Derivation of the entropic formula for the statistical mechanics of space plasmas

George Livadiotis
Southwest Research Institute, Space Science and Engineering, San Antonio, TX, USA

Correspondence: George Livadiotis (glivadiotis@swri.edu)

Received: 19 September 2017 – Discussion started: 25 September 2017
Revised: 16 December 2017 – Accepted: 1 January 2018 – Published: 7 February 2018

Abstract. Kappa distributions describe velocities and energies of plasma populations in space plasmas. The statistical origin of these distributions is associated with the framework of nonextensive statistical mechanics. Indeed, the kappa distribution is derived by maximizing the $q$ entropy of Tsallis, under the constraints of the canonical ensemble. However, the question remains as to what the physical origin of this entropic formulation is. This paper shows that the $q$ entropy can be derived by adapting the additivity of energy and entropy.

1 Introduction

Space plasmas are collisionless and correlated particle systems characterized by a non-Maxwellian behavior, typically described by the formulations of kappa distributions. The origin of this vastly different statistical behavior between classical systems and space plasmas is the manifestation of correlations between the plasma particles. These systems are characterized by long-range interactions that induce correlations, resulting in a collective behavior among particles (e.g., see Jund et al., 1995; Salazar and Toral, 1999; Villain, 2008; Tsallis, 2009; Grassi, 2010; Tiranaki and Borges, 2016). The induction of any type of correlations among particles (more accurately, among particle energies or particle phase space) departs the system from thermal equilibrium to be restabilized into other stationary states out of thermal equilibrium described by kappa distributions, or combinations/superposition thereof. (Note that as we will see in Sect. 3, single kappa distributions induce a certain type of correlation, which, however, can be further generalized when a combination or superposition of kappa distributions is taken into account; e.g., see Spectral Statistics, Tsallis, 2009, Linear/Nonlinear superposition, chap. 6.2.1; Livadiotis and McComas, 2013a, Appendix A; Livadiotis, 2017a, chap. 4.3.4.)

Kappa distributions describe numerous space plasma populations. Several examples are the following: (i) the inner heliosphere, including solar wind (e.g., Maksimovic et al., 1997; Pierrard et al., 1999; Mann et al., 2002; Marsch, 2006; Zouganelis, 2008; Štverák et al., 2009; Livadiotis and McComas, 2013a; Yoon, 2014; Pierrard and Pieters, 2015; Pavlos et al., 2016), solar spectra (e.g., Dzifčaková and Dudík, 2013; Dzifčaková et al., 2015), solar corona (e.g., Owocki and Scudder, 1983; Vocks et al., 2008; Lee et al., 2013; Cranmer, 2013), solar energetic particles (e.g., Xiao et al., 2008; Laming et al., 2013), corotating interaction regions (e.g., Chotoo et al., 2000), and related solar flares (e.g., Mann et al., 2009; Livadiotis and McComas, 2013b; Bian et al., 2014; Jeffrey et al., 2016); (ii) planetary magnetospheres, including magnetosheath (e.g., Formisano et al., 1973; Ogasawara et al., 2013), magnetopause (e.g., Ogawara et al., 2013), solar corona (e.g., Owocki and Scudder, 1983; Vocks et al., 2008; Lee et al., 2013; Cranmer, 2014), solar energetic particles (e.g., Xiao et al., 2008; Laming et al., 2013), corotating interaction regions (e.g., Chotoo et al., 2000), and related solar flares (e.g., Mann et al., 2009; Livadiotis and McComas, 2013b; Bian et al., 2014; Jeffrey et al., 2016); (iii) the outer heliosphere and the inner heliosheath (e.g., Decker and Krimigis, 2003;
Decker et al., 2005; Heerikhuisen et al., 2008, 2015; Zank et al., 2010; Livadiotis et al., 2011, 2012, 2013; Livadiotis and McComas, 2011a, 2012, 2013; Livadiotis and McComas, 2013b, 2013c; Livadiotis, 2014, 2016; Fuselier et al., 2014; Zirnstein and McComas, 2015; Zank, 2015); (iv) beyond the heliosphere, including HII regions (e.g., Nicholls et al., 2012), planetary nebula (e.g., Nicholls et al., 2013; Zhang et al., 2014), and supergiant magnetospheres (e.g., Raymond et al., 2010), and on cosmological scales (e.g., Hou et al., 2017); and (v) other space-plasma related analyses (e.g., Milovanov and Zelenyi, 2000; Saito et al., 2000; Du, 2004; Yoon et al., 2006, 2012; Raadu and Shafiq, 2007; Livadiotis, 2009, 2015a, b, c, 2018; Tribeche et al., 2009; Hellberg et al., 2009; Livadiotis and McComas, 2010b, 2014; Baluku et al., 2010; Le Roux et al., 2010; Elsami et al., 2011; Kourakis et al., 2012; Randol and Christian, 2014, 2016; Varotsos et al., 2014; Fisk and Gloeckler, 2014; Viñas et al., 2014, 2015; Ourabah et al., 2015; Dos Santos et al., 2016; Nicolaou and Livadiotis, 2016). (See also the book Livadiotis, 2017a, and references therein.) Finally, it has to be noted that the kappa distributions and its associated statistical mechanics have been applied in a variety of disciplines other than space and plasma physics. A few examples are the following: in sociometry, e.g., the internet (Abe and Suzuki, 2003) and in citation networks of scientific papers (Tsallis and De Albuquerque, 2000) and urban agglomeration (Malacarne et al., 2001); in linguistics (Montemurro, 2001); in economics (Borland, 2002); in biochemistry (Andricioaei and Straub, 1996); in applied statistics (Habeck et al., 2005); in nonlinear dynamics (Borges et al., 2002); in physical chemistry (Livadiotis, 2009); and in ecology (Livadiotis et al., 2015, 2016).

