On the most concise set of axioms and the uniqueness theorem for Tsallis entropy

2002 J. Phys. A: Math. Gen. 35 10731

(http://iopscience.iop.org/0305-4470/35/50/304)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 10731–10738 PII: S0305-4470(02)38518-4

On the most concise set of axioms and the uniqueness theorem for Tsallis entropy

Hiroki Suyari

Department of Information and Image Sciences, Faculty of Engineering, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522 Japan

Received 19 June 2002
Published 4 December 2002
Online at stacks.iop.org/JPhysA/35/10731

Abstract
We present the most concise set of axioms for Tsallis entropy, and rigorously prove the uniqueness theorem. This set of axioms consists of only two distinct additivities: pseudoadditivity and Shannon additivity. We then compare our axioms with the axioms presented by Santos. The peculiarity of pseudoadditivity as an axiom for Tsallis entropy is also discussed.

PACS number: 89.70.+c
Mathematics Subject Classification: 94A17, 54C70

1. Introduction

The introduction of Tsallis entropy, a one-parameter generalization of Shannon entropy, has led to a successful generalization of Boltzmann–Gibbs statistical mechanics especially for the analysis of long-range correlations in space or time, multifractal systems and physical systems emerging from scale-invariant structure [1]. In these studies, the theoretical results derived from the generalization provide us with some consistent interpretations for the actual nonlinear behaviours such as chaotic dynamics [1, 2]. For a probability distribution $(p_1, \ldots, p_n)$ $\left(\sum_{i=1}^n p_i = 1, p_i \geqslant 0, i = 1, \ldots, n\right)$ and $q \in R^+$, Tsallis entropy is defined by

$$S_q(p_1, \ldots, p_n) \equiv k \frac{1 - \sum_{i=1}^n p_i^q}{q - 1} \tag{1.1}$$

where k is a positive constant. Tsallis entropy recovers Shannon entropy when q approaches 1:

$$S_q = k \frac{1 - \sum_{i=1}^n p_i^q}{q - 1} \quad \xrightarrow{q \to 1} \quad S_1 = -k \sum_{i=1}^n p_i \ln p_i. \tag{1.2}$$

Shannon entropy and Tsallis entropy are bases of traditional Boltzmann–Gibbs statistical mechanics [3] and its generalization [1, 4, 5], respectively. Therefore, the above correspondence (1.2) between the two cases $q \in R^+$ and $q = 1$ is very important in the sense that generalized Boltzmann–Gibbs statistical mechanics based on Tsallis entropy recovers the Boltzmann–Gibbs statistical mechanics based on Shannon entropy.

For example, the equilibrium state in generalized Boltzmann–Gibbs statistical mechanics recovers the well-known equilibrium state when $q \to 1$. The generalized equilibrium state $\left\{p_i^{(q)}\right\}$ can be obtained by maximizing Tsallis entropy (1.1) with the energy constraint $\langle\langle\varepsilon_i\rangle\rangle_q = U_q$ as follows (see [6]):

$$p_i^{(q)} = \frac{\exp_q[-\beta(\varepsilon_i - U_q)]}{Z_q} \qquad i = 1, \ldots, n \tag{1.3}$$

where $\langle\langle\cdot\rangle\rangle_q$ is the normalized q-expectation defined by

$$\langle\langle X\rangle\rangle_q \equiv \frac{\sum_{i=1}^n p_i^q x_i}{\sum_{j=1}^n p_j^q}. \tag{1.4}$$

$\exp_q$ and Z_q are the q-exponential function and the generalized partition function, respectively given by

$$\exp_q[x] \equiv [1 + (1-q)x]^{\frac{1}{1-q}} \tag{1.5}$$

$$Z_q \equiv \sum_{i=1}^n \exp_q[-\beta(\varepsilon_i - U_q)] \tag{1.6}$$

and β is the Lagrange parameter. The q-exponential function $\exp_q$ recovers the exponential function when q goes to 1:

$$\exp_q[x] = [1 + (1-q)x]^{\frac{1}{1-q}} \quad \stackrel{q\to 1}{\longrightarrow} \quad \exp[x] \tag{1.7}$$

so that all of the above mathematical formulae (1.3)–(1.6) in the foundations of generalized statistical mechanics recover the well-known formulae in traditional Boltzmann–Gibbs statistical mechanics. The correspondence between $q \in R^+$ and $q = 1$ is found everywhere in the formulae of generalized statistical mechanics. These generalized formulae using appropriate $q \in R^+$ for each application enable us to present nice consistent interpretations for each observed dynamics [1, 2].

