We describe in great generality features concerning constrained entropic, functional variational problems that allow for a broad range of applications. Our discussion encompasses not only entropies but, potentially, any functional of the probability distribution, like Fisher-information or relative entropies, etc. In particular, in dealing with generalized statistics in straightforward fashion one may sometimes find that the first thermal law \( \frac{dS}{d\beta} = \beta \frac{d\langle U \rangle}{d\beta} \) seems to be not respected. We show here that, on the contrary, it is indeed obeyed by any system subject to a Legendre extremization process, i.e., in all constrained entropic variational problems.

I. INTRODUCTION

Generalized entropies have become in the last 25 years a very important sub-field of statistical mechanics, with multiple applications to many scientific disciplines \([1-18]\). Among the variegated set of physical scenarios to which these entropic measures have been applied we can mention the thermostatistics of systems with long range interactions \([1, 2]\), thermodynamics of many-particle systems in the overdamped motion regime \([3]\), plasma physics \([4, 5]\), diverse aspects of stellar dynamics \([6, 7]\), chaotic dynamical systems \([8, 9]\) (specially, systems exhibiting weak chaos \([10]\)), Bose-Einstein condensation \([11]\), thermodynamic-like description of the ground state of quantum systems \([12]\), nonlinear Schrödinger equations \([13]\), speckle patterns generated by rough surfaces \([14]\), metal melting \([15]\), and the statistics of postural sway in humans \([16]\). Tsallis’ entropy is the paramount example of a generalized entropy and the associated thermostatistic is, by far, the one that has been most intensively investigated. The above list of recent developments on generalized entropies and their applications (most of them concerning Tsallis entropy) is only illustrative. In spite of the mind blowing diversity of subjects to which Tsallis theory has been applied, there actually are a few underlying basic themes that connect many of these applications. Arguably, among these common threads the three most important ones are (1) many-body systems with interactions whose range is of the same order as the size of the system (that is, long-range interactions), (2) systems governed by nonlinear Fokker-Planck equations involving power-law diffusion terms, and (3) weak chaos. For a more detailed discussion of the vast research literature dealing with these matters see \([19, 20]\) and references therein. In this effort we focus attention on the statistical derivation of thermodynamics’ first law in the guise

\[
\frac{dS}{d\beta} = \beta \frac{d\langle U \rangle}{d\beta},
\]

where \( \beta \) is the inverse temperature, \( S \) the entropy and \( U \) the internal energy. This is trivial in the case of Boltzmann-Gibbs’ logarithmic entropy and can be looked up in any text-book \([21, 22]\). However, for general entropies such is not the case (see, as one of many possibilities, Ref. \([23]\)). Let us look in detail at a famous example so as to clearly illustrate the problem we are talking about.

A. A typical abeyance example

An example is appropriate to appreciate the difficulties we are here referring to. Our probability density functions (PDFs) are designed with the letter \( p \), and \( p_{ME} \) would stand for the MaxEnt PDF.

We will use the q-functions \([19]\)

\[
e_q(x) = \begin{cases} 1 + (1-q)x \sqrt{1-q} & \text{for } q < 1 \\ \exp(x) & \text{for } q = 1 \end{cases}
\]
\[
\ln_q(x) = \frac{x^{(1-q)} - 1}{1 - q}; \quad \ln_q(x) = \ln(x) \text{ for } q = 1.
\] (3)

We define the Tsallis q-entropy, for any real \( q \), as
\[
S_T = \int dx f(p),
\] (4)
with
\[
f(p) = \frac{p - p^q}{q - 1}.
\] (5)

Our a priori knowledge is that of the mean energy \( \langle U \rangle \) (canonical ensemble). The MaxEnt variational problem becomes, with Lagrange multipliers \( \lambda_N, \lambda_U \)
\[
\frac{1 - qp^{q-1}}{q - 1} - \lambda_U U - \lambda_N = 0,
\] (6)
\[
f'(p) = \frac{1 - qp^{q-1}}{q - 1}.
\] (7)

One conveniently defines here \( g(p) \) as the inverse of \( f'(p) \) such that \( g[f'(p)] = p \). One has
\[
g(\nu) = q^{1-q}[1 - (q - 1)p^\nu]^{1/(q-1)} = q^{1-q}e_{(2-q)}(p^\nu).
\] (8)

