Nonextensive entropies derived from form invariance of pseudoadditivity

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 August 2001; revised manuscript received 20 March 2002; published 21 June 2002)

The form invariance of pseudoadditivity is shown to determine the structure of nonextensive entropies. Nonextensive entropy is defined as the appropriate expectation value of nonextensive information content, similar to the definition of Shannon entropy. Information content in a nonextensive system is obtained uniquely from generalized axioms by replacing the usual additivity with pseudoadditivity. The satisfaction of the form invariance of the pseudoadditivity of nonextensive entropy and its information content is found to require the normalization of nonextensive entropies. The proposed principle requires the same normalization as that derived previously [A.K. Rajagopal and S. Abe, Phys. Rev. Lett. **83**, 1711 (1999)], but is simpler and establishes a basis for the systematic definition of various entropies in nonextensive systems.

DOI: 10.1103/PhysRevE.65.066118 PACS number(s): 02.50.−r, 05.20.−y, 89.70.+c

I. INTRODUCTION

Since the first proposal of nonextensive entropy by Tsallis [1,2], there have been many successful studies and applications analyzing physical systems such as long-range interactions, long-time memories, and multifractal structures in the nonextensively generalized Boltzmann-Gibbs statistical mechanics [3]. In the rapid progress in this field, some modifications have been made to mathematical formulations in the generalized statistical mechanics in order to maintain the physical consistency. One of the most important modifications was the introduction of an appropriate definition of the generalized expectation value. This modification has already appeared in the literature [4], but has been applied as a candidate for satisfying the physical requirements in a given situation without a systematic framework. The necessity to apply such modifications invites the establishment of guiding principles that will provide a clear basis for proposed generalizations of Boltzmann-Gibbs statistical mechanics. In a recent paper [5], Rajagopal and Abe presented a principle for determining the structure of nonextensive entropies. Their principle was the form invariance of the Kullback-Leibler entropy when generalized to nonextensive situations.

In the present paper, a much simpler principle for determining the structure of nonextensive entropies is presented. The original Tsallis entropy is determined for the given appropriate axioms [6–8]. In contrast to these axiomatic approaches, we define nonextensive entropy in another way; in terms of the appropriate expectation value of nonextensive information content similar to the definition for Shannon entropy [9]. This definition has already been applied to the generalization of the Shannon source coding theorem using the *normalized q-expectation value* of *nonextensive information content* [10]. Nonextensive information content $I_q(p)$ is defined by $I_q(p) = -\ln_q p$ in Ref. [10] as an intuitively natural generalization of the standard information content $I_1(p) = -\ln p$ (referred to as *self-information* or *self-entropy* in Shannon information theory [9]), where $\ln_q x$ is a q-logarithm function defined by $\ln_q x \equiv (x^{1-q}-1)/(1-q)$. However, in Ref. [10], the form invariance presented in this paper was not mentioned. We introduce the axioms for nonextensive information content $I_q(p)$ as a slight generalization of that for the standard information content, and obtain $I_q(p)$ uniquely from the generalized axioms. The requirement of form invariance of pseudoadditivity when we define nonextensive entropy $S_q(p)$ as the appropriate expectation value of $I_q(p)$,

$$S_q(p) \equiv E_{q,p}[I_q(p_i)], \tag{1}$$

leads to the determination of the structure of the nonextensive entropy, where $E_{q,p}[\cdot]$ is the expectation value satisfying the following form invariance of the pseudoadditivity:

$$\frac{I_q(p_1 p_2)}{k} = \frac{I_q(p_1)}{k} + \frac{I_q(p_2)}{k} + \varphi(q)\frac{I_q(p_1)}{k}\frac{I_q(p_2)}{k} \tag{2}$$

and

$$\frac{S_q(p^{AB})}{k} = \frac{S_q(p^A)}{k} + \frac{S_q(p^B)}{k} + \varphi(q)\frac{S_q(p^A)}{k}\frac{S_q(p^B)}{k}, \tag{3}$$

where k is a positive constant. $\varphi(q)$ is any function of the nonextensivity parameter q and satisfies the conditions (7) given below.

Note that information content means the amount of information provided by a result of an observation in a physical sense. The standard information content $I_1(p) = -\ln p$ has been called *surprise* by Watanabe [11], and *unexpectedness* by Barlow [12].

II. NONEXTENSIVE INFORMATION-CONTENT

The axioms of standard information content $I_1 : [0,1] \to R^+$ satisfying $I_1(1) = 0$ are given as follows [9]. [S1] I_1 is differentiable with respect to any $p \in (0,1)$, [S2] $I_1(p_1 p_2) = I_1(p_1) + I_1(p_2)$ for any $p_1, p_2 \in [0,1]$. Axiom [S2] means that the information content for two stochastically independent events is given by the sum of the two sets of information.

*Electronic address: suyari@tj.chiba-u.ac.jp

For the above axioms, $I_1(p)$ is determined uniquely by

$$I_1(p) = -k \ln p, \tag{4}$$

where k is a positive constant [9].