All of these formulae are derived from the definition of the entropies. Thus, according to the above correspondence (1.2), two sets of axioms for the two entropies, Tsallis entropy and Shannon entropy, should have the following correspondence:

$$\begin{array}{l}\text{a set of axioms}\\ \text{for Tsallis entropy } S_q\end{array} \quad \stackrel{q\to 1}{\longrightarrow} \quad \begin{array}{l}\text{a set of axioms}\\ \text{for Shannon}\\ \text{entropy } S_1\end{array} \begin{cases}\text{[SK1] continuity}\\ \text{[SK2] maximality}\\ \text{[SK3] additivity}\\ \text{[SK4] expandability.}\end{cases} \tag{1.8}$$

Along the lines of this correspondence, Santos presents a set of axioms for Tsallis entropy [7]. Among his four axioms, two additivities, pseudoadditivity [1] and Shannon additivity [5], are included, and expandability is missing. Shannon additivity corresponds to the additivity [SK3] in (1.8), and pseudoadditivity is a characteristic property of Tsallis entropy [4]. These two additivities appear to be similar in the sense that both are given by the formulations of the entropy $S_q(AB)$ for the composite systems A and B, but they are actually different. When two systems, A and B, are mutually independent, in extensive systems ($q = 1$), pseudoadditivity coincides with Shannon additivity, but in nonextensive systems ($q \in R^+$), this coincidence generally does not hold.

In this paper, we show that pseudoadditivity and Shannon additivity can constitute the most concise set of axioms for Tsallis entropy. In other words, we rigorously prove that

Tsallis entropy is uniquely determined by these two distinct additivities only. This result is in agreement with the requirement that a set of axioms for a function should be as concise as possible. Based on our result, we compare our axioms with Santos' axioms [7] for Tsallis entropy. Taking into account our axioms and the uniqueness theorem, the four axioms in [7] are obviously redundant for unique determination of Tsallis entropy. We explain in detail the factor leading to the redundancy in his axioms by comparing Santos' axioms with the Shannon–Khinchin axioms [SK1]–[SK4]. As a result, our axioms are the most concise of all conceivable axioms, but they are not a generalization of the Shannon–Khinchin axioms along the lines of the correspondence (1.8) because of the use of pseudoadditivity as an axiom. In other words, it is found that Shannon additivity rather than pseudoadditivity should be used as an axiom for the generalization of the Shannon–Khinchin axioms.

2. The most concise set of axioms and the uniqueness theorem for Tsallis entropy

Let Δ_n be defined by the n-dimensional simplex:

$$\Delta_n \equiv \left\{ (p_1, \ldots, p_n) \,\middle|\, p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1 \right\}. \tag{2.1}$$

Our present set of axioms for Tsallis entropy consists of only two out of the four axioms in [7], as follows.

Let $S_q(p_1, \ldots, p_n)$ be a function defined for any integer $n \in N$, for any positive real number $q \in R^+$ and for all probability distributions $(p_1, \ldots, p_n) \in \Delta_n$. It is assumed that there exist $(p_1, \ldots, p_n) \in \Delta_n$ and $q \in R^+$ such that $S_q(p_1, \ldots, p_n)$ does not take zero. If for any $n \in N$ and $q \in R^+$ this function satisfies the following properties [I] and [II], then $S_q(p_1, \ldots, p_n)$ is uniquely determined to be (1.1):