It is obvious that
\[
p_{ME} = g(\lambda_N + \lambda_U U),
\] (9)
or
\[
p_{ME} = g(\lambda_N + \lambda_U U) = q^{1-q}e_{(2-q)}(\lambda_N + \lambda_U U),
\] (10)
so that one cannot extract \( \lambda_N \) from that expression. Moreover, you do not obtain explicitly the relation between \( Z \) and \( \lambda_N \). Since one can not immediately derive from it a value for \( \lambda_N \), a heuristic alternative is to introduce
\[
\lambda_N = -\frac{q}{q - 1}Z_T^{q-1} + \frac{1}{q - 1} = \frac{1}{q - 1}[1 - qZ_T^{q-1}],
\] (11)
with \( Z_T \) unknown for the time being, and re-express \( \lambda_U \) in the guise
\[
\lambda_U = qZ_T^{1-q}\beta.
\] (12)

where \( \beta \) is determined by the above equation. The variational problem becomes
\[
\frac{1 - qp^{q-1}}{(q - 1)} = -\frac{q}{q - 1}Z_T^{1-q} + \frac{1}{q - 1} + Z_T^{1-q}q\beta U = 0
\] (13)
\[
p^{q-1} = Z_T^{1-q}[1 - (q - 1)\beta U],
\] (14)
and yields
\[
p_{ME} = Z_T^{-1}[1 - (q - 1)\beta U]^{1/(q-1)},
\] (15)
where \( \beta \) is definitely NOT the variational multiplier \( \lambda_U \). Moreover, we can now have an expression for

\[
Z_T = \int dx \left[ 1 - (q - 1)\beta U \right]^{1/(q-1)}.
\]

(16)

Thus, we have

\[
p^q p^{1-q} = p; \quad p^q \ln q(p) = \frac{p^q - p^q}{1 - q},
\]

(17)

and then

\[
S_q = -\int dx p^q \ln q(p) = \int dx p \left[ (1 - \frac{1}{Z_T})^{q-1} \right] / (q - 1) = \int dx \left[ (1 - \frac{1}{Z_T})^{q-1} \right] + (1/Z_T) \beta U
\]

(18)

\[
= \int dx \left[ (1 - \frac{1}{Z_T})^{q-1} \right] / (q - 1) + (1/Z_T) \beta U
\]

(19)

\[
= \int dx \left[ (1 - \frac{1}{Z_T})^{q-1} \right] + (1/Z_T) \beta U
\]

(20)

\[
S_q = \ln q(Z_T) + Z_T^{1-q} \beta \langle U \rangle = \ln q(Z_T) + \beta \langle U \rangle / q,
\]

(22)

entailing

\[
\frac{dS_q}{d\beta} = \frac{Z_T^{(2-q)/(1-q)} c_q^{2-q} \langle U \rangle}{1 - q} + \langle U \rangle / q + \beta \frac{\delta \langle U \rangle / q}{\delta \beta}.
\]

(23)

We encounter now, as a result, that Eq. (1) is violated for \( \beta \). This is a fact that has created some confusion in the Literature [23].

**B. Our present goal**

We will proceed, starting with the next Section, to overcome the difficulties posed by the above kind of situations. The paper is organized as follows. Section II contains a very general proof. It applies to any functional of the probability distribution, like generalized entropies, Fisher information, relative entropies, etc. We will demonstrate the fact that Eq. (1) always holds, no matter what the quantifier one has in mind might be, becoming in fact a basic result of the variational problem. In order to further clarify the issue at hand we specify this proof in several particular instances of interest. In Section III we revisit Tsallis’ quantifier. Renenyi’s entropy is discussed in Sect. IV. An arbitrary, trace form entropic quantifier is the focus of Section V and, finally, an also arbitrary entropic functional lacking trace form is examined in Sect. VI. Some conclusions are drawn in Sect. VII.