The above axioms are generalized in nonextensive situations as follows. Nonextensive information content $I_q: [0,1] \to R^+$ for any fixed $q \in R^+$, satisfying

$$\lim_{q\to 1} I_q(p) = I_1(p) = -k \ln p, \tag{5}$$

should have the following properties: [T1] I_q is differentiable with respect to any $p \in (0,1)$ and $q \in R^+$, [T2] $I_q(p)$ is convex with respect to $p \in [0,1]$ for any fixed $q \in R^+$, and [T3] there exists a function φ: $R \to R$ such that

$$\frac{I_q(p_1 p_2)}{k} = \frac{I_q(p_1)}{k} + \frac{I_q(p_2)}{k} + \varphi(q) \frac{I_q(p_1)}{k} \frac{I_q(p_2)}{k} \tag{6}$$

for any $p_1, p_2 \in [0,1]$, where $\varphi(q)$ is differentiable with respect to any $q \in R^+$,

$$\lim_{q\to 1} \frac{d\varphi(q)}{dq} \neq 0, \quad \lim_{q\to 1} \varphi(q) = \varphi(1) = 0, \quad \varphi(q) \neq 0\ (q \neq 1). \tag{7}$$

Equation (6) is called *pseudoadditivity* in many studies [3] as a special form of *composability* [13,14].

Note that in these generalized axioms, [T2] is needed to maintain non-negativity of the Kullback-Leibler entropy for any $q \in R^+$ when generalized to nonextensive situations [15,16]. In general, Kullback-Leibler entropy $K(p^A \| p^B)$ is defined by the appropriate expectation value of the difference between two information contents [9],

$$K(p^A \| p^B) \equiv E_{p^A}[I(p_i^B) - I(p_i^A)]. \tag{8}$$

Therefore, the non-negativity of the Kullback-Leibler entropy leads to Gibbs inequality [17,18],

$$K(p^A \| p^B) \geqslant 0 \;\Leftrightarrow\; S(p^A) = E_{p^A}[I(p_i^A)] \leqslant E_{p^A}[I(p_i^B)]. \tag{9}$$

When E_p is a normalized expectation value (i.e., $E_p[1] = 1$) and $p^B = (1/W, \ldots, 1/W)$, the right-hand side $E_{p^A}[I(p_i^B)]$ of the above Gibbs inequality is equal to the maximum entropy,

$$E_{p^A}\left[I\left(\frac{1}{W}\right)\right] = I\left(\frac{1}{W}\right) = E_{1/W}\left[I\left(\frac{1}{W}\right)\right] = S\left(\frac{1}{W}, \ldots, \frac{1}{W}\right). \tag{10}$$

This inequality coincides with the maximality condition that is one of the Shannon-Khinchin axioms [19,20], that is, the simplest form of the maximum entropy principle without constraints. Therefore, the satisfaction of the non-negativity of the Kullback-Leibler entropy for any $q \in R^+$ is needed when generalized to nonextensive systems. In order to satisfy the requirement of the non-negativity, the information content $I_q(p)$ should be convex with respect to $p \in [0,1]$ for any fixed $q \in R^+$ and its corresponding appropriate expectation value should be applied. Such examples of expectation value are *q-expectation value* and *normalized q*-expectation value [2,4]. Under these two conditions (convex information content and appropriate expectation value), the non-negativity of the nonextensive Kullback-Leibler can be proved. The proof is given in Appendix A.

Using the axioms [T1]–[T3], we determine $I_q(p)$ uniquely in the following procedures. Using Eq. (6), for any $1+\Delta \in (0,1)$, we have

$$\frac{I_q[p(1+\Delta)]}{k} = \frac{I_q(p+\Delta p)}{k} = \frac{I_q(p)}{k} + \frac{I_q(1+\Delta)}{k} + \varphi(q)\frac{I_q(p)}{k}\frac{I_q(1+\Delta)}{k}. \tag{11}$$

This can be rewritten as

$$\frac{I_q(p+\Delta p) - I_q(p)}{\Delta p} = \frac{1}{k}\frac{I_q(1+\Delta)}{\Delta}\frac{k+\varphi(q) I_q(p)}{p}. \tag{12}$$

Taking the limit $\Delta \to 0$ on both sides of Eq. (12), we obtain

$$\frac{dI_q(p)}{dp} = \frac{\beta}{k}\frac{k+\varphi(q)I_q(p)}{p}, \tag{13}$$

where $\beta \equiv \lim_{\Delta\to 0-}[I_q(1+\Delta)/\Delta]$ and the first axiom [T1] is applied. The differential equation is given by

$$\frac{1}{k+\varphi(q)y} dy = \frac{\beta}{k}\frac{1}{x} dx, \tag{14}$$

where $x \equiv p$ and $y \equiv I_q(p)$. This can be solved analytically; the rigorous solution is

$$y = k\frac{(Cx^\beta)^{\varphi(q)/k} - 1}{\varphi(q)}, \quad \text{that is,} \quad I_q(p) = k\frac{(Cp^\beta)^{\varphi(q)/k} - 1}{\varphi(q)}, \tag{15}$$

where C is a constant. Then, from the initial condition $\lim_{q\to 1} I_q(p) = I_1(p) = -k \ln p$, we have that $C = 1$ and $\beta = -k$, where conditions (7) are applied. Thus, the nonextensive information content $I_q(p)$ is derived as $I_q(p) = k[(p^{-\varphi(q)} - 1)/\varphi(q)]$.