Empirical kappa distributions were introduced in mid-1960s by Binsack (1966), Olbert (1968), and Vasyliunas (1968), while their connection with statistical mechanics was shown and studied in detail about half a century later (see Livadiotis and McComas, 2009, and references therein). In particular, the statistical origin of these distributions is now widely accepted to be determined within the framework of nonextensive statistical mechanics (Tsallis, 2009). This is a consistent generalization of the classical statistical mechanics, which is based on a monoparametric (q index) entropic formula (Tsallis, 1988). The theoretical q exponential distribution, which results from the maximization of entropy in the canonical ensemble, has the same formulation as the empirical kappa distribution; the two distributions are identical under the transformation of their characteristic indices (q = 1 + 1/κ).

Though a consistent connection of the mathematical model of kappa distributions has been attained with the physical means of entropy maximization, this does not precisely answer the main question regarding the origin of these distributions. We have only shifted the modeling from the distributions to the entropic formulation. Therefore, we may understand now that the statistical origin of kappa distributions is given by the Tsallis entropy maximization in the canonical ensemble, but still, the origin of this specific entropic formulation remains unknown.

Certainly, there are various mechanisms responsible for generating kappa distributions in space and other plasmas, for example, the presence of pickup ions (Livadiotis and McComas, 2010a, 2011a) or weak turbulence (Yoon et al., 2012; Yoon, 2014). Moreover, kappa distributions belong to the framework of nonextensive statistical mechanics. Thus, once a kappa distribution is generated and stabilized into a plasma population, the whole “tool package” of nonextensive statistical mechanics is applicable for describing the statistical physics of this population; for instance, the entropy is given by the Tsallis formulation, while the temperature can be determined by the mean kinetic energy.

Here, we do not argue which mechanisms generate kappa distributions in space plasmas, rather the physical reasons that these distributions sustain themselves in space plasmas once generated. The typical answer is that this is an effect of the presence and preservation of correlations in the collisionless environment that governs space plasmas. (For particle systems such as space plasmas, collisions can destroy correlations, and thus, their collective behavior.) The collisionless environment conserves the energy. Moreover, weakly coupled plasmas (mutual electron and ion potential energy is small compared to the average kinetic energy) can be described as ideal gases. Interparticle energy terms can be ignored, leading to the additivity of energy: the energy of a multi-particle state is the sum of the energies of all the one-particle states involved. On the other hand, the preservation of local correlations among particles creates a conceptual separation of particles in correlation clusters. Debye spheres are correlation clusters that may include up to trillions of particles since space plasmas are weakly coupled (Bryant, 1996; Rubab and Murata, 2006; Gougam and Tribeche, 2011; Livadiotis and McComas, 2014). This structure can lead to the additivity of entropy: the entropy of a multi-particle state is the sum of the entropies of all the one-particle states involved. (Note that ideal gases are considered to have (i) zero interparticle interactions, and (ii) zero particle correlations. While ideal gases are characterized by short-range interactions that cannot induce correlations among particles, space plasmas have interactions weak enough to be negligible but with a long enough range that correlations can be induced.)