**II. THE GENERAL VARIATIONAL PROBLEM**

**A. Functional derivatives: a brief reminder**

A functional \( F \) of a distribution \( g \) is a mapping between a collection of \( g \)’s and a set of numbers [24]. The functional derivative can be introduced via the Taylor expansion

\[
f[g + \epsilon h] = F[g] + \epsilon \int dx \frac{\delta F}{\delta g(x)} h(x) + O(\epsilon^2),
\]

(24)

for any reasonable \( h(x) \). Here \( \frac{\delta F}{\delta g(x)} \) becomes the definition of a functional derivative. Note that it is both a function of \( x \) and a functional of \( g \). In our case, generalized entropies constitute our foremost example of a functional.
B. The general problem

Let $F$ and $G$ be functionals of a normalized probability density function (PDF) $f$.

$$F = F[f]; \quad G = G[f]; \quad \int dx f(x) = 1. \quad (25)$$

Given two functionals $F$ and $G$, one wishes to extremize $F$ subject to a fixed value for $G$. The ensuing variational problem reads

$$\delta [F - bG - af] \Rightarrow \delta F - b \frac{\delta G}{\delta f} - a = 0, \quad (26)$$

while

$$\int dx f(a, b, x) = 1 \Rightarrow \int dx [\frac{\partial f}{\partial b} + \frac{\partial f}{\partial a} \frac{\partial a}{\partial b}] = 0. \quad (27)$$

Eq. (27) plays a very important role in our endeavors, as we will presently see. We now face

$$\frac{dF}{db} = \int dx \frac{\delta F}{\delta f} \frac{\partial f}{\partial b} + \frac{\partial f}{\partial a} \frac{\partial a}{\partial b}, \quad (28)$$

so that, using (27), as just promised

$$\frac{dF}{db} = [b \frac{\delta G}{\delta f} + a] \frac{\partial f}{\partial b} + \frac{\partial f}{\partial a} \frac{\partial a}{\partial b}. \quad (29)$$

Use now $f$–normalization to derive the fundamental relation

$$\frac{dF}{db} = [b \frac{\delta G}{\delta f}] \frac{\partial f}{\partial b} + \frac{\partial f}{\partial a} \frac{\partial a}{\partial b} = b(dG/db). \quad (30)$$

QED. The theme has been broached in different manners to ours, for example, in [26, 41, 42], but without (i) our specific details and (2) our generality. For further clarification we address below important particular cases.

III. TSALLIS’ MAXENT VARIATIONAL PROBLEM REVISITED

Since the Lagrange multipliers are the focus of the problems we are trying to solve, we change notations and call them simply $a, b$. We have for $S_T$

$$\delta \left( \int dx \left[ \frac{f - f^q}{q-1} + bUf + af \right] \right) = 0, \quad (31)$$

and then

$$q f^{q-1} = 1 - (q - 1)(a + bU), \quad (32)$$

so that Tsallis’ canonical MaxEnt distribution $f$ with linear constraints is

$$f = \left[ \frac{1 - [(q - 1)(a + bU(x))]^{1/(q-1)}}{q} \right], \quad (33)$$

with $a, b$ Lagrange multipliers, $b$ the inverse temperature $T$. The first Law states that

$$\frac{dS}{db} = b \frac{d < U >}{db}. \quad (34)$$
Now set

$$G = (1 - [(q - 1)(a + bU)])^{(2-q)/(q-1)},$$  \hspace{1cm} (35)$$

$$Q(q) = (\frac{1}{q})^{1/(q-1)},$$  \hspace{1cm} (36)$$

$$D = \frac{1}{q - 1},$$  \hspace{1cm} (37)$$

$$K = \left[ \frac{da}{db} + U \right],$$  \hspace{1cm} (38)$$

entailing

$$(df/db) = QDGK,$$  \hspace{1cm} (39)$$

and, because of $f$–normalization, we derive the fundamental relation

$$QD \int dxGK = 0.$$  \hspace{1cm} (40)$$

Tsallis entropy is

$$S = \frac{1 - \int dx f^q}{q - 1},$$  \hspace{1cm} (41)$$

so that

$$\frac{dS}{db} = -Dq \int dx f^{q-1} DGK,$$  \hspace{1cm} (42)$$

but, since

$$f^{q-1} = \left( \frac{1}{q} \right) \left( 1 - [(q - 1)(a + bU)] \right).$$  \hspace{1cm} (43)$$

Accordingly,

$$\frac{dS}{db} = -QD^2 \int dx \left[ (1 - [(q - 1)(a + bU)]) GK, \right.$$

$$\left. -QD^2 \int dx [GK[(q - 1)(a + bU)]], \right.$$  \hspace{1cm} (44)$$

that we decompose so as to take advantage of Eq. (40),

$$\frac{dS}{db} = QD^2 \int dx [GK[(q - 1)(a + bU)]],$$  \hspace{1cm} (45)$$

and re-using Eq. (40)

$$\frac{dS}{db} = QD \int dx bUGK.$$  \hspace{1cm} (46)$$

Now:
\[
\frac{d < U >}{db} = \int dx U \frac{df}{db},
\]
(47)

that is
\[
\frac{d < U >}{db} = QD \int dx U G K.
\]
(48)

Comparing Eq. (48) with Eq. (46) we see that
\[
\frac{dS}{db} = b \frac{d < U >}{db}.
\]
(49)

QED.