Moreover, by [T2], the second differential of $I_q(p)$ with respect to p should be non-negative for any fixed $q \in R^+$. Thus, we can derive a constraint $\varphi(q) + 1 \geqslant 0$ for any $q \in R^+$.

Summarizing these results, the nonextensive information content $I_q(p)$ obtained from the axioms [T1]–[T3] is

$$I_q(p) = k\frac{p^{-\varphi(q)} - 1}{\varphi(q)}, \tag{16}$$

where k is a positive constant and

$$\varphi(q) + 1 \geqslant 0 \quad \text{for any} \quad q \in R^+. \tag{17}$$

For example, $\varphi(q)=q-1$ implies $I_q(p)=-k\ln_q p$.

Note that there have already been remarks on the alternative candidates for nonextensive information content to define the original Tsallis entropy [3]. They are

$$I_q^{(1)}(p)\equiv -k\ln_q p \quad\text{and}\quad I_q^{(2)}(p)\equiv k\ln_q p^{-1}, \tag{18}$$

and they only coincide for $q=1$.

$$S_q^{\mathrm{org}}(p)\equiv k\frac{1-\sum_{i=1}^{W} p_i^q}{q-1}=\sum_{i=1}^{W} p_i^q I_q^{(1)}(p_i)=\sum_{i=1}^{W} p_i I_q^{(2)}(p_i), \tag{19}$$

$I_q^{(1)}(p)$ and $I_q^{(2)}(p)$ correspond to $\varphi^{(1)}(q)\equiv q-1$ and $\varphi^{(2)}(q)\equiv 1-q$ as $\varphi(q)$ in Eq. (16), respectively. However, the latter case $\varphi^{(2)}(q)=1-q$ does *not* satisfy the identity (17) for any $q\in R^+$, that is, $I_q^{(2)}(p)$ does *not* possess the property of *convexity* [T2]. Therefore, $I_q^{(2)}(p)$ *cannot* be information content. Even if $I_q^{(2)}(p)$ is applied as information content, then the non-negativity property of the nonextensive Kullback-Leibler entropy is *not* held for any $q\in R^+$ for lack of the convexity of $I_q^{(2)}(p)$, as stated above. The convexity of information content is applied to Jensen's inequality in order to prove the non-negativity of the nonextensive Kullback-Leibler entropy. See Appendix A for the details.

In case of $\varphi(q)=q-1$ and $k=1$, the pseudoadditivity (6) of $I_q(p)$ is remarkably similar to the relation of the Jackson basic number in q-deformation theory [21,22] as follows. Let $[X]_q$ be the Jackson basic number of a physical quantity X, that is, $[X]_q\equiv(q^X-1)/(q-1)$. Then the Jackson basic number of $X+Y$ satisfies the identity $[X+Y]_q=[X]_q+[Y]_q+(q-1)[X]_q[Y]_q$. The surprising similarity to pseudoadditivity (6) can be seen if we consider a quantity $f(x)=p^{-(x-1)}$. Clearly, $f(1)=1$. Standard information content $I_1(p)$ is expressed as $I_1(p)=df(x)/dx|_{x=1}$; nonextensive information content is given by $I_q(p)=D_q f(x)|_{x=1}\equiv[f(qx)-f(x)]/(qx-x)|_{x=1}$, where D_q is the Jackson differential. According to q-deformation theory, the property $\lim_{q\to1}I_q(p)=I_1(p)$ originates from the convergence $\lim_{q\to1}D_q=d/dx$.

III. EFFECTS OF RENORMALIZATION OF NONEXTENSIVE ENTROPIES

The normalized nonextensive entropies follow naturally from the form invariance between entropy and its information content. In this section we assume $k=1$ for simplicity.

Similar to Shannon entropy, nonextensive entropy $S_q(p)$ is defined as the expectation value of the information content $I_q(p)$ obtained in Eq. (16). For example, the nonextensive entropy S_q^{org} using the unnormalized expectation value $E_{q,p}^{\mathrm{org}}[\cdot]$ is given by

$$S_q^{\mathrm{org}}(p)=E_{q,p}^{\mathrm{org}}[I_q(p)]\equiv\sum_{i=1}^{W} p_i^q I_q(p_i), \tag{20}$$

where W is the total number of microscopic configurations with probabilities $\{p_i\}$. In the definition of $E_{q,p}^{\mathrm{org}}$ in Eq. (20), the *q-expectation value* [2,4] is used. If we let $\varphi(q)=q-1$, then $S_q^{\mathrm{org}}(p)$ is concretely derived from Eqs. (16) and (20) as follows:

$$S_q^{\mathrm{org}}(p)=\frac{1-\sum_{i=1}^{W} p_i^q}{q-1}. \tag{21}$$

This is the original Tsallis entropy [1].