[I] (pseudoadditivity):

$$\frac{S_q(AB)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (1-q)\frac{S_q(A)}{k}\frac{S_q(B)}{k} \tag{2.2}$$

where A and B are mutually independent finite systems:

$$A = \begin{pmatrix} A_1 & \cdots & A_n \\ p_1^A & \cdots & p_n^A \end{pmatrix} \qquad B = \begin{pmatrix} B_1 & \cdots & B_m \\ p_1^B & \cdots & p_m^B \end{pmatrix} \tag{2.3}$$

$$AB = \begin{pmatrix} A_1B_1 & A_1B_2 & \cdots & A_nB_m \\ p_{11}^{AB} & p_{12}^{AB} & \cdots & p_{nm}^{AB} \end{pmatrix} = \begin{pmatrix} A_1B_1 & A_1B_2 & \cdots & A_nB_m \\ p_1^A p_1^B & p_1^A p_2^B & \cdots & p_n^A p_m^B \end{pmatrix}. \tag{2.4}$$

[II] (Shannon additivity):

$$S_q\left(p_{11}, \ldots, p_{nm_n}\right) = S_q(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i^q S_q\left(\frac{p_{i1}}{p_i}, \ldots, \frac{p_{im_i}}{p_i}\right) \tag{2.5}$$

where two probability distributions $\{p_i : i = 1, \ldots, n\}$ and $\{p_{ij} : i = 1, \ldots, n, j = 1, \ldots, m_i\}$ satisfy

$$p_{ij} \geqslant 0 \qquad p_i = \sum_{j=1}^{m_i} p_{ij} \qquad \sum_{i=1}^{n} p_i = 1 \tag{2.6}$$

for any $i = 1, \ldots, n$ and $j = 1, \ldots, m_i$.

Note that (2.5) is a generalization of the Shannon additivity in [5], which is given by

$$S_q(p_1, \ldots, p_W) = S_q(p_L, p_M) + p_L^q S_q\left(\frac{p_1}{p_L}, \ldots, \frac{p_{W_L}}{p_L}\right) + p_M^q S_q\left(\frac{p_{W_L+1}}{p_M}, \ldots, \frac{p_W}{p_M}\right) \quad (2.7)$$

where $W_L + W_M = W$, $p_L = \sum_{i=1}^{W_L} p_i$ and $p_M = \sum_{i=W_L+1}^{W} p_i$. Comparing (2.5) with (2.7), (2.7) is clearly the special case of $n = 2$ in (2.5). Moreover, in the axioms presented in [7], (2.7) is given as one of the four axioms, but actually (2.5) is applied in the proof of the uniqueness theorem (see formula (22) in [7]). Throughout this paper, we shall use formula (2.5) as Shannon additivity.

The uniqueness theorem for our axioms can be easily proved as follows.

Consider the condition that two systems A and B are mutually independent as in the precondition (2.4) for pseudoadditivity, that is,

$$\text{for any} \quad i = 1, \ldots, n \quad \text{and any} \quad j = 1, \ldots, m \qquad p_{ij}^{AB} = p_i^A p_j^B. \quad (2.8)$$

Here, we take $m_i = m$ for all $i = 1, \ldots, n$. Under condition (2.8), Shannon additivity (2.5) can be simplified to

$$S_q\left(p_1^A p_1^B, \ldots, p_n^A p_m^B\right) = S_q\left(p_1^A, \ldots, p_n^A\right) + \left(\sum_{i=1}^{n} \left(p_i^A\right)^q\right) S_q\left(p_1^B, \ldots, p_m^B\right). \quad (2.9)$$

In accordance with the notation (2.3) and (2.4) for pseudoadditivity (2.2), we introduce the following notation:

$$\text{for any} \quad i = 1, \ldots, n \qquad p_i = p_i^A. \quad (2.10)$$

Thus the obtained Shannon additivity (2.9) can be written as

$$S_q(AB) = S_q(A) + \left(\sum_{i=1}^{n} p_i^q\right) S_q(B). \quad (2.11)$$

Eliminating $S_q(AB)$ from both (2.2) and (2.11) yields

$$S_q(A) + S_q(B) + \frac{1-q}{k} S_q(A) S_q(B) = S_q(A) + \left(\sum_{i=1}^{n} p_i^q\right) S_q(B). \quad (2.12)$$

This equality (2.12) holds for any $(p_1, \ldots, p_n) \in \Delta_n$ and $q \in R^+$. As the existence of nonzero S_q is assumed, Tsallis entropy (1.1) is directly obtained from (2.12) as follows:

$$S_q(A) = S_q(p_1, \ldots, p_n) = k \frac{1 - \sum_{i=1}^{n} p_i^q}{q - 1}. \quad (2.13)$$

Thus the proof is complete.