IV. THE CASE OF RENYI’S ENTROPY

Renyi’s quantifier \( S_R \) is an important quantity in several areas of scientific effort. One can cite as examples ecology, quantum information, the Heisenberg XY spin chain model, theoretical computer science, conformal field theory, quantum quenching, diffusion processes, etc. [30–39], and references therein. An important Renyi-characteristic lies in its lack of trace form. We have

\[
S_R = \frac{1}{1 - q} \ln \left( \int f^q dx \right) = \frac{1}{1 - q} \ln J,
\]
(50)

\[
J = \int f^q dx,
\]
(51)

with

\[
\frac{dJ}{db} = q \int f^{q-1} \frac{df}{db} dx.
\]
(52)

The variational problem is

\[
\delta \left( \ln \left\{ \int dx \frac{f^q}{(1 - q)} \right\} - \int dx [bU f + a f] \right) = 0,
\]
(53)
yielding

\[
\frac{q f^{q-1}}{J(1 - q)} - bU - a = 0,
\]
(54)
i.e.,

\[
f^{q-1} = \frac{J(1 - q)}{q} [a + bU],
\]
(55)

and

\[
f = \left( \frac{J(1 - q)}{q} [a + bU] \right)^{1/(q-1)},
\]
(56)
is the MaxEnt solution, with $a$, $b$ Lagrange multipliers, $b$ the inverse temperature $T$.

\[
\frac{df}{db} = \frac{1}{q-1} \left( \frac{J(q-1)}{q} [a + bU] \right) \left( \frac{1}{q} \frac{da}{db} + U \right) + \frac{1}{q} [(a + bU)(q-1) \frac{dJ}{db}] ,
\]

(57)

\[
G = \left( \frac{J(1-q)}{q} [a + bU] \right)^{(2-q)/(q-1)},
\]

(58)

\[
K = \left( \frac{J(1-q)}{q} \frac{da}{db} + U \right) + \frac{1}{q} [(a + bU)(q-1) \frac{dJ}{db}] .
\]

(59)

Now:

\[
\frac{df}{db} = \frac{1}{q-1} GK,
\]

(60)

so that, on account of $f$–normalization, we derive the fundamental relation

\[
\int dx GK = 0.
\]

(61)

The first Law states that

\[
\frac{dS_R}{db} = b \frac{d < U >}{db},
\]

(62)

with

\[
\frac{d < U >}{db} = \int dx U \frac{df}{db},
\]

(63)

that is

\[
\frac{d < U >}{db} = \frac{1}{q-1} \int dx U G K.
\]

(64)

According to (50)

\[
\frac{dS_r}{db} = \frac{q}{J(1-q)} \int dx f^{q-1} \frac{df}{db},
\]

(65)

i.e.,

\[
\frac{dS_r}{db} = \frac{q}{J(1-q)} \int dx \frac{df}{db} \frac{J(1-q)}{q} [a + bU],
\]

(66)

or,

\[
\frac{dS_r}{db} = \frac{1}{q-1} \int dx [a + bU] G K,
\]

(67)

and using (61)
Comparing with (64) we obtain
\[
\frac{dS_R}{db} = b \frac{d < U >}{db},
\] (69)

QED.

