Available online at www.sciencedirect.com
SCIENCE DIRECT
Physica A 365 (2006) 76–84
PHYSICA A
www.elsevier.com/locate/physa

Towards information theory for q-nonextensive statistics without q-deformed distributions

Petr Jizba[a,*], Toshihico Arimitsu[b]

[a]*FNSPI, Czech Technical University, Břehová 7, 115 19 Praha 1, Czech Republic*
[b]*Institute of Physics, Tsukuba University, Ibaraki 305-8571, Japan*

Available online 3 February 2006

Abstract

In this paper we extend our recent results [P. Jizba, T. Arimitsu Physica A 340 (2004) 110] on q-nonextensive statistics with non-Tsallis entropies. In particular, we combine an axiomatics of Rényi with the q-deformed version of Khinchin axioms to obtain the entropy which accounts both for systems with embedded self-similarity and q-nonextensivity. We find that this entropy can be uniquely solved in terms of a one-parameter family of information measures. The corresponding entropy maximizer is expressible via a special function known under the name of the Lambert W-function. We analyze the corresponding "high" and "low-temperature" asymptotics and make some remarks on the possible applications.

Keywords: Information theory; Rényi's information measure; Tsallis–Havrda–Charvát entropy; MaxEnt principle; Lambert W-function

1. Introduction

The idea that Gibbsian statistical thermodynamics and Shannon's communication theory share the same line of reasoning was originally introduced by Edwin Jaynes in his influential 1957 papers [1]. There he proposed the *Maximum Entropy Principle* (MaxEnt) as a general inference procedure with a direct relevance to statistical physics. A standard frame of statistical thermodynamics appeared as soon as the notion of entropy was introduced. In particular, Jaynes's MaxEnt utilized Shannon's entropy, or better, Shannon's information measure ($= S$), as an inference functional. The central rôle of Shannon's entropy as a tool for inductive inference (i.e., inference where prior information are given in terms of expectation values) was further demonstrated in works of Faddeyev [2], Shore and Johnson [3], Wallis [4] and others. Following Jaynes, one should view the MaxEnt distribution (or maximizer) as a distribution that is maximally noncommittal with regard to missing information and that agrees with all what is known about prior information, but expresses maximum uncertainty with respect to all other matters [1].

With the advancement in information theory it has become clear that Shannon's entropy is not the only feasible information measure. Indeed, many modern communication processes, including signals, images and coding systems, often operate in complex environments dominated by conditions that do not match the basic

*Corresponding author. Tel.: +420 224 35 8320; fax: +420 222 32 0861.
E-mail addresses: p.jizba@fjfi.cvut.cz (P. Jizba), arimitsu@cm.ph.tsukuba.ac.jp (T. Arimitsu).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2006.01.038

tenets of Shannon's communication theory. For instance, buffer memory (or storage capacity) of a transmitting channel is often finite, coding can have a non-trivial cost function, codes might have variable-length codes, sources and channels may exhibit memory or losses, etc. Post-Shannon developments of information theory offer various generalized measures of information to deal with such situations. Measures of Havrda-Charvát [5], Sharma-Mittal [6], Rényi's [7] and Kapur's [8] can serve as examples.

If the parallel between information theory and statistical physics has a deeper reason, as advocated by Jaynes, then one should expect similar progress also in statistical physics. Indeed, in the past 20 years, physicists have begun to challenge the assumptions of Gibbs's statistics such as ergodicity or metric transitivity. This happened when evidence accumulated showing that there are many situations of practical interest requiring statistics which do not conform with Gibbs's exponential maximizers. Examples include percolation, cosmic rays, turbulence, granular matter, clustered volatility, etc.

When trying to generalize Gibbs's entropy, the information-theoretic parallel with statistical thermodynamics provides a useful conceptual guide. The natural strategy that fits this framework is to revisit the axiomatic rules governing Shannon's entropy and potential extensions translate into a language of statistical physical. The usual axiomatics of Khinchin [9] is prone to several cogent generalizations. Among those, the additivity of independent mean information is a natural axiom to attack. In this way, two fundamentally distinct generalization schemes have been pursued in the literature; one redefining the statistical mean and another generalizing the additivity rule. While the first leads to Rényi's entropies [7,10] that are nature tool in systems with embedded self-similarity [11], the second scheme yields various deformed entropies [12] that play important rôle in long-range/time correlated systems.