Note that Renyi entropy $S_q^R(p_1, \ldots, p_n)$ has the following simple relation [4, 5] with Tsallis entropy $S_q(p_1, \ldots, p_n)$:

$$S_q^R(p_1, \ldots, p_n) = \frac{\ln[1 + (1-q) S_q(p_1, \ldots, p_n)/k]}{1 - q}. \quad (2.14)$$

The axioms of Renyi entropy S_q^R are given in [8].

3. Comparison of our axioms with Santos' axioms

3.1. Redundancy in Santos' axioms

In [7], Santos shows a set of axioms for Tsallis entropy as follows.

[S1] $S_q(p_1, \ldots, p_n)$ must be a continuous function of the probabilities p_i ($p_i \in (0, 1)\ \forall i$).

[S2] $S_q(p_1, \ldots, p_n)$ must be a monotonic increasing function of the number of states n, in the case of equiprobability.

[S3] $S_q(p_1, \ldots, p_n)$ must satisfy the pseudoadditivity relation:

$$\frac{S_q(AB)}{k} = \frac{S_q(A)}{k} + \frac{S_q(B)}{k} + (1-q)\frac{S_q(A)}{k}\frac{S_q(B)}{k} \tag{3.1}$$

where A and B are mutually independent systems and k is a positive constant.

[S4] S_q must satisfy the relation:

$$S_q(p_1, \ldots, p_n) = S_q(p_L, p_M) + p_L^q S_q\left(\frac{p_1}{p_L}, \ldots, \frac{p_{n_L}}{p_L}\right) + p_M^q S_q\left(\frac{p_{n_L+1}}{p_M}, \ldots, \frac{p_n}{p_M}\right) \tag{3.2}$$

where $n_L + n_M = n$, $p_L + p_M = 1$ and $p_L = \sum_{i=1}^{n_L} p_i$ and $p_M = \sum_{i=n_L+1}^{n} p_i$.

Under the above four axioms, Santos states that the unique function satisfying all these properties [S1]–[S4] is Tsallis entropy (1.1). As mentioned before, despite presenting axiom [S4], Santos applies the generalized formula (2.5) in his proof.

Comparing Santos' axioms [S1]–[S4] with ours [I]–[II], [I] and [II] coincide with [S3] and [S4], respectively. Thus Santos' axioms are clearly redundant for unique determination of Tsallis entropy.

3.2. The peculiarity of pseudoadditivity as an axiom

The Shannon–Khinchin axioms are well established as axioms for Shannon entropy in Shannon information theory [9]. Therefore, in order to construct axioms for Tsallis entropy according to the correspondence (1.8), they must be the least generalization of the Shannon–Khinchin axioms. Comparing Santos' axioms with the Shannon–Khinchin axioms, the peculiarity of pseudoadditivity as an axiom can be revealed as follows.

Firstly, we briefly review the Shannon–Khinchin axioms. The Shannon–Khinchin axioms [9] are given by the following four conditions:

[SK1] *continuity*. For any $n \in N$ the function $S_1(p)$ is continuous with respect to $p \in \Delta_n$.

[SK2] *maximality*. For given $n \in N$ and for $(p_1, \ldots, p_n) \in \Delta_n$, the function $S_1(p_1, \ldots, p_n)$ takes its largest value for $p_i = \frac{1}{n}$ $(i = 1, \ldots, n)$.

[SK3] *additivity*. If

$$p_{ij} \geqslant 0 \qquad p_i = \sum_{j=1}^{m_i} p_{ij} \quad \text{for any} \quad i = 1, \ldots, n \quad \text{and} \quad j = 1, \ldots, m_i$$

$$\text{and} \quad \sum_{i=1}^{n} p_i = 1 \tag{3.3}$$

then the following equality holds:

$$S_1\big(p_{11}, \ldots, p_{nm_n}\big) = S_1(p_1, \ldots, p_n) + \sum_{i=1}^{n} p_i S_1\left(\frac{p_{i1}}{p_i}, \ldots, \frac{p_{im_i}}{p_i}\right). \tag{3.4}$$

[SK4] *expandability*.