Let A and B be two independent systems in the sense that the probabilities p_{ij}^{AB} of the total system $A+B$ factorize into those of A and of B, i.e.,

$$p_{ij}^{AB}=p_i^A p_j^B \quad\text{for any}\quad i=1,\ldots,W_A \quad\text{and}\quad j=1,\ldots,W_B. \tag{22}$$

The nonextensive entropy $S_q^{\mathrm{org}}(p^{AB})$ of the total system $A+B$ can then be expanded using definitions (20) and (22) as follows:

$$\begin{aligned}S_q^{\mathrm{org}}(p^{AB})&=E_{q,p}^{\mathrm{org}}[I_q(p^{AB})]\\&=\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q I_q(p_{ij}^{AB})\\&=\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_i^A p_j^B)^q I_q(p_i^A p_j^B).\end{aligned} \tag{23}$$

Applying the pseudoadditivity (6) for information content $I_q(p)$, we obtain

$$\begin{aligned}S_q^{\mathrm{org}}(p^{AB})&=\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_i^A p_j^B)^q\{I_q(p_i^A)+I_q(p_j^B)\\&\quad+\varphi(q)I_q(p_i^A)I_q(p_j^B)\}\\&=\left(\sum_{j=1}^{W_B}(p_j^B)^q\right)S(p^A)+\left(\sum_{i=1}^{W_A}(p_i^A)^q\right)S(p^B)\\&\quad+\varphi(q)S(p^A)S(p^B),\end{aligned} \tag{24}$$

where we used

$$S^{\mathrm{org}}(p^A)=\sum_{i=1}^{W_A}(p_i^A)^q I_q(p_i^A)$$

and

$$S^{\mathrm{org}}(p^B)=\sum_{j=1}^{W_B}(p_j^B)^q I_q(p_j^B). \tag{25}$$

Dividing both sides of Eq. (24) by

$$\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q=\left(\sum_{i=1}^{W_A}(p_i^A)^q\right)\left(\sum_{j=1}^{W_B}(p_j^B)^q\right)(\neq 0), \tag{26}$$

yields

$$\frac{S_q^{\rm org}(p^{AB})}{\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q}=\frac{S_q^{\rm org}(p^A)}{\sum_{i=1}^{W_A}(p_i^A)^q}+\frac{S_q^{\rm org}(p^B)}{\sum_{j=1}^{W_B}(p_j^B)^q}+\varphi(q)\frac{S_q^{\rm org}(p^A)}{\sum_{i=1}^{W_A}(p_i^A)^q}\frac{S_q^{\rm org}(p^B)}{\sum_{j=1}^{W_B}(p_j^B)^q}. \quad (27)$$

In order to preserve the form of pseudoadditivity between nonextensive entropy $S_q(p)$ and the corresponding information content $I_q(p)$, the nonextensive entropy is modified as

$$S_q^{\rm nor}(p)\equiv\frac{S_q^{\rm org}(p)}{\sum_{j=1}^{W}p_j^q}=\frac{\sum_{i=1}^{W}p_i^qI_q(p_i)}{\sum_{j=1}^{W}p_j^q}, \quad (28)$$

which is the expectation value of the information content $I_q(p_i)$ with respect to the *escort distribution* [24] of p of order q. Then, by substituting (28) into (27), the following equation with respect to $S_q^{\rm nor}(p)$ is obtained:

$$S_q^{\rm nor}(p^{AB})=S_q^{\rm nor}(p^A)+S_q^{\rm nor}(p^B)+\varphi(q)S_q^{\rm nor}(p^A)S_q^{\rm nor}(p^B). \quad (29)$$

This is the same pseudoadditivity of nonextensive entropy $S_q(p)$ as Eq. (6). We have thus derived from the expectation value of $I_q(p)$ the form invariance of the pseudoadditivity of nonextensive entropy and its information content. Moreover, the nonextensive entropy $S_q^{\rm nor}(p)$ defined by Eq. (28) is actually the *q-normalized nonextensive entropy* [5,23]. Thus, according to the principle of the form invariance of pseudoadditivity between nonextensive entropy $S_q(p)$ and its information content $I_q(p)$, $S_q^{\rm nor}(p)$ should be used as nonextensive entropy instead of $S_q^{\rm org}(p)$. If $E_{q,p}^{\rm nor}$ denotes the normalized q-expectation value with respect to $\{p_i\}$, defined by

$$E_{q,p}^{\rm nor}[A]\equiv\frac{E_{q,p}^{\rm org}[A]}{\sum_{j=1}^{W}p_j^q}=\frac{\sum_{i=1}^{W}p_i^qA_i}{\sum_{j=1}^{W}p_j^q}, \quad (30)$$

where A is a physical quantity, then q-normalized nonextensive entropy $S_q^{\rm nor}(p)$ is given by

$$S_q^{\rm nor}(p)=E_{q,p}^{\rm nor}[I_q(p_i)]. \quad (31)$$

For the most typical case $\varphi(q)=q-1$, $S_q^{\rm nor}(p)$ is concretely given by

$$S_q^{\rm nor}(p)=\frac{1-\sum_{i=1}^{W}p_i^q}{(q-1)\sum_{j=1}^{W}p_j^q}. \quad (32)$$

This normalized Tsallis entropy (32) is concave only if the nonextensivity parameter q lies in (0,1) [3,5].