It is to be expected that a suitable merger of the above generalizations could provide a new conceptual frame suitable for a statistical description of systems possessing both self-similarity and long-range correlations. Such systems are quite pertinent with examples spanning from the early cosmological phase transitions to currently much studied quantum phase transitions (frustrated spin systems, Fermi liquids, etc.). Our aim was to study one particular merger, namely merger of Rényi and Tsallis–Havrda–Charvát (THC) entropies.

The structure of this paper is the following: in Section 2 axiomatics of Rényi and THC entropies are reviewed. In Section 3 we formulate a new axiomatics which aims at unifying the Rényi and THC entropies. Such an axiomatics allows for only one one-parameter family of information measures. Basic properties of this new class of entropies are discussed in Section 4. The ensuing maximizer is calculated in Section 5. There we show that MaxEnt distribution is expressible through the Lambert W-function. We analyze the corresponding "high" and "low-temperature" asymptotics and discuss the corresponding non-trivial structure of the parameter space. A final discussion is given in Section 6.

2. Rényi's and THC entropies—axiomatic viewpoint

As already said, RE represents a step towards more realistic situations encountered in information theory. Since RE's have a firm operational characterization given in terms of block coding and hypotheses testing (see, e.g., Ref. [13]), it can be directly measured. This is typically happening, e.g., in communication systems with the buffer overflow problem or in variable-length coding with an exponential cost constraint. RE's are also indispensable in various branches of physics that require self-similar sample spaces. Examples being chaotic dynamical systems or multifractals. RE of order q that is assigned to a discrete distribution $\mathscr{P} = \{p_1, \dots, p_n\}$ is defined as

$$\mathscr{I}_q(\mathscr{P}) = \frac{1}{(1-q)} \ln\left(\sum_{k=1}^{n} (p_k)^q\right), \quad q > 0. \tag{1}$$

For simplicity's sake we use the base e of natural logarithms. RE thus defined is then measured in natural units—nats, rather than bits.[1]

[1] To convert, note that 1 bit = 0.693 nats.

In his original work Rényi [10] introduced a one-parameter family of information measures (=RE) which he based on axiomatic considerations. His axioms have been further sharpened by Daróczy [14] and others [15]. It has been recently shown in Ref. [11] that RE can be conveniently characterized by the following set of axioms:

(1) For a given integer n and given $\mathscr{P} = \{p_1, p_2, \ldots, p_n\}$ ($p_k \geqslant 0, \sum_k^n p_k = 1$), $\mathscr{I}(\mathscr{P})$ is a continuous with respect to all its arguments.
(2) For a given integer n, $\mathscr{I}(p_1, p_2, \ldots, p_n)$ takes its largest value for $p_k = 1/n$ ($k = 1, 2, \ldots, n$) with the normalization $\mathscr{I}\left(\frac{1}{2}, \frac{1}{2}\right) = \ln 2$.
(3) For a given $q \in \mathbb{R}$; $\mathscr{I}(A \cap B) = \mathscr{I}(A) + \mathscr{I}(B|A)$ with $\mathscr{I}(B|A) = g^{-1}(\sum_k \varrho_k(q) g(\mathscr{I}(B|A = A_k)))$, and $\varrho_k(q) = p_k^q / \sum_k p_k^q$ with $p_k = \mathscr{P}(A_k)$.
(4) g is invertible and positive in $[0, \infty)$.
(5) $\mathscr{I}(p_1, p_2, \ldots, p_n, 0) = \mathscr{I}(p_1, p_2, \ldots, p_n)$, i.e., adding an event of probability zero (impossible event) we do not gain any new information.

These axioms markedly differ from those utilized in Refs. [10,14,15]. Important distinction is the emergence of the zooming (or escort) distribution $\varrho(q)$ in axiom 3. Note also that RE of two independent experiments is additive. In fact, it was proved in Ref. [10] that RE is the most general information measure compatible with additivity of independent information and the Kolmogorov system of probability.