The purpose of this paper is to show that there is a deeper connection of Tsallis q entropy and space plasmas: namely, we will show that two simple first principles, such as the additive energy and additive entropy, which apply to plasma particle populations, are sufficient for indicating the specific formula of q entropy (Fig. 1). Therefore, the main objective of this work is to demonstrate the theory which determines that the entropic form given by the q-entropy formula proposed originally by Tsallis (1988) follows from certain assumptions regarding the (microscopic) state of the system. The importance of this discussion for the (space) plasma physics community resides mostly in the fact that the kappa
independent discrete distributions, i.e., labeling the two particles with no correlations among particle velocities or energies. Therefore, the joint two-particle probability distribution can be additive if the energy is also additive (red arrow 2). In the same way, it can be shown that these entropic formulations can lead to additive energy if the entropy is additive (purple arrow 1). The objective here is to show that the entropic formula can be derived from the additivity of energy and entropy (blue arrow 3).

velocity/energy distribution functions, ubiquitously observed in space and astrophysical environments, can be derived from the maximization of the $q$ entropy, under the constraints of a canonical ensemble.

In Sect. 2, we describe the physical motive of this paper in detail. In Sect. 3, we show in detail a similar property for both the entropic formalisms of Boltzmann–Gibbs (BG) and Tsallis: the entropy is non-additive in general for some arbitrary probability distribution, but it can become additive specifically for the canonical probability distribution (the one that maximizes the corresponding entropy). In Sect. 4, we show how we can determine the entropic formula appropriate for describing the plasma particle populations, simply by setting two first principles’ properties, obvious for collisionless plasmas: energy and entropy are additive, at least macroscopically. Finally, Sect. 5 briefly summarizes the conclusions.

2 Physical motive

Classical BG statistical mechanics characterizes systems with no correlations among particle velocities or energies. Therefore, the joint two-particle probability distribution can be expressed as the product of the one-particle identical and independent discrete distributions, i.e., labeling the two particles with A and B, $p_{ij}^{A+B} = p_i^A \cdot p_j^B$. Hereafter, we consider a particle system described by a discrete energy spectrum $\{e_k\}_{k=1}^W$, which is associated with a discrete probability distribution $\{p_k\}_{k=1}^W$. The same semantics is used when the system is separated in two subsystems, A and B, where the two-particle distribution describes a two-particle state, with one particle at each subsystem. The logarithm of the probability is an additive function, $\ln p_{ij}^{A+B} = \ln p_i^A + \ln p_j^B$, from which we obtain the additivity of entropy (here the BG entropy), $S^{A+B} = S^A + S^B$. For special cases, however, where the independence relationship does not apply, $p_{ij}^{A+B} \neq p_i^A \cdot p_j^B$, the entropy is non-additive, $S^{A+B} \neq S^A + S^B$. The logical rec-
(i.e., $q$ exponential or kappa distribution) conforms to the specific correlations expressed by $g(x) \propto (x^q-1)/(q-1)$ or $(p_i^{A+B})^{q-1}= (p_i^A)^{q-1} + (p_i^B)^{q-1} - 1$, which again makes the entropy additive. In Sect. 3 we show this similar property of the two statistical formalisms in detail.

Then, we may ask the following question: is the property of BG and Tsallis entropies described above a general feature of any physically meaningful entropic function? Or can we reverse the question, and ask which specific entropic function follows the above properties? It will be really intriguing if we can determine the entropic formula appropriate for describing the plasma particle populations, simply by setting the following two first principles' properties: (1) additive energy, (2) additive entropy; i.e., the probability distribution derived by maximizing the entropy under the constraints of the canonical ensemble makes the entropy additive. This will be the main purpose of this paper and will be examined in Sect. 4.

3 Canonical ensemble distributions with additive energy lead to additive entropy

3.1 The Gibbs’ path

The Gibbs’ path (1902) for the maximization of the entropy $S(p_1, p_2, \ldots, p_W)$ under the constraints of a canonical ensemble (i.e., (i) normalization $1 = \sum_{k=1}^{W} p_k$ and (ii) fixed internal energy $U = \sum_{k=1}^{W} p_k e_k$) involves maximizing the functional

$$G(p_1, p_2, \ldots, p_W) = S(p_1, p_2, \ldots, p_W) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k e_k.$$  

(1)

Next, we examine the BG and Tsallis entropic formulations.