Note that definition (20) can be easily replaced with a more general unnormalized expectation value, leading to almost the same conclusion as that derived in this study. For example, if we use a more general form (16) as information content, then the expectation value $E_{q,p}^{\text{g-org}}$ defined by Eq. (A4) can be applied to the definition of the generalized original Tsallis entropy $S_q^{\text{g-org}}$,

$$\begin{aligned}S_q^{\text{g-org}}(p)&\equiv E_{q,p}^{\text{g-org}}[I_q(p_i)]\\&=\sum_{i=1}^{W}p_i^{\varphi(q)+1}I_q(p_i)\\&=\frac{1-\sum_{i=1}^{W}p_i^{\varphi(q)+1}}{\varphi(q)}.\end{aligned} \quad (33)$$

Then, along the same procedure as that presented in this section, the following $S_q^{\text{g-nor}}$ can be obtained in order to preserve the form invariance of pseudoadditivity between nonextensive entropy and its information content,

$$S_q^{\text{g-nor}}(p)\equiv\frac{S_q^{\text{g-org}}(p)}{\sum_{j=1}^{W}p_j^{\varphi(q)+1}}=\frac{1-\sum_{i=1}^{W}p_i^{\varphi(q)+1}}{\varphi(q)\sum_{j=1}^{W}p_j^{\varphi(q)+1}}. \quad (34)$$

In fact, when $\varphi(q)=q-1$, the formulas (33) and (34) coincide with Eqs. (21) and (32), respectively.

A dissatisfaction of the form invariance of the pseudoadditivity in the original Tsallis entropy can be revealed through the following simple calculation. Here we take $k=1$ for simplicity. When $\varphi(q)=q-1$, and substituting Eq. (16) into Eq. (20), $S_q^{\rm org}(p)$ coincides with the original Tsallis entropy [1] as shown in Eq. (21). The original Tsallis entropy $S_q^{\rm org}(p)$ given by Eq. (21) is widely known to satisfy the following pseudoadditivity [3]:

$$S_q^{\rm org}(p^{AB})=S_q^{\rm org}(p^A)+S_q^{\rm org}(p^B)+(1-q)S_q^{\rm org}(p^A)S_q^{\rm org}(p^B). \quad (35)$$

However, the pseudoadditivity (6) of $I_q(p)$ for the same condition [i.e., $\varphi(q)=q-1$] is given by

$$I_q(p_1p_2)=I_q(p_1)+I_q(p_2)+(q-1)I_q(p_1)I_q(p_2). \quad (36)$$

By comparing Eqs. (35) and (36), the coefficient $(q-1)$ of the cross term of pseudoadditivity in Eq. (36) *differs* from the $(1-q)$ in Eq. (35) when $E_{q,p}^{\rm org}[\cdot]$ defined by Eq. (20) is used. This clearly reveals that the form of the pseudoadditivity of $S_q^{\rm org}(p)$ and $I_q(p)$ is *not* invariant in the computation of $E_{q,p}^{\rm org}[\cdot]$. In other words, the form of the pseudoadditivity is *not* fixed when the unnormalized expectation value $E_{q,p}^{\rm org}[\cdot]$ is applied to the definition of Tsallis entropy.

More generally, for the generalized original Tsallis entropy $S_q^{\rm g\text{-}org}$ obtained in Eq. (33), the following pseudoadditivity is held:

$$S_q^{\rm g\text{-}org}(p^{AB}) = S_q^{\rm g\text{-}org}(p^A) + S_q^{\rm g\text{-}org}(p^B) - \varphi(q) S_q^{\rm g\text{-}org}(p^A) S_q^{\rm g\text{-}org}(p^B). \tag{37}$$

By comparing Eqs. (6) and (37), a "$\varphi(q)$ versus $-\varphi(q)$ inconsistency" can be found as similar as the above discussion. Note that when $q=1$, the form invariance discussed here holds because both $E_{q,p}^{\rm org}[\cdot]$ and $E_{q,p}^{\rm g\text{-}org}[\cdot]$ become normalized expectation values when $q=1$.

Therefore, the unnormalized expectation value such as $E_{q,p}^{\rm org}[\cdot]$ results in an inconsistency in the form invariance of the pseudoadditivity for the original Tsallis entropy.

If we let $\varphi(q)=q-1$, then from Eq. (29) the following pseudoadditivity holds:

$$S_q^{\rm nor}(p^{AB}) = S_q^{\rm nor}(p^A) + S_q^{\rm nor}(p^B) + (q-1) S_q^{\rm nor}(p^A) S_q^{\rm nor}(p^B). \tag{38}$$

In other words, $S_q^{\rm nor}(p)$ given by Eq. (32) satisfies the same pseudoadditivity (38) as Eq. (36). Therefore, the form invariance of pseudoadditivity requires the change from the familiar identity (35) to the modified one (38). This follows clearly from the above discussion because when $E_{q,p}^{\rm org}[\cdot]$, defined by Eq. (20), is applied to the definition of $S_q(p)$, the form invariance of the pseudoadditivity is *not* held as shown above.