Among variety of deformed entropies the currently popular one is the q-deformed Shannon's entropy, better known as THC entropy. As the classical additivity of independent information is not valid there, one may infer that the typical playground for THC entropy should be in systems with non-vanishing long-range/time correlations: e.g., in statistical systems with quantum non-locality or in various option-price models. In the case of discrete distributions $\mathscr{P} = \{p_1, \ldots, p_n\}$ THC entropy takes the form:

$$\mathscr{S}_q(\mathscr{P}) = \frac{1}{(1-q)}\left[\sum_{k=1}^{n} (p_k)^q - 1\right], \quad q > 0. \tag{2}$$

Axiomatic treatment was recently proposed in Ref. [16] and it consists of four axioms

(1) For a given integer n and given $\mathscr{P} = \{p_1, p_2, \ldots, p_n\}$ ($p_k \geqslant 0, \sum_k^n p_k = 1$), $\mathscr{S}(\mathscr{P})$ is a continuous with respect to all its arguments.
(2) For a given integer n, $\mathscr{S}(\mathscr{P})$ takes its largest value for $p_k = 1/n$ ($k = 1, 2, \ldots, n$).
(3) For a given $q \in \mathbb{R}$; $\mathscr{S}(A \cap B) = \mathscr{S}(A) + \mathscr{S}(B|A) + (1-q)\mathscr{S}(A)\mathscr{S}(B|A)$ with $\mathscr{S}(B|A) = \sum_k \varrho_k(q)$ $\mathscr{S}(B|A = A_k)$.
(4) $\mathscr{S}(p_1, p_2, \ldots, p_n, 0) = \mathscr{S}(p_1, p_2, \ldots, p_n)$.

As said before, one keeps here the linear mean but generalizes the additivity law. In fact, the additivity law in axiom 3 is nothing but the Jackson sum of the q calculus.

3. Axiomatic merger

As a natural axiomatic merger of previous two axiomatics one can choose

(1) For a given integer n and given $\mathscr{P} = \{p_1, p_2, \ldots, p_n\}$ ($p_k \geqslant 0, \sum_k^n p_k = 1$), $\mathscr{D}(\mathscr{P})$ is a continuous with respect to all its arguments.
(2) For a given integer n, $\mathscr{D}(\mathscr{P})$ takes its largest value for $p_k = 1/n$ ($k = 1, 2, \ldots, n$).
(3) For a given $q \in \mathbb{R}$; $\mathscr{D}(A \cap B) = \mathscr{D}(A) + \mathscr{D}(B|A) + (1-q)\mathscr{D}(A)\mathscr{D}(B|A)$ with $\mathscr{D}(B|A) = f^{-1}\left(\sum_k \varrho_k(q)\right.$ $\left. f(\mathscr{D}(B|A = A_k))\right)$.
(4) f is invertible and positive in $[0, \infty)$.
(5) $\mathscr{D}(p_1, p_2, \ldots, p_n, 0) = \mathscr{D}(p_1, p_2, \ldots, p_n)$.

In Refs. [17,18] it has been shown that the above axioms allow for only one one-parameter class of solutions given by

$$\mathscr{D}_q(\mathscr{A}) = \frac{1}{1-q}(\mathrm{e}^{-(1-q)\sum_k \varrho_k(q)\ln p_k} - 1) = \frac{1}{1-q}\left(\prod_k (p_k)^{-(1-q)\varrho_k(q)} - 1\right). \tag{3}$$

Here $\langle \ldots \rangle_q$ is defined with respect to the distribution $\varrho_k(q)$. We can further recast the relation (3) into another, more convenient form

$$\mathscr{D}_q(\mathscr{A}) = \frac{1}{1-q}\left(\mathrm{e}^{-(1-q)^2\, \mathrm{d}\mathscr{I}_q/\mathrm{d}q} \sum_{k=1}^{n} (p_k)^q - 1\right). \tag{4}$$

Eqs. (3)–(4) represent the sought information measure.

