Dear editor,

We would like to thank all three referees for their comments that helped to improve the paper. We have worked diligently and revised the paper according to all of their comments.

Together with this response, we have included a “track changes” version as well as a clean version of the revised manuscript. In the “track changes” version we mark all the changes-corrections in regards to the referees’ comments.

We are very thankful for your effort to improve our paper.

Best regards,
The authors
Respond to Referee #1

Thank you for your valuable comments that have improved the paper significantly. Thank you also for your good words, especially for finding the paper “very hot” and that it is very important and may become a cornerstone in this field. The revised version has been prepared by taking care all of your comments and corrections.

Together with this response, we have included a “track changes” version as well as a clean version of the revised manuscript. In the “track changes” version we mark all the changes-corrections in regards to the comments of all referees.

George Livadiotis
MS No.: npg-2017-54
Title: “Derivation of the entropic formula for the statistical mechanics of space plasmas”
Author: George Livadiotis

Respond to Referee #2

The author is grateful for your valuable comments that have improved the paper significantly. Thank you also for finding the paper interesting and that deserves publication. The revised version has been prepared by taking care all of your comments and resolving the confusing issues you have mentioned. Below there is a reply to each of your comments separately.

Together with this response, we have included a “track changes” version of the revised manuscript. In the “track changes” version we mark all the changes-corrections in regards to your comments:

Again, thank you for your valuable comments and the effort to improve this work.

George Livadiotis

Reply to each comment:

1. Page 1. Line 2. It is said that the fact that space plasmas follow kappa distributions is a “vastly different statistical behavior between classical systems and space plasmas”. The sentence suggests that space plasmas are very special in this sense, but they are rather just one example of a large family of systems where non-Maxwellian behavior is found. Kappa-like or Tsallis-like distribution functions can be found in spin systems, high-energy physics, turbulent fluids, etc., as well as many examples in biological or social systems. So this should be put in the proper context.

The referee is right. The text has been revised accordingly. Thanks!

2. Page 1. Line 16. “The induction of any type of correlations. . . departs the system from thermal equilibrium to be re-stabilized to other stationary states. . . described by kappa distributions.” This sentence is too strong. When correlations are absent, one should expect that the thermal equilibrium is Maxwellian. However, is there any guarantee, in general, that
(a) any correlation leads to a non-Maxwellian equilibrium?
Yes. This has been proved in Livadiotis & McComas 2011b, and generalized further in Livadiotis 2015c and 2017a, Chapter 5.
(b) if it does, the final state is described by a kappa distribution?
Yes! The existence of particle stationary states characterized by both (i) temperature, and (ii) correlations, means necessarily the formation of kappa distributions (or, combinations thereof). Particle systems, with or without correlations, may exist in other formulations, but they will not be characterized by a physically meaningful definition