3.2 BG entropy

First, we start from the classical case of BG entropy

$$S(p_1, p_2, \ldots, p_W) = - \sum_{k=1}^{W} p_k \ln(p_k),$$  

(2)

where we ignored the Boltzmann constant $k_B$ for simplicity. Then, setting $(\partial/\partial p_j)G(p_1, p_2, \ldots, p_W) = 0$ to

$$G(p_1, p_2, \ldots, p_W) = - \sum_{k=1}^{W} p_k \ln(p_k) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k e_k,$$  

(3)

we find

$$p_j(\varepsilon_j) = \exp(\lambda_1 - 1) \cdot \exp(\lambda_2 \varepsilon_j).$$  

(4)

We may write Eq. (4) in a logarithmic form, $\ln p_j = \lambda_2 \varepsilon_j + \lambda_1 - 1$. Then, we separate the particle system into two parts, $A$ and $B$, so that each part is a new subsystem for which Eq. (4) holds:

$$\ln p_j^A = \lambda_2 \varepsilon_j^A + \lambda_1 - 1 \quad \text{and} \quad \ln p_j^B = \lambda_2 \varepsilon_j^B + \lambda_1 - 1.$$  

(5)

The whole system is characterized by the joint probability, $p_{ij}^{A+B}$, meaning the probability of a particle in the subsystem $A$ of residing in the state $i$ and a particle in the subsystem $B$ of residing in the state $j$. This is related with the energy $\varepsilon_{ij}^{A+B}$ of the two-particle state,

$$\ln p_{ij}^{A+B} = \lambda_2 \varepsilon_{ij} + \lambda_1 - 1.$$  

(6)

Trivially, the energy of the two-particle state energy $\varepsilon_{ij}^{A+B}$ equals the summation of the energy of each particle (since no interparticle force is considered); i.e., the system’s energy is additive:

$$\varepsilon_{ij}^{A+B} = \varepsilon_i^A + \varepsilon_j^B.$$  

(7)

Hence, by eliminating energies from Eqs. (5) and (6), we find

$$\ln p_{ij}^{A+B} + (\lambda_1 - 1) = \lambda_2 \varepsilon_{ij}^A + (\lambda_1 - 1) + \lambda_2 \varepsilon_{ij}^B + (\lambda_1 - 1)$$

$$= \ln p_{ij}^A + \ln p_{ij}^B \quad \text{or}$$

$$p_{ij}^{A+B} = p_i^A \cdot p_j^B \cdot e^{-(\lambda_1-1)}.$$  

(8)

(9)

At this point we recall that the Lagrange multipliers, $\lambda_1$ and $\lambda_2$, are related with the partition function $Z = e^{-(\lambda_1-1)}$, and the inverse temperature $\beta = -\lambda_2$, respectively, and they are not necessarily equal for the two subsystems $A$ and $B$, or the whole system $A + B$. Nevertheless, the logarithm of the partition function or $(\lambda_1 - 1)$ is an extensive parameter, i.e., $(\lambda_1 - 1)^{A+B} = (\lambda_1 - 1)^A + (\lambda_1 - 1)^B$, while the temperature is not an extensive parameter and can be considered the same: $\lambda_2^{A+B} = \lambda_2^A = \lambda_2^B$. Then, instead of Eqs. (8) and (9), we obtain

$$\ln p_{ij}^{A+B} = \lambda_2 \varepsilon_{ij}^{A+B} + (\lambda_1 - 1)^{A+B}$$

$$= \lambda_2 \varepsilon_i^A + (\lambda_1 - 1)^A + \lambda_2 \varepsilon_j^B + (\lambda_1 - 1)^B$$

$$= \ln p_{ij}^A + \ln p_{ij}^B.$$  

(10)

which clearly shows that the canonical probabilities are independent:

$$\ln \left( p_{ij}^{A+B} \right) = \ln \left( p_i^A \right) + \ln \left( p_j^B \right) \Rightarrow p_{ij}^{A+B} = p_i^A \cdot p_j^B.$$  

(11)

Equation (9) indicates that the result in Eq. (11) can be obtained simply by setting $\lambda_1 = 1$. Certainly, this restricts the generality, but it can be used as a trick to simplify the calculations. Furthermore, we can easily obtain the additivity of entropy. Indeed, applying the operator $\sum_{i=1}^{W} \sum_{j=1}^{W} p_{ij}^{A+B} \times$ to
both sides of Eq. (11), we obtain
\[ p_{ij}^{A+B} \ln p_{ij}^{A+B} = p_{ij}^A \ln p_{ij}^A + p_{ij}^{A+B} \ln p_{ij}^B \]
\[ \Rightarrow - \sum_{i=1}^{W} \sum_{j=1}^{W} p_{ij}^{A+B} \ln p_{ij}^{A+B} = - \sum_{i=1}^{W} p_{ij}^A \ln (p_{ij}^A) \]
\[ - \sum_{i=1}^{W} p_{ij}^B \ln (p_{ij}^B) \]
(12)

because \( \sum_{j=1}^{W} p_{ij}^{A+B} = p_{ij}^A \) and \( \sum_{i=1}^{W} p_{ij}^{A+B} = p_{ij}^B \). Hence, we arrive at the additivity of the entropy of the system to the entropies of the subsystems,
\[ S^{A+B} = S^A + S^B. \]
(13)