$$S_1(p_1, \ldots, p_n, 0) = S_1(p_1, \ldots, p_n). \tag{3.5}$$

Comparing Santos' axioms [S1]–[S4] with the above Shannon–Khinchin axioms [SK1]–[SK4], [S1], [S2] and [S4] correspond to [SK1], [SK2] and [SK3], respectively. However, pseudoadditivity given in [S3] does not correspond to any axiom in the Shannon–Khinchin axioms. Moreover, an axiom corresponding to expandability given in [SK4] is missing in Santos' axioms. In other words, when we consider Santos' axioms as a generalization of the Shannon–Khinchin axioms, the expandability given in [SK4] is replaced by the pseudoadditivity given in [S3] in Santos' axioms.

In nonextensive systems, there exist two additivities: pseudoadditivity (2.2) and Shannon additivity (2.5). When $q \to 1$, it can be easily verified that pseudoadditivity and Shannon additivity coincide with each other. However, when $q \neq 1$ (nonextensive systems), they are completely different from each other in the sense that one cannot be derived from the other. Even if the associated systems A and B are mutually independent, each of (2.2) and (2.11) cannot be derived from the other. Therefore, we can obtain Tsallis entropy (1.1) uniquely from these two distinct additivities in the previous section. But the set of axioms [I] and [II] is not clearly a generalization of the Shannon–Khinchin axioms.

Then, for a natural generalization of the Shannon–Khinchin axioms to nonextensive systems, we need to consider which additivity is suitable to be one of the axioms of Tsallis entropy. In the Shannon–Khinchin axioms, Shannon additivity (3.4) is given as additivity. Thus, we need to consider two kinds of comparison: '(3.4) versus (2.5)', and '(3.4) versus (2.2)'.

For the former comparison: (3.4) versus (2.5), the expectation '$\sum_{i=1}^{n} p_i \times$' on the right-hand side of (3.4) is generalized to '$\sum_{i=1}^{n} p_i^q \times$' on the right-hand side of (2.5).

$$\text{extensive } (q = 1): \sum_{i=1}^{n} p_i \times \quad \overset{\text{generalized}}{\longrightarrow} \quad \text{nonextensive } (q \in R^+): \sum_{i=1}^{n} p_i^q \times \tag{3.6}$$

On the other hand, for the latter comparison: (3.4) versus (2.2), consider the same condition such that the two systems A and B are mutually independent. Under such a condition in extensive systems ($q = 1$), (3.4) implies that

$$S_1(AB) = S_1(A) + S_1(B). \tag{3.7}$$

Equation (3.7) can be obtained by substituting $q = 1$ into (2.11) as seen in the previous section. Compare (3.7) in extensive systems with (2.2) in nonextensive systems. For simplicity, we consider the case $k = 1$. The sum of two entropies $S_1(A) + S_1(B)$ on the right-hand side of (3.7) is generalized to $S_q(A) + S_q(B) + (1 - q)S_q(A)S_q(B)$ on the right-hand side of (2.2):

$$\text{extensive } (q = 1): S_1(A) + S_1(B)$$
$$\overset{\text{generalized}}{\longrightarrow} \text{nonextensive } (q \in R^+): S_q(A) + S_q(B) + (1 - q)S_q(A)S_q(B) \tag{3.8}$$

Compare the two generalizations (3.6) and (3.8). Clearly (3.6) is a more natural generalization than (3.8).

Moreover, when we construct a set of axioms for a function in general, we cannot use any knowledge of the concrete form of the function. If we set a special form such as '$+(1 - q)S_q(A)S_q(B)$' in (3.8) as one of the axioms for Tsallis entropy, we need a clear and meaningful reason why such an axiom can be given without any knowledge of the concrete form of Tsallis entropy. In fact, we cannot find any reason why $1 - q^2$ is not suitable instead of $1 - q$ in (3.8) without any knowledge of the concrete form (1.1). There are many other functions $\varphi(q)$ of q such that $\lim_{q \to 1} \varphi(q) = 0$. Therefore, the requirement of pseudoadditivity (2.2) as an axiom is not suitable and unnatural for an axiom of Tsallis entropy for natural generalization of the Shannon–Khinchin axioms.