Note that the obtained pseudoadditivity (38) is the same as the relation of the Jackson basic number: $[X+Y]_q = [X]_q + [Y]_q + (q-1)[X]_q[Y]_q$, where $[X]_q \equiv (q^X-1)/(q-1)$ [5]. Consider a quantity $\tilde{f}(x) \equiv 1/\Sigma_i (p_i)^x$. Clearly, $\tilde{f}(1)=1$.
Shannon entropy $S_1(p)$ is expressed as $S_1(p) = d\tilde{f}(x)/dx|_{x=1}$; normalized Tsallis entropy is given by $S_q^{\rm nor}(p) = D_q \tilde{f}(x)|_{x=1} \equiv [\tilde{f}(qx) - \tilde{f}(x)]/(qx-x)|_{x=1}$, where D_q is the Jackson differential.

IV. CONCLUSION

We have established a self-consistent principle for the form invariance of the pseudoadditivity of nonextensive entropy and its information content. The present principle is drawn from Shannon information theory and leads to the same normalization of the original Tsallis entropy as that derived in Ref. [5].

Once a set of an information content and an expectation value is given, various entropies such as Kullback-Leibler entropy (relative entropy) and mutual entropy can be formulated systematically. In nonextensive systems, an information content (16) and two expectation values (20) and (30) are given in the previous sections. Therefore, we can formulate two sets of various entropies based on two sets of the information content (16) and the expectation value (20) or (30), respectively. Please see the concrete formulas in Appendix B.

Note that the alternative selection from the original Tsallis entropy or the normalized Tsallis entropy should be careful in each application. From the mathematical point of view, the normalized Tsallis entropy has nice properties such as the form invariance of the pseudoadditivity, the unified application of the normalized q-expectation value, the form invariance of the statement of the maximum entropy principle, and so on. However, from the physical point of view, the original Tsallis entropy has many advantages over the normalized version. For example, the results derived from the Kolmogorov-Sinai entropy using the original Tsallis entropy have the perfect matching with nonlinear dynamical behavior such as the sensitivity to the initial conditions in chaos. On the other hand, the normalized Tsallis entropy does not have these convenient properties. Please see Refs. [25–28] for the details.

The principle discussed here is based on the usual formulation for information in Shannon information theory. Therefore, the ideas presented in this paper are an application of information theory to statistical mechanics, similar to the philosophy of Jaynes' work [29]. There remain many other applications of Shannon information theory to this interesting field.

ACKNOWLEDGMENTS

The author would like to thank Professor Yoshinori Uesaka and Professor Makoto Tsukada for their valuable comments and discussions. The author is grateful to Professor Sumiyoshi Abe for reading the first draft and his useful comments.

APPENDIX A: GIBBS INEQUALITY DERIVED FROM CONVEXITY OF INFORMATION CONTENT AND APPROPRIATE EXPECTATION VALUE

As presented in Eq. (8), the Kullback-Leibler entropy is generally defined by means of the information content,

$$K_q(p^A \| p^B) = E_{q,p^A}[I_q(p_i^B) - I_q(p_i^A)]. \tag{A1}$$

Our result (16) implies that

$$I_q(p_2) - I_q(p_1) = p_1^{-\varphi(q)} I_q\left(\frac{p_2}{p_1}\right). \tag{A2}$$

Substituting this relation into Eq. (A1), the Kullback-Leibler entropy is

$$K_q(p^A \| p^B) = E_{q,p^A}\left[(p_i^A)^{-\varphi(q)} I_q\left(\frac{p_i^B}{p_i^A}\right)\right]. \tag{A3}$$

If we take an unnormalized expectation value,

$$E_{q,p}^{\rm g\text{-}org}[X] \equiv \sum_{i=1}^{W} p_i^{\varphi(q)+1} X_i, \tag{A4}$$

then

$$K_q(p^A \| p^B) = \sum_{i=1}^{W} p_i^A I_q\left(\frac{p_i^B}{p_i^A}\right). \tag{A5}$$

$I_q(p)$ is a convex function with respect to $p \in [0,1]$ for any $q \in R^+$. Thus, we can use Jensen's inequality [9]: If f is a convex function and X is a physical quantity (random variable in mathematics), then

$$E[f(X)] \geqslant f(E[X]), \tag{A6}$$

where E is the usual expectation value when $q=1$. Therefore, Eq. (A5) satisfies

$$\begin{aligned} K_q(p^A \| p^B) &= \sum_{i=1}^{n} p_i^A I_q\left(\frac{p_i^B}{p_i^A}\right) \\ &\geqslant I_q\left(\sum_{i=1}^{W} p_i^A \frac{p_i^B}{p_i^A}\right) \\ &= I_q\left(\sum_{i=1}^{W} p_i^B\right) \\ &= I_q(1) \\ &= 0. \end{aligned} \tag{A7}$$

If we take a normalized expectation value:

$$E_{q,p}^{\text{g-nor}}[X] \equiv \sum_{i=1}^{W} \frac{p_i^{\varphi(q)+1}}{\sum_{j=1}^{n} p_j^{\varphi(q)+1}} X_i, \tag{A8}$$

then the same Gibbs inequality as Eq. (A7) can be obtained. When $\varphi(q)=q-1$, the expectation values (A4) and (A8) coincide with *q-expectation value* defined by Eq. (20) and *normalized q-expectation value* defined by Eq. (30), respectively. Thus, if an expectation value is chosen appropriately for a given information content, then the non-negativity condition of the Kullback-Leibler entropy is held.