### 3.3 Tsallis entropy

Next, we continue with the Tsallis q entropy.
\[ S(p_1, p_2, \ldots, p_W) = \frac{1 - \varphi(p_1, p_2, \ldots, p_W)}{q - 1} \]
\[ = \frac{1}{q - 1} \cdot \sum_{k=1}^{W} (p_k - p_k^q). \]
(14)

(e.g., Havrda and Charvát, 1967; Tsallis, 1988) where the argument \( \varphi \) is defined by
\[ \varphi(p_1, p_2, \ldots, p_W) = \sum_{k=1}^{W} p_k^q. \]
(15)

Again, the maximization of the entropy under the constraints of the canonical ensemble involves maximizing the functional
\[ G(p_1, p_2, \ldots, p_W) = \frac{1}{q - 1} \cdot \sum_{k=1}^{W} (p_k - p_k^q) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k \]
\[ \lambda_2 \sum_{k=1}^{W} p_k \]
(16)

Note that for simplicity we do not use the formulation of escort distributions (Beck and Schlogl, 1993). The dyadic formalism of ordinary/escort distributions is of fundamental importance in modern nonextensive statistical mechanics (Livadiotis, 2017a; chap. 1). It has been shown that this dyadic formalism of distributions can be avoided in order to simplify the theory, but it leads to a dyadic formulation of entropy (Livadiotis, 2017b).

Hence, \( (\partial/\partial p_j) G(p_1, p_2, \ldots, p_W) = 0 \) gives
\[ p_j(\varepsilon_j) = \left[ 1 + (1 - q^{-1}) \cdot (\lambda_1 - 1) \right]^{-q^{-1} - 1} \]
\[ \times \left[ 1 + (1 - q^{-1}) \cdot \frac{\lambda_2 \varepsilon_j}{1 + (1 - q^{-1}) \cdot (\lambda_1 - 1)} \right]^{q^{-1} - 1} \]
(17a)
or
\[ p_j(\varepsilon_j) = \exp_{q^{-1}}^{-(\lambda_1 - 1)} \cdot \exp_{q^{-1}}^{-(\lambda_2 \varepsilon_j / (1 - q^{-1} \cdot (\lambda_1 - 1)))}, \]
(17b)

which reflects a generalization of Eq. (4). We used the \( \lambda \)-deformed exponential function, and its inverse, the \( \lambda \)-logarithm function (Silva et al., 1998; Yamano, 2002), defined by
\[ \exp\lambda(x) = [1 + (1 - \lambda) \cdot x]^{1/\lambda}, \]
\[ \ln\lambda(x) = 1 - x^{\lambda - 1}. \]
(18a)

We also used the \( \lambda \)-deformed “unity function” (Livadiotis and McComas, 2009), defined by
\[ \lambda_0(x) = [1 + (1 - \lambda) \cdot x]^+. \]
(18b)

The subscript “+” in \( \ldots \) denotes the cut-off condition, where \( \exp\lambda(x) \) becomes zero if its base \( \ldots \) is non-positive. Therefore, Eq. (17b) leads to
\[ p_j^{q-1} = 1 + (1 - q^{-1}) \cdot (\lambda_1 - 1) + (1 - q^{-1}) \cdot \lambda_2 \varepsilon_j, \]
\[ 1 - p_j^{q-1} = \ln \lambda q^{-1}(p_j^{q-1}) = -q \cdot \ln q(p_j^{q-1}) = \lambda_2 \varepsilon_j + (\lambda_1 - 1). \]
(19)

Dividing again the whole system into two subsystems, A and B, using the additivity of energy, and setting \( \lambda_1 = 1 \), the independence relation (Eq. 11) is generalized to
\[ \ln q \left( p_{ij}^{A+B} \right) = \ln q (p_i^A) + \ln q (p_j^B) \]
\[ \Rightarrow \left( p_{ij}^{A+B} \right)^{q-1} = \left( p_i^A \right)^{q-1} + \left( p_j^B \right)^{q-1} - 1, \]
(21)

which is sometimes called the \( q \) independence relationship (Umurov et al., 2008). Then, we apply the operator \( \sum_{i=1}^{W} \sum_{j=1}^{W} p_{ij}^{A+B} \times \),
\[ \sum_{i=1}^{W} \sum_{j=1}^{W} \left( p_{ij}^{A+B} \right)^{q-1} = \sum_{i=1}^{W} \left( p_i^A \right)^{q-1} + \sum_{j=1}^{W} \left( p_j^B \right)^{q-1} - 1 \Rightarrow \varphi^{A+B} = \varphi^A + \varphi^B - 1, \]
(22)

and using the entropic formula (Eq. 14), we end up with the additivity of entropy, as shown in Eq. (13).