4. Basic properties of $\mathscr{D}_q$

Before studying the implications of the formulas (3)–(4), there is one immediate consequence which warrants special mention. In particular, from the condition $\mathrm{d}\mathscr{I}_q/\mathrm{d}q \leqslant 0$ (see, e.g., Ref. [7]) one has

$$\mathscr{D}_q(A)\begin{cases} \geqslant \mathscr{S}_q(A) & \text{if } q \leqslant 1, \\ \leqslant \mathscr{S}_q(A) & \text{if } q \geqslant 1, \end{cases} \tag{5}$$

with equality, iff $q = 1$ or $\mathrm{d}\mathscr{I}_q/\mathrm{d}q = 0$. This happens only when $\mathscr{P}$ is uniform or trivial, i.e., $\{1, 0, \ldots, 0\}$. By utilizing the known properties of $\mathscr{I}_q$ and $\mathscr{S}_q$ we have

$$\begin{aligned} &0 \leqslant S(\mathscr{P}) \leqslant \mathscr{I}_q(\mathscr{P}) \leqslant \mathscr{S}_q(\mathscr{P}) \leqslant \mathscr{D}_q(\mathscr{P}) \leqslant \ln_q n \quad \text{for } 0 < q \leqslant 1, \\ &0 \leqslant \mathscr{D}_q(\mathscr{P}) \leqslant \mathscr{S}_q(\mathscr{P}) \leqslant \mathscr{I}_q(\mathscr{P}) \leqslant S(\mathscr{P}) \leqslant \ln n \quad \text{for } q \geqslant 1. \end{aligned} \tag{6}$$

This means that by investigating the information measure $\mathscr{D}_q$ with the given $q < 1$ we receive more information than restricting to $\mathscr{I}_q$ or $\mathscr{S}_q$ only. On the other hand, when $q > 1$ then both $\mathscr{I}_q$ and $\mathscr{S}_q$ are more informative than $\mathscr{D}_q$. In practice one usually requires more than one q to gain more complete information about a system. In fact, when entropies $\mathscr{I}_q$ or $\mathscr{S}_q$ are used, it is necessary to know them for all q in order to obtain a full information on a given statistical system [11]. For applications in strange attractors the reader may see Ref. [19], for reconstruction theorems see, e.g., Refs. [7,11].

Let us state here some of the basic characteristics of $\mathscr{D}_q$. Among properties that are common to both Rényi's and THC entropies we find

(a) $\mathscr{D}_q(\mathscr{P} = \{1, 0, \ldots, 0\}) = 0$,
(b) $\mathscr{D}_q(\mathscr{P}) \geqslant 0$,
(c) $\mathscr{D}_q$ is decisive, i.e., $\mathscr{D}_q(0, 1) = \mathscr{D}_q(1, 0)$,
(d) $\mathscr{D}_q$ is expansible, i.e., $\mathscr{D}_q(p_1, \ldots, p_n) = \mathscr{D}_q(0, p_1, \ldots, p_n)$,
(e) $\mathscr{D}_1 = \mathscr{I}_1 = \mathscr{S}_1 = S$,
(f) $\mathscr{D}_q$ involves a single free parameter—q,
(g) $\mathscr{D}_q$ is symmetric, i.e., $\mathscr{D}_q(p_1, \ldots, p_n) = \mathscr{D}_q(p_{k_{(1)}}, \ldots, p_{k_{(n)}})$,
(h) $\mathscr{D}_q$ is bounded.

Among features inherited from Rényi's entropy we can find

(i) $\mathscr{D}_q(A) = f^{-1}(\sum_k \varrho_k(q) f(\mathscr{D}_q(A_k)))$,
(j) $\mathscr{D}_q$ is a strictly decreasing function of q, i.e., $\mathrm{d}\mathscr{D}_q/\mathrm{d}q \leqslant 0$, for any $q > 0$.

Result (i) follows from the fact that $\mathscr{D}_q$ is a monotonically decreasing function of $\langle \ln \mathscr{P} \rangle_q$ and that $\langle \ln \mathscr{P} \rangle_q$ is a monotonically increasing function of q. Finally, properties imprinted from Tsallis entropy include

(k) $\max_{\mathscr{P}} \mathscr{D}_q(\mathscr{P}) = \mathscr{D}_q(\mathscr{P} = \{1/n, \ldots, 1/n\}) = \ln_q n$,
(l) $\mathscr{D}_q$ is q non-extensive, i.e., $\mathscr{D}(A \cap B) = \mathscr{D}(A) + \mathscr{D}(B|A) + (1-q)\mathscr{D}(A)\mathscr{D}(B|A)$.

The issue of thermodynamic stability will be discussed separately in Section 5.1.

5. MaxEnt distributions for $\mathscr{D}_q$

According to information theory, the MaxEnt principle yields distributions which reflect least bias and maximum ignorance about information not provided to a recipient (or observer). Important feature of the usual Gibbsian MaxEnt formalism is that maximizers are all greater than zero and that the maximal entropy is a concave function of the values of the prescribed constraints [20].

