though, since they entail information loss. The main physical consequence of this fact is that appeal to the methodology of Sec. 3.3 is unavoidable.
- There is a price to pay for our  $\lambda_U$ -extraction procedure. The ensuing reciprocity relation for entropy/energy,  $(\partial S_q / \partial \langle U \rangle)$ , equals not  $\lambda_U$  (as it should) but

$\lambda_U/q$ , for both cases  $S_1$  and  $S_2$ , as a result of not solving the determining equation of Subsec. 2.2. It is nonetheless gratifying that our two wrong equations coincide, since the entropy is just one and the physics, *i.e.*, the reciprocity relation, should not depend on whether we use  $S_1$  or  $S_2$ .

- Remark that  $S_1 = S_2$  as functions but their associated variational problems are not identical.
- In abstract form one can show that the reciprocity relations are indeed valid, as we showed in Sec. 3.2, without need for appealing to explicit knowledge of the Legendre multipliers. One only requires that a special function that we called  $g$  does exist.
- This  $g$  is different in the  $S_1$  or the  $S_2$  cases.
- Obtaining the Lagrange normalization multiplier  $\lambda_U$  in explicit fashion, although accomplished via a seemingly legitimate symbolic manipulation, entails however some information loss, as the reciprocity relation entropy-energy is not exactly re-obtained, as a result of not solving the determining equation of Subsec. 2.2. Usually, this is a very difficult numerical problem. A practical alternative, as we saw above, is to numerically solve, once we have  $\xi_{ME}$ , the  $M + 1$  set of equations (36).

- 
1. E.T. Jaynes, in: *Statistical physics*, ed. by W.K. Ford (Benjamin, New York, 1963); A. Katz, *Statistical mechanics* (Freeman, San Francisco, 1967).
  2. M. Gell-Mann and C. Tsallis, Eds. *Nonextensive Entropy: Interdisciplinary applications*, Oxford University Press, (Oxford, 2004); C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, (New York, 2009).
  3. See <http://tsallis.cat.cbpf.br/biblio.htm> for a regularly updated bibliography on the subject.
  4. R. B. Lindsay, H. Margenau, *Foundations of physics* (Dover, NY, 1957).
  5. E. A. Desloge, *Thermal physics* NY, Holt, (Rhinehart and Winston, 1968).
  6. A. Plastino, A. R. Plastino, *Phys. Lett. A* **226** (1997) 257.
  7. E. M. F. Curado, C. Tsallis, *J. Phys. A* **24** (1991) L69.
  8. C. Tsallis, R. S. Mendes, A.R. Plastino, *Physica A* **261** (1998) 534.
  9. J. Skilling, *Massive Inference and Maximum Entropy*, in *Maximum Entropy and Bayesian Methods*, pages 1-14, (1997). Proceedings of the 17th International Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis, Editors J. Erickson, Joshua T. Rychert, C. Ray Smith (Springer, New York).
  10. A. Plastino, M. C. Rocca, F. Pennini, *Phys. Rev. E* **94** (2016) 012145.