Note that the additivity leads to the extensivity: the additivity for a function \( f \) is expressed by \( f(A+B) = f(A) + f(B) \), or considering \( N \) different subsystems,
\[ f \left( \bigcup_{n=1}^{N} A_n \right) = \sum_{n=1}^{N} f(A_n), \]
(23a)

while the extensivity is expressed by
\[ f \left( \bigcup_{n=1}^{N} A_0 \right) = N \cdot f(A_0). \]
(23b)
Therefore, the canonical probability distribution, the one that maximizes the entropy under the constraints of a canonical ensemble, makes the entropy additive (and therefore extensive) if the energy is additive. Several special conditions can simplify this result, e.g., constant Lagrange constraints with \( \lambda_1 = 1 \) (i.e., independent of the probability distribution). This is, however, the case for both the entropic formulation of classical BG and Tsallis nonextensive statistical mechanics.

Next, we reverse the problem and seek to find the specific entropic formula for which both the energy and entropy are additive.

4 Additive energy and entropy leads to Tsallis entropic formalism

The general entropic form is still function of the probabilities: \( S = S\{\{p_k\}_{k=1}^{W}\} \). Then, its derivative with respect to any of the probability components, for example the \( k \)th, is also a function of all of these components, i.e., \( \partial S/\partial p_i = f_i(\{p_k\}_{k=1}^{W}) \), for any \( i: 1, \ldots, W \). However, the second constraint (i.e., fixed internal energy) of the canonical ensemble connects the \( k \)th entropic derivative to a function \( h_k \) of the \( k \)th energy, \( e_k \), namely, \( \partial S/\partial p_i = h_i(e_i) \). On the other hand, the entropy distribution distribution derived from the entropy maximization constitutes an expression of the \( k \)th probability component with some invertible function \( g \) of the \( i \)th energy, \( p_i = g(e_i) \). Therefore, we conclude that \( \partial S/\partial p_i = f_i(p_i) \), where \( f_i = h_i \circ g^{-1} \); in other words, the entropy can be factorized as a summation of functions of each probability component, \( S = \sum_{k=1}^{W} f_k(p_k) \), where we set \( f_k(p_k) = \int f_i(p_i)d\pi_i \). Finally, we consider that none of the states \( (k = 1, \ldots, W) \) should have a special effect on the entropy; i.e., each state “weights” the same, so the entropic functional \( S = S\{\{p_k\}_{k=1}^{W}\} \) should be symmetric to any permutation of each components, e.g., \( S = S\{\ldots, p_k, \ldots, p_k, \ldots\} = S\{\ldots, p_k, \ldots, p_k, \ldots\} \) (i.e., the entropy is invariant under any relabeling of the states). This leads to \( f_k = f^\prime \); hence, considering (1) entropy maximization and (2) no weighting, we obtain

\[
S = \sum_{k=1}^{W} f(p_k). \tag{24a}
\]

For example, in the cases of Boltzmann (Eq. 2) and Tsallis (Eq. 14) entropies, the function \( f \) is respectively given by

\[
f(x) = -x \ln(x) \quad \text{and} \quad f(x) = (x - x^q)/(q - 1). \tag{24b}
\]

The maximization of entropy under the constraints of a canonical ensemble, i.e., \( 1 = \sum_{k=1}^{W} p_k \) and \( U = \sum_{k=1}^{W} p_k e_k \), involves maximizing the functional \( G(\{p_k\}_{k=1}^{W}) = \sum_{k=1}^{W} f(p_k) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k e_k \). Hence, setting \( \partial G(\{p_k\}_{k=1}^{W})/\partial p_i = 0 \), we obtain

\[
F(p_i) + \lambda_1 + \lambda_2 e_i = 0, \quad \text{or} \quad p_i(e_i) = F^{-1}(-\lambda_1 - \lambda_2 e_i),
\]