Let us first address the issue of the $\mathscr{D}_q$ maximizer. We start by seeking the conditional extremum of $\mathscr{D}_q$ subject to the constraints imposed by the q-averaged value of energy E

$$\langle E \rangle_q = \sum_k \varrho_k(q) E_k. \tag{7}$$

By considering the normalization condition for p_i we should extremize the functional

$$L_q(\mathscr{P}) = \mathscr{D}_q(\mathscr{P}) - \Omega \, \frac{\sum_k (p_k)^q E_k}{\sum_k (p_k)^q} - \Phi \sum_k p_k, \tag{8}$$

with Ω and Φ being the Lagrange multipliers. Setting the derivatives of $L_q(\mathscr{P})$ with respect to $p_1, \ldots, p_n$ to zero, we obtain

$$\begin{aligned} \frac{\partial L_q(\mathscr{P})}{\partial p_i} &= \mathrm{e}^{(q-1)\langle \ln \mathscr{P} \rangle_q} (q(\langle \ln \mathscr{P} \rangle_q - \ln p_i) - 1) \frac{(p_i)^{q-1}}{\sum_k (p_k)^q} \\ &\quad - \Omega q (E_i - \langle E \rangle_q) \frac{(p_i)^{q-1}}{\sum_k (p_k)^q} - \Phi = 0, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{9}$$

Note that when $q \to 1$ then (9) reduces to the usual condition for Shannon's maximizer. This, in turn, ensures that in the $q \to 1$ limit the maximizer boils down to Gibbs's distribution. To proceed we note that Eq. (9) can be cast to the form

$$\Phi (p_i)^{1-q} \sum_k (p_k)^q = \mathrm{e}^{(q-1)\langle \ln \mathscr{P} \rangle_q} (q(\langle \ln \mathscr{P} \rangle_q - \ln p_i) - 1) - q\Omega (E_i - \langle E \rangle_q). \tag{10}$$

By multiplying both sides by $\varrho_i(q)$, summing over i and taking the normalization condition $\sum_k p_k = 1$, we obtain

$$\Phi = -\mathrm{e}^{(q-1)\langle \ln \mathscr{P} \rangle_q} \quad \Rightarrow \quad \frac{\ln(-\Phi)}{q-1} = \langle \ln \mathscr{P} \rangle_q \quad \Rightarrow \quad \mathscr{D}_q(\mathscr{P})|_{\max} = \frac{1}{q-1}(\Phi + 1). \tag{11}$$

Plugging result (11) back into (10) we have

$$\sum_k (p_k)^q = (p_i)^{q-1} \left[q \ln p_i + \left(1 - \frac{q \ln(-\Phi)}{q-1} - \frac{q\Omega}{\Phi}(E_i - \langle E \rangle_q) \right) \right], \tag{12}$$

which must hold for any index i. On the substitution

$$\mathscr{E}_i = 1 - \frac{q \ln(-\Phi)}{q-1} - \frac{q\Omega}{\Phi} \Delta_q E_i, \quad \Delta_q E_i = E_i - \langle E \rangle_q, \tag{13}$$

we finally obtain the equation

$$\kappa(p_i)^{1-q} = q \ln p_i + \mathscr{E}_i, \quad \sum_k (p_k)^q \equiv \kappa. \tag{14}$$

This has the solution

$$p_i = \left[\frac{q}{\kappa(q-1)} W\left(\frac{\kappa(q-1)}{q} \, e^{(q-1)\mathscr{E}_i/q}\right)\right]^{1/(1-q)} = \exp\left\{\frac{W\left(\frac{\kappa(q-1)}{q} \, e^{(q-1)\mathscr{E}_i/q}\right)}{(q-1)} - \mathscr{E}_i/q\right\}, \tag{15}$$

with $W(x)$ being the Lambert W-function [21].