APPENDIX B: SYSTEMATIC FORMULATIONS OF VARIOUS NONEXTENSIVE ENTROPIES

The proposed procedure for defining nonextensive entropy using an information content and an expectation value is applicable to the systematic formulations of various nonextensive entropies such as Kullback-Leibler entropy (relative entropy) and mutual entropy in accordance with the formulations in Shannon information theory [9].

In nonextensive systems, an information content and two expectations are given in Sec. III. Therefore, we can formulate various nonextensive entropies in the following two cases: case 1, the information content (16) and the q-expectation value (20); case 2, the information content (16) and the normalized q-expectation value (30). In each formulation $\varphi(q)=q-1$ is used as the most typical function of $\varphi(q)$.

1. Various nonextensive entropies using q-expectation value $E_q^{\text{org}}[\cdot]$

The information content (16) and the q-expectation value (20) are applied to the definitions of various entropies as follows:

(O-1) (nonextensive joint entropy),

$$S_q^{\text{org}}(p^{AB}) \equiv E_{q,p^{AB}}^{\text{org}}[I_q(p_{ij}^{AB})] = \frac{1-\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q}{q-1}; \tag{B1}$$

(O-2) (nonextensive conditional entropy),

$$\begin{aligned} S_q^{\text{org}}(p^{B|A}) &\equiv E_{q,p^A}^{\text{org}}\left[S_q\left(\frac{p_{i*}^{AB}}{p_i^A}\right)\right] \\ &= \sum_{i=1}^{W_A}(p_i^A)^q\left[\frac{1-\sum_{j=1}^{W_B}\left(\frac{p_{ij}^{AB}}{p_i^A}\right)^q}{q-1}\right] \\ &= \frac{\sum_{i=1}^{W_A}(p_i^A)^q-\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q}{q-1}, \end{aligned} \tag{B2}$$

where

$$\left\{\frac{p_{i*}^{AB}}{p_i^A}\right\} = \left\{\frac{p_{i1}^{AB}}{p_i^A}, \ldots, \frac{p_{iW_B}^{AB}}{p_i^A}\right\} \quad \text{for} \quad i=1,\ldots,W_A;$$

(O-3) (nonextensive Kullback-Leibler entropy),

$$K_q^{\text{org}}(p^A \| p^B) \equiv E_{q,p^A}^{\text{org}}[I_q(p_i^B)-I_q(p_i^A)] = \frac{1-\sum_{i=1}^{W} p_i^B\left(\frac{p_i^A}{p_i^B}\right)^q}{1-q}; \tag{B3}$$

(O-4) (nonextensive mutual entropy),

$$\begin{aligned} \mathcal{I}_q^{\text{org}}(p^A;p^B) &\equiv K_q^{\text{org}}(p^{AB} \| p^A p^B) \\ &= E_{q,p^{AB}}^{\text{org}}[I_q(p_i^A p_j^B)-I_q(p_{ij}^{AB})] \\ &= \frac{1-\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_i^A p_j^B)\left(\frac{p_{ij}^{AB}}{p_i^A p_j^B}\right)^q}{1-q}. \end{aligned} \tag{B4}$$

2. Various nonextensive entropies using normalized q-expectation value $E_q^{\text{nor}}[\cdot]$

The information content (16) and the q-expectation value (30) are applied to the definitions of various entropies as follows:

(N-1) (nonextensive joint entropy),

$$S_q^{\mathrm{nor}}(p^{AB}) \equiv E_{q,p^{AB}}^{\mathrm{nor}}[I_q(p_{ij}^{AB})] = \frac{1-\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q}{(q-1)\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_{ij}^{AB})^q}; \tag{B5}$$

(N-2) (nonextensive conditional entropy) ,

$$S_q^{\mathrm{nor}}(p^{Bt|A}) \equiv E_{q,p^A}^{\mathrm{nor}}\left[S_q\left(\frac{p_{i*}^{AB}}{p_i^A}\right)\right] = \frac{\sum_{i=1}^{W_A}\left[(p_i^A)^q \frac{1-\sum_{s=1}^{W_B}\left(\frac{p_{is}^{AB}}{p_i^A}\right)^q}{(q-1)\sum_{t=1}^{W_B}\left(\frac{p_{it}^{AB}}{p_i^A}\right)^q}\right]}{\sum_{j=1}^{W_A}(p_j^A)^q}, \tag{B6}$$

where

$$\left\{\frac{p_{i*}^{AB}}{p_i^A}\right\} = \left\{\frac{p_{i1}^{AB}}{p_i^A}, \ldots, \frac{p_{iW_B}^{AB}}{p_i^A}\right\} \quad \text{for} \quad i=1,\ldots,W_A;$$