V. GENERAL ENTROPIES OF TRACE FORM

\[
S = \left[ \int dx R[f(x)] \right],
\] (70)

with \( R \) an arbitrary smooth function. Then
\[
S' = \int dx R'.
\] (71)

(Here \( R' \) denotes the functional derivative). The MaxEnt variational problem is
\[
\delta \left[ \int dx (R - af - bUf) \right] = 0.
\] (72)

\[
\left[ \int dx (R' - a - bU) \right] = 0.
\] (73)

\[
R' = a + bU
\] (74)

Define now the inverse function if \( R' \)
\[
g = (R')^{(-1)}; \text{ so that } g[R'] = f = R'[g].
\] (75)

\[
f = g[a + bU]
\] (76)

\[
\int dy g[a + bU] = 1.
\] (77)

\[
\frac{d}{db} \int dx g[a + bU] = 0.
\] (78)

\[
\int dx g'[a + bU] \frac{da}{db} + U = 0.
\] (79)

\[
\frac{d < U >}{db} = \int dx g'[a + bU] \frac{da}{db} + U. \] (80)
We use now \(75\) to set \(R'[g] = a + bU\), and then
\[
\frac{dS}{db} = \int dx \frac{dR}{db} = \int dx R'[g] \frac{dg}{db} = \int dx (a + bU) g'[\frac{da}{db} + U]. \tag{81}
\]
We use now normalization \(79\) and obtain the fundamental relation
\[
\frac{dS}{db} = \int dx (a + bU) \frac{dg}{db} \frac{d}{db} = \int dx bU g'\frac{da}{db} + U, \tag{82}
\]
so that comparing \(80\) with \(82\) we satisfy the first Law.

VI. GENERAL ENTROPIES LACKING TRACE FORM

\[
S = B \left[ \int dx R[f] \right], \tag{83}
\]
with \(B\) an arbitrary smooth functional. Define the number \(J = B'[\int dx R(f)]\).
\[
\frac{dS}{db} = J \int dx R'(df/db). \tag{84}
\]
(Here \(S'\) denotes the functional derivative). Define \(F = R'\) and consider the inverse function of \(F\), namely,
\[
g = F^{(-1)}; \quad F[g(f)] = f; \quad g[F(f)] = f. \tag{85}
\]
The MaxEnt variational problem ends up being
\[
JF(f) - a - bU = 0, \tag{86}
\]
so that the MaxEnt solution’s PD \(f_{ME}\) is
\[
f = g\{(a + bU)/J\}, \tag{87}
\]
and the MaxEnt entropy reads
\[
S_{ME} = B \left[ \int dx R[g\{(a + bU)/J\}] \right]. \tag{88}
\]
One also has
\[
0 = \frac{d}{db} \int dx f. \tag{89}
\]
\[
\int dx g'\{(a + bU)/J\} \left\{ \frac{\partial a}{\partial b} + U, J^{-1} - (1/J^2)(dJ/db)[a + bU] \right\} = 0, \tag{90}
\]
Now,
\[
\frac{d < U >}{db} = \int dx U g'\{(a + bU)/J\} \left\{ \frac{\partial a}{\partial b} + U, J^{-1} - (1/J^2)(dJ/db)[a + bU] \right\}. \tag{91}
\]
We will use now (85) to set $F'[g] = (a + bU)/J$.

$$
\frac{dS}{db} = J \int dx R'[\{g(a + bU)/J\}]g'[(a + bU)/J] = J \int dx (a + bU)J^{-1}g'\left\{\left[\frac{\partial a}{\partial b} + U\right]J^{-1} - \left(1/J^2\right)(dJ/db)[a + bu]\right\}.
$$

(92)

We use now $f$–normalization and derive the fundamental relation

$$
\frac{dS}{db} = J \int dx [(a + bU)/J]g'[(a + bU)/J] \left\{\left[\frac{\partial a}{\partial b} + U\right]J^{-1} - \left(1/J^2\right)(dJ/db)[a + bu]\right\},
$$

(93)

so that

$$
\int dx bUg'[(a + bU)/J] \left\{\left[\frac{\partial a}{\partial b} + U\right]J^{-1} - \left(1/J^2\right)(dJ/db)[a + bu]\right\},
$$

(94)

so that comparing (91) with (94) we satisfy the first Law.

VII. CONCLUSIONS

We have conclusively shown that the first law $\frac{dS}{dT} = \beta \frac{d\langle U\rangle}{dT}$ is obeyed by any system subject to a Legendre extremization process, i.e., in any constrained entropic variational problems, no matter what form the entropy adopts and what kind of constraints are used, We will demonstrate the fact that Eq. (1) always holds, no matter what the quantifier one has in mind might be. The essential tool of our proofs is a judicious use of the normalization requirement.

Note that the treatment of Section IIB encompasses the three different forms of non-linear averaging that have been proposed for Tsallis’ statistics in [43].