Some comments are now in order. First, p_i's as prescribed by (15) are positive for any value of $q>0$. This is a straightforward consequence of the following two identities [21]:

$$W(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} n^{n-2}}{(n-1)!} \, x^n, \quad W(x) = x e^{-W(x)}. \tag{16}$$

Indeed, the first relation ensures that for $x<0$ also $W(x)<0$ and hence $W(x)/x>0$. Thus for $0<q<1$ the positivity of p_i's is proven. Positivity for $q \geqslant 1$ follows directly from the second relation. Second, as $q \to 1$ the entropy $\mathscr{D}_q \to S$ and hence p_i's defined by (15) approaches the Gibbs distribution. To see that this is the case, let us realize that

$$\Phi|_{q=1} = -1, \quad \mathscr{E}_i|_{q=1} = 1 + \mathscr{H} + \Omega(E_i - \langle E \rangle) \quad \text{and} \quad \kappa|_{q=1} = 1. \tag{17}$$

Then

$$p_i|_{q=1} = \exp(1 - (1 + \mathscr{H} + \Omega(E_i - \langle E \rangle))) = \exp(\Omega F - \Omega E_i) = e^{-\Omega E_i}/Z, \tag{18}$$

which after identification $\Omega|_{q=1} = \beta$ leads to the desired result. Note also that (15) is invariant under uniform translation of the energy spectrum, i.e., the corresponding p_i is insensitive to the choice of the ground-state energy. Third, there does not seem to be any simple method for determining Φ and Ω in terms of $\langle E \rangle_q$. In fact, only asymptotic situations for large and vanishingly small Ω can be successfully tackled. For this purpose we briefly remark on the asymptotic behavior of p_i in regard to Ω.

We first assume that $\Omega \ll 1$—"*high-temperature expansion*"—then from (16) follows

$$W\left(\frac{\kappa(1-q)}{\Phi q} \, e^{(q-1)/q} \exp\left((1-q)\frac{\Omega}{\Phi} \, \Delta_q E_i\right)\right) \approx W(x)[1 - (1-q)\Omega^* \Delta_q E_i],$$

with

$$\Omega^* = -\frac{\Omega}{\Phi(W(x)+1)}, \quad x = -\frac{\kappa(q-1)}{\Phi q} \exp\left(\frac{q-1}{q}\right).$$

The relation (15) then implies that

$$p_i = \frac{[1 - (1-q)\Omega^* \Delta_q E_i]^{1/(1-q)}}{\sum_k [1 - (1-q)\Omega^* \Delta_q E_k]^{1/(q-1)}} = Z^{-1}[1 - (1-q)\Omega^* \Delta_q E_i]^{1/(1-q)}, \tag{19}$$

with the partition function

$$Z = \sum_k [1 - (1-q)\Omega^* \Delta_q E_k]^{1/(1-q)} = \left[\frac{q}{\kappa(q-1)} \, W(x)\right]^{1/(q-1)}. \tag{20}$$

The distribution (19) agrees with the so-called third version of thermostatics introduced by Tsallis et al. [22]. It can by also formally identified with the maximizer for RE [23]. Clearly, Ω^* is not a Lagrange multiplier, but Ω^* passes to β at $q \to 1$ (in fact, $\Phi \to -1$, $\Omega \to \beta$ and $W(x) \to 0$ at $q \to 1$). Note also that when $\Omega = 0$ (i.e., no energy constraint) then $p_i = 1/n$ which reconfirms that $\mathscr{D}_q$ reaches its maximum for uniform distribution.

From the physical standpoint it is the asymptotic behavior at $\Omega|(q-1)/\Phi| \gg 1$—"*low-temperature*" *expansion*—that is most intriguing. This is because the branching properties of the Lambert W-function at negative argument make the structure of $\mathscr{P}$ non-trivial. To this end one can distinguish four distinct situations

$$(a_1)\ (q-1)>0 \text{ and } \Delta_q E_i<0, \quad (a_2)\ (q-1)>0 \text{ and } \Delta_q E_i>0,$$
$$(b_1)\ (q-1)<0 \text{ and } \Delta_q E_i<0, \quad (b_2)\ (q-1)<0 \text{ and } \Delta_q E_i>0.$$