We remark that in [10], Abe uses a generalized form of pseudoadditivity (2.2) such as

$$S_q(AB) = S_q(A) + S_q(B|A) + (1-q)S_q(A)S_q(B|A) \tag{3.9}$$

as one of the axioms. The requirement (3.9) is also inappropriate for an axiom of Tsallis entropy for the same reasons as stated above. The right-hand side of (3.9) is rather more complicated than that of (2.2).

4. Conclusion

We have presented the most concise set of axioms for Tsallis entropy. Our axioms consist of only two additivities: pseudoadditivity and Shannon additivity. We have rigorously proved that Tsallis entropy is uniquely determined by only these two additivities. Moreover, we have compared our axioms with Santos' axioms [7]. Our axioms consist of two out of Santos' four axioms. Thus Santos' axioms are redundant for unique determination of Tsallis entropy. So, our axioms in this paper are found to be the most concise set of axioms for Tsallis entropy. This is in agreement with the general requirement that a set of axioms for a function should be as concise as possible.

However, our axioms are *not* a natural generalization of the Shannon–Khinchin axioms [9] in extensive systems from the point of view of (1.8). This is due to the inappropriate choice of additivity for an axiom of Tsallis entropy. We then compared Santos' axioms with the Shannon–Khinchin axioms carefully and discussed in detail which additivity of pseudoadditivity and Shannon additivity is suitable for an axiom of Tsallis entropy. As a result, we have revealed that pseudoadditivity is unnatural and inappropriate for an axiom of Tsallis entropy in view of the correspondence (1.8). This is because there is no clear way of deriving it without any knowledge of the concrete form of Tsallis entropy. Therefore, Shannon additivity (2.5) and not pseudoadditivity (2.2) should be used as an axiom from the point of view of (1.8).

Generalization of the Shannon–Khinchin axioms to nonextensive systems is given by the author in [11]. The result in [11] coincides with another definition of Tsallis entropy by means of information content [12]. In the axioms in [11], only Shannon additivity is applied as additivity, which differs from the axioms presented above.

In summary, if pseudoadditivity (2.2) is accepted as an axiom of Tsallis entropy, the most concise set of axioms for Tsallis entropy can be constructed as shown in section 2. This result is in agreement with the general requirement that a set of axioms for a function should be as concise as possible. But these axioms are not a generalization of the Shannon–Khinchin axioms. In contrast, if pseudoadditivity (2.2) is not accepted as an axiom of Tsallis entropy, a natural generalization of the Shannon–Khinchin axioms can be constructed [11], but it is not as concise as the axioms shown in section 2.

References

[1] Tsallis C *et al* 2001 *Nonextensive Statistical Mechanics and Its Applications* ed S Abe and Y Okamoto (Heidelberg: Springer)

[2] Tsallis C, Rapisarda A, Latora V and Baldovin F 2002 Nonextensivity: from low-dimensional maps to Hamiltonian systems *Preprint* cond-mat/0209168

[3] Landsberg P T 1990 *Thermodynamics and Statistical Mechanics* (New York: Dover)

[4] Tsallis C 1988 *J. Stat. Phys.* **52** 479–87

[5] Curado E M F and Tsallis C 1991 *J. Phys. A: Math. Gen.* **24** L69–72
Curado E M F and Tsallis C 1991 *J. Phys. A: Math. Gen.* **24** 3187 (corrigendum)
Curado E M F and Tsallis C 1992 *J. Phys. A: Math. Gen.* **25** 1019 (corrigendum)

[6] Tsallis C, Mendes R S and Plastino A R 1998 *Physica* A **261** 534–54
[7] dos Santos R J V 1997 *J. Math. Phys.* **38** 4104–7
[8] Aczél J and Daróczy Z 1975 *On Measures of Information and Their Characterizations* (New York: Academic)
[9] Khinchin A I 1957 *Mathematical Foundations of Information Theory* (New York: Dover)
[10] Abe S 2000 *Phys. Lett.* A **271** 74–9
[11] Suyari H 2002 Generalization of Shannon–Khinchin axioms to nonextensive systems and the uniqueness theorem *Preprint* math-ph/0205004 submitted
[12] Suyari H 2002 *Phys. Rev.* E **65** 066118