(N-3) (nonextensive Kullback-Leibler entropy),

$$K_q^{\mathrm{nor}}(p^A \| p^B) \equiv E_{q,p^A}^{\mathrm{nor}}[I_q(p_i^B) - I_q(p_i^A)] = \frac{1-\sum_{i=1}^{W} p_i^B\left(\frac{p_i^A}{p_i^B}\right)^q}{(1-q)\sum_{j=1}^{W}(p_j^A)^q}; \tag{B7}$$

(N-4) (nonextensive mutual entropy),

$$\begin{aligned}\mathcal{I}_q^{\mathrm{nor}}(p^A;p^B) &\equiv K_q^{\mathrm{nor}}(p^{AB}\|p^A p^B) \\ &= E_{q,p^{AB}}^{\mathrm{nor}}[I_q(p_i^A p_j^B) - I_q(p_{ij}^{AB})] \\ &= \frac{1-\sum_{i=1}^{W_A}\sum_{j=1}^{W_B}(p_i^A p_j^B)\left(\frac{p_{ij}^{AB}}{p_i^A p_j^B}\right)^q}{(1-q)\sum_{s=1}^{W_A}\sum_{t=1}^{W_B}(p_{st}^{AB})^q}.\end{aligned} \tag{B8}$$

Note that both of the nonextensive mutual entropies, respectively, defined by Eqs. (B4) and (B8) are clearly *symmetric* with respect to "$\{p_i^A\} \leftrightarrow \{p_j^B\}$" in each formulation. Furthermore, the *non-negativity* of mutual entropies (B4) and (B8) is directly derived from that of the nonextensive Kullback-Leibler entropy, i.e.,

$$\mathcal{I}_q(p^A;p^B) = K_q(p^{AB}\|p^A p^B) \geqslant 0 \quad \text{for any} \quad q \in R^+, \tag{B9}$$

with equality $p_{ij}^{AB} = p_i^A p_j^B$ for any $i=1,\ldots,W_A$ and $j=1,\ldots,W_B$.

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

[2] E.M.F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991); **24**, 3187 (1991); **25**, 1019 (1992).

[3] C. Tsallis, *et al.*, *Nonextensive Statistical Mechanics and Its Applications*, edited by S. Abe and Y. Okamoto (Springer, Heidelberg, 2001); see also the comprehensive list of references in http://tsallis.cat.cbpf.br/biblio.htm

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

[5] A.K. Rajagopal and S. Abe, Phys. Rev. Lett. **83**, 1711 (1999).

[6] R.J.V. dos Santos, J. Math. Phys. **38**, 4104 (1997).

[7] K.S. Fa, J. Phys. A **31**, 8159 (1998).

[8] S. Abe, Phys. Lett. A **271**, 74 (2000).

[9] A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding* (McGraw-Hill, New York, 1979); T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991); J. G. Proakis and M. Salehi, *Communication Systems Engineering* (Prentice-Hall, Englewood Cliffs, NJ, 1994).

[10] T. Yamano, Phys. Rev. E **63**, 046105 (2001).

[11] S. Watanabe, *Knowing and Guessing* (Wiley, New York, 1969).

[12] H. Balow, Vision Res. **30**, 1561 (1990).

[13] M. Hotta and I. Joichi, Phys. Lett. A **262**, 302 (1999).

[14] S. Abe, Phys. Rev. E **63**, 061105 (2001).

[15] C. Tsallis, Phys. Rev. E **58**, 1442 (1998).

[16] L. Borland, A.R. Plastino and C. Tsallis, J. Math. Phys. **39**, 6490 (1998); **40**, 2196(E) (1999).

[17] J.W. Gibbs, *Collected Works* (Yale University Press, New Haven, CT, 1948).

[18] A. Isihara, *Statistical Physics* (Academic Press, New York, 1971).

[19] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1963).

[20] A.I. Khinchin, *Mathematical Foundations of Information Theory* (Dover, New York, 1957).

[21] S. Abe, Phys. Lett. A **224**, 326 (1997); **244**, 229 (1998).

[22] R.S. Johal, Phys. Rev. E **58**, 4147 (1998).

[23] P.T. Landsberg and V. Vedral, Phys. Lett. A **247**, 211 (1998).

[24] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems* (Cambridge University Press, Cambridge, 1993).

[25] C. Tsallis, A.R. Plastino, and W.-M. Zheng, Chaos, Solitons Fractals **8**, 885 (1997).

[26] U.M.S. Costa, M.L. Lyra, A.R. Plastino, and C. Tsallis, Phys. Rev. E **56**, 245 (1997).

[27] M.L. Lyra and C. Tsallis, Phys. Rev. Lett. **80**, 53 (1998).

[28] E.P. Borges, C. Tsallis, G.F.J. Ananos, and P.M.C. Oliveira, e-print cond-mat/0203342.

[29] E.T. Jaynes, *Papers on Probability, Statistics and Statistical Physics*, edited by R.D. Rosenkrantz (D. Reidel, Boston, 1983).