Cases (a_1) and (a_2) are much simpler to start with as the argument of W is positive. W is then a real and single valued function which belongs to the principal branch of W known as W_0. When $\Delta_q E_i<0$ then (a_1) implies $W(z) \approx z$ and hence

$$p_i = \left(\frac{1}{|\Phi|}\right)^{1/(1-q)} \mathrm{e}^{-1/q} \exp\left(-\frac{\Omega}{|\Phi|}\, \Delta_q E_i\right) = Z_1^{-1} \exp\left(-\frac{\Omega}{|\Phi|}\, \Delta_q E_i\right), \tag{21}$$

with

$$Z_1 = \left(\frac{1}{|\Phi|}\right)^{1/(q-1)} \mathrm{e}^{1/q}.$$

Note that in this case p_i is of a Boltzmann type. On the other hand, the (a_2) situation implies the asymptotic expansion [21]:

$$W(z) \approx \ln(z) - \ln(\ln(z)) \ \Rightarrow\ p_i = Z_2^{-1}[1-(1-q)\Omega^* \Delta_q E_i]^{1/(1-q)}, \tag{22}$$

with

$$Z_2 = \left[\frac{q}{\kappa(q-1)} \ln\left(\frac{\kappa(q-1)}{|\Phi| q}\, \mathrm{e}^{(q-1)/q}\right)\right]^{1/(q-1)}, \quad \Omega^* = \frac{\Omega}{|\Phi| \ln\left(\frac{\kappa(q-1)}{|\Phi| q} \exp\left(\frac{q-1}{q}\right)\right)}.$$

Although the distribution (22) formally agrees with Tsallis et al. distribution, it cannot be identified with it as Ω^* does not tend to β in the $q \to 1$ limit. In fact, the limit $q \to 1$ is prohibited as it violates the "low-temperature" condition $\Omega|(q-1)/\Phi| \gg 1$. Note particularly that our MaxEnt distribution represents in the "low-temperature" regime a heavy tailed distribution with Boltzmannian outset—behavior typical, e.g., for income distributions. When Ω and $q>1$ are fixed one may find κ and Φ from the normalization condition and sewing condition at $\Delta_q E = 0$. However, because the "low-temperature" approximation does not allow to probe regions with small $\Delta_q E$ one must numerically optimize the sewing by interpolating the forbidden parts of $\Delta_q E$ axis [18].

Cases (b_1) and (b_2) have much richer structure than (a_1) and (a_2). This is due to the negativity of the argument that enters the W function. A remarkable upshot of this is an existence of a strongly suppressive effect in the occupation of the high-energy states. In addition, the suppression appears in two different ways depending on the value of $(1-q)/|\Phi|$. Analogous type of behavior is known in quantum phase transitions [24]. Complete discussion of this phenomenon will be presented in Ref. [18].

5.1. Thermodynamic stability—concavity issue

In the following we are going to address the issue of thermodynamic stability. Note that in contrast to information-theoretic entropy $\mathscr{D}_q$, $\mathscr{D}_q|_{\max}$ is the system entropy, i.e., it depends on the system state variables. Thermodynamic stability then consists of showing that $\mathscr{D}_q|_{\max}$ is a concave function of the energy constraint [20]. So we wish to show that

$$\frac{\partial^2 \mathscr{D}_q(\mathscr{P})|_{\max}}{\partial \langle E \rangle_q^2} = \frac{\partial^2 \mathscr{D}_q(\langle E \rangle_q)}{\partial \langle E \rangle_q^2} \leqslant 0. \tag{23}$$

This can be done by observing that [21]

$$\frac{\mathrm{d}W(x)}{\mathrm{d}x} = \frac{W(x)}{x(W(x)+1)} \quad \text{and} \quad \frac{\mathrm{d}^2W(x)}{\mathrm{d}x^2} = -\frac{W(x)^2(W(x)+2)}{x^2(W(x)+1)^3}. \tag{24}$$

If we combine (24) with the fact that $\mathrm{d}\mathscr{D}_q(\mathscr{P})|_{\max}/\mathrm{d}\Phi = 1/(q-1)$ we obtain

$$\frac{\partial^2 \mathscr{D}_q(\mathscr{P})|_{\max}}{\partial \langle E \rangle_q^2} = (1-q)^2 \frac{|\Phi|}{\Omega} \langle \ln \mathscr{P} \rangle_q \leqslant 0. \tag{25}$$

Thus $\mathscr{D}_q$ is thermodynamically stably for any q.

6. Conclusions

In this paper we have reviewed the main aspects of the recently proposed information measure $\mathscr{D}_q$. In contrast to presently popular generalizations based on deformed entropies, we have aimed here at a strictly axiomatic approach. This is because we hold that one cannot proceed the formal generalization of the entropy in physics by ignoring the consistency with information theory. As a rule, axiomatic treatments of information measures have the benefit of a closer passage to operational characterizations and hence to a systematic use in practical applications.