with \( F(x) \equiv f^\prime(x) \). \tag{25}

We now consider two systems A and B, with respective energy spectra \( \{\epsilon_i^A\}_{i=1}^{W} \) and \( \{\epsilon_j^B\}_{j=1}^{W} \), associated with the discrete probability distributions \( \{p_i^A\}_{i=1}^{W} \) and \( \{p_j^B\}_{j=1}^{W} \). The total system A + B has an energy spectrum \( \{\epsilon_{ij}^{A+B}\}_{ij=1}^{W+W} \), associated with the joint probability distribution \( \{p_{ij}^{A+B}\}_{ij=1}^{W+W} \). The probability distributions \( \{p_i^A\}_{i=1}^{W} \) and \( \{p_j^B\}_{j=1}^{W} \) are marginal of the joint distribution, i.e., \( \sum_{i=1}^{W} p_{ij}^{A+B} = p_i^A \) and \( \sum_{j=1}^{W} p_{ij}^{A+B} = p_j^B \). As we will find further below, the joint probability can be expressed as a function of the marginal probabilities, \( p_{ij}^{A+B} = H(p_i^A, p_j^B) \). On the other hand, the relation between the joint energies \( \epsilon_{ij}^{A+B} \) is rather trivial to be derived: particles in A with energy \( \epsilon_i^A \) and particles in B with energy \( \epsilon_j^B \) ensemble the particles in A + B with energy \( \epsilon_{ij}^{A+B} = \epsilon_i^A + \epsilon_j^B \). Trivially, the same additivity holds for their mean values – the internal energies:

\[
U^{A+B} = \sum_{i,j} p_{ij}^{A+B} \epsilon_{ij}^{A+B} = \sum_{i,j} p_{ij}^{A+B} \epsilon_i^A + \sum_{i,j} p_{ij}^{A+B} \epsilon_j^B
\]

\[
= \sum_i p_i^A \epsilon_i^A + \sum_j p_j^B \epsilon_j^B = U^A + U^B. \tag{26}
\]

Now, the probability distributions are related to their energies, according to Eq. (7). According to Eq. (25), we have

\[
F\left(p_i^A\right) + \lambda_1 + \lambda_2 \epsilon_i^A = 0, \quad F\left(p_j^B\right) + \lambda_1 + \lambda_2 \epsilon_j^B = 0,
\]

\[
F\left(p_{ij}^{A+B}\right) + \lambda_1 + \lambda_2 \epsilon_{ij}^{A+B} = 0, \tag{27}
\]

and due to the additivity of energies, we obtain

\[
F\left(p_{ij}^{A+B}\right) + \lambda_1 = F\left(p_i^A\right) + F\left(p_j^B\right). \tag{28}
\]

Again, the Lagrange constants, \( \lambda_1 \) and \( \lambda_2 \), are considered to be independent of the probability distribution. Setting \( F \equiv
\[
\frac{1}{\lambda_1} F, \text{ Eq. (28) becomes}
\frac{1}{\lambda_1} \left[ F \left( p_{ij}^{A+B} \right) - 1 \right] = \frac{1}{\lambda_1} \left[ F \left( p_i^A \right) - 1 \right] + \frac{1}{\lambda_1} \left[ F \left( p_j^B \right) - 1 \right], \quad \text{or,}
\]
\[p_{ij}^{A+B} = H(p_i^A, p_j^B),\]
\[\text{with } H(x, y) = F^{-1} \left( F(x) + F(y) - 1 \right). \quad (30)\]

Then, we apply \( \sum_{i,j=1}^W \left[ F \left( p_{ij}^{A+B} \right) - 1 \right] p_{ij}^{A+B} = \sum_{i=1}^W \left[ F \left( p_i^A \right) - 1 \right] \sum_{j=1}^W p_{ij}^{A+B} \]
\[+ \sum_{i=1}^W \left[ F \left( p_i^B \right) - 1 \right] \sum_{j=1}^W p_{ij}^{A+B}, \quad \text{or}
\]
\[\sum_{i,j}^W \left[ F \left( p_{ij}^{A+B} \right) - 1 \right] p_{ij}^{A+B} = \sum_i^W \left[ F \left( p_i^A \right) - 1 \right] p_i^A \]
\[+ \sum_j^W \left[ F \left( p_j^B \right) - 1 \right] p_j^B. \quad (31)\]

(Note that the number of states allowed may be different for the two subsystems, \( W_A \neq W_B \), but here it does not make any difference to consider \( W_A = W_B = W \).)