We hope that the proposed axiomatics might serve as a novel playground for q-nonextensive systems with embedded self-similarity. Indeed, our conclusions hint that $\mathscr{D}_q$ could play a relevant rôle in quantum phase transitions and/or in econophysics.

The reader may note that we have not checked $\mathscr{D}_q$ for Lesche's observability criterium [25] (also known as experimental robustness). This is because in our view the use of Lesche's condition as the stability criterion is rather doubtful, see e.g., Refs. [26,27]. In this connection Yamano's local stability criterion [27] would seem more appropriate concept to use. Work along those lines is currently in progress.

Finally we should stress that the presented entropy $\mathscr{D}_q$ has many desirable attributes: like THC entropy it satisfies the nonextensive q-additivity, involves a single parameter q, goes over into S in the limit $q \to 1$, it complies with thermodynamic stability, continuity, symmetry, expansivity, decisivity, etc. On that basis it would appear that both $\mathscr{S}_q$ and $\mathscr{D}_q$ have an equal right to serve as a generalization of statistical thermodynamics.

References

[1] E.T. Jaynes, Phys. Rev. 106 (1957) 171;
E.T. Jaynes, Phys. Rev. 108 (1957) 620.
[2] D.K. Faddeyev, Uspekhi Mat. Nauk 11 (1956) 227.
[3] J.E. Shore, R.W. Johnson, IEEE Trans. Inf. Theory 26 (1980) 26.
[4] E.T. Jaynes, Probability Theory, The Logic of Science, Cambridge University Press, Cambridge, 2003.
[5] J.H. Havrda, F. Charvát, Kybernatica 3 (1967) 30.
[6] B.D. Sharma, D.P. Mittal, J. Math. Sci. 10 (1975) 28.
[7] A. Rényi, Selected Papers of Alfred Rényi, vol. 2 , Akadémia Kiado, Budapest, 1976.
[8] J.N. Kapur, Ind. J. Pure Appl. Math. 17 (1986) 429.
[9] A.I. Khinchin, Mathematical Foundations of Information Theory, Dover, London, 1957.
[10] A. Rényi, Probability Theory, North-Holland, Amsterdam, 1970.
[11] P. Jizba, T. Arimitsu, Ann. Phys. 312 (2004) 17.
[12] see e.g., C. Tsallis, J. Stat. Phys. 52 (1988) 479;
C. Tsallis, Braz. J. Phys. 29 (1999) 1;
J. Naudts, Physica A 316 (2002) 323;
G. Kaniadakis, Phys. Rev. E 66 (2002) 056125.
[13] I. Csiszár, IEEE Trans. Inf. Theory 41 (1995) 26.
[14] Z. Daróczy, Acta Math. Acad. Sci. Hungaricae 15 (1964) 203.
[15] S. Guias, Information Theory with Applications, McGraw-Hill, New York, 1977.
[16] S. Abe, Phys. Lett. A 271 (2000) 74.
[17] P. Jizba, T. Arimitsu, Physica A 340 (2004) 110.

[18] P. Jizba, T. Arimitsu, in progress.
[19] H.G.E. Hentschel, I. Procaccia, Physica D 8 (1983) 435.
[20] H.B. Callen, Thermodynamics and An Introduction to Thermostatistics, Wiley, New York, 1985.
[21] R.M. Corless, G.H. Gonnet, D.E. Hare, D.J. Jeffrey, D.E. Knuth, Adv. Comput. Math. 5 (1996) 329.
[22] C. Tsallis, R.S. Mendes, A.R. Plastino, Physica A 261 (1998) 534.
[23] A.G. Bashkirov, A.D. Sukhanov, J. Exp. Theor. Phys. 95 (2002) 440;
A.G. Bashkirov, Physica A 340 (2004) 153.
[24] S. Sachdev, Quantum Phase Transitions, Cambridge University Press, New York, 1999.
[25] B. Lesche, J. Stat. Phys. 27 (1982) 419.
[26] P. Jizba, T. Arimitsu, Phys. Rev. E 69 (2004) 026128.
[27] T. Yamano, Phys. Lett. A 329 (2004) 268.