We recall that \( F(x) = \frac{1}{\lambda_1} f' \left( \frac{x}{\lambda_1} \right) \); thus, we find
\[\sum_{i,j}^W \left[ \frac{1}{-\lambda_1} f' \left( \frac{p_{ij}^{A+B}}{-\lambda_1} \right) - 1 \right] p_{ij}^{A+B} = \quad (32)\]
\[\sum_i^W \left[ \frac{1}{-\lambda_1} f' \left( \frac{p_i^A}{-\lambda_1} \right) - 1 \right] p_i^A + \sum_j^W \left[ \frac{1}{-\lambda_1} f' \left( \frac{p_j^B}{-\lambda_1} \right) - 1 \right] p_j^B.\]

We compare this relationship with the additivity of entropy:
\[S^{A+B} = \sum_{i,j}^W f \left( p_{ij}^{A+B} \right) = \sum_i^W f \left( p_i^A \right) + \sum_j^W f \left( p_j^B \right) \]
\[= S^A + S^B. \quad (33)\]

The two functions \( f(x) \) and \( \frac{1}{-\lambda_1} f' \left( \frac{x}{-\lambda_1} \right) - 1 \cdot x \) have the same additivity property. Therefore, one function \( f \) that can ensure the additivity of entropy is the one that obeys conforms to the proportionality, \( f(x) \propto \left[ \frac{1}{-\lambda_1} f' \left( \frac{x}{-\lambda_1} \right) - 1 \right] \cdot x \), or the differential equation
\[f(x) = c \cdot \left[ \frac{1}{-\lambda_1} f' \left( \frac{x}{-\lambda_1} \right) - 1 \right] \cdot x, \quad \text{or}
\]
\[f'(x) + \frac{\lambda_1}{c} f(x) = -\lambda_1, \quad (34)\]

with solution
\[f(x) = \lambda_1 \cdot \frac{x - \frac{\lambda_1}{c}}{c - \lambda_1} + f(1) \cdot x \cdot \frac{-\lambda_1}{c}. \quad (35)\]

(Note that the selection of proportionality between the two functions \( f(x) \) and \( \frac{1}{-\lambda_1} f'(x) - 1 \cdot x \) makes the derivation of Eq. (34) a sufficient but not necessary condition. Other functional forms may also exist, for example, a linear combination of the two functions mentioned.)

A fully organized system has zero entropy, so that \( S(p_i = 1, p_j = 0 \forall j \neq i, \ldots, W, \text{ with } j \neq i = 0 \). Then, from Eq. (24a) we find \( S = 0 = f(1) + (W - 1) f(0) \). Equation (35) gives \( f(0) = 0 \); hence, we find \( f(1) = 0 \). Then, we set \( q = \frac{-\lambda_1}{c} \), where we find
\[f(x) = \lambda_1 \cdot \frac{x - x^q}{q - 1}, \quad (36)\]

or, setting also \( \lambda_1 = 1 \) (that is, setting the entropic unit \( k_B \) equal to 1), we end up with
\[f(x) = \frac{x - x^q}{q - 1}. \quad (37)\]

Therefore, the entropic function \( S = \sum_{k=1}^W f(p_k) \) becomes
\[S = \frac{1}{q-1} \sum_{k=1}^W (p_k - p_k^q), \quad (38)\]

that is, the Tsallis entropic formulation that builds the nonextensive statistical mechanics.

We note that Eq. (33) is invariant under linear transformations:
\[f(x) \rightarrow f(x) + a(x + b) \quad \text{with} \quad a = \frac{\lambda_1 - 1}{1 + (q - 1) \frac{W-1}{W-2}}, \quad \frac{1}{W(W-2)} \quad (39)\]

which leads again to Eq. (38).

5 Conclusions

The paper resolved a basic problem about the origin of the distributions and statistical mechanics applied in space plasmas. Kappa distributions, or combinations/superposition thereof, can describe the velocities and energies of the plasma populations in space plasmas. While these empirical distributions have been used since the mid-1960s for modeling space plasma datasets, their statistical origin remained unknown. It was just about a decade ago that the connection of these distributions with the statistical framework of nonextensive statistical mechanics was completed and understood (Livadiotis, 2017a; chap. 1). Indeed, the kappa distribution is the outcome of the maximization of the \( q \) entropy of Tsallis under the constraints of a canonical ensemble (identifying the \( q \) exponential distributions, first used in a statistical framework context in Tsallis, 1988, as kappa distributions). Once this concept was understood by the science community,
the next question was about the physical origin and reasoning of this entropic formula. This paper showed that the \( q \) entropy, which is the entropic formula that when maximized leads to the kappa distribution, can be derived under simple first principles and conditions, namely, by considering that energy and entropy are both additive physical quantities.

**Competing interests.** The authors declare that they have no conflict of interest.

**Special issue statement.** This article is part of the special issue “Nonlinear Waves and Chaos”. It is a result of the 10th International Nonlinear Wave and Chaos Workshop (NWCW17), San Diego, United States, 20–24 March 2017.

**Acknowledgements.** The work was supported in part by the project NNX17AB74G of NASA’s HGI Program.

Edited by: Yasuhito Narita
Reviewed by: three anonymous referees

**References**


G. Livadiotis: Derivation of the entropic formula for the statistical mechanics of space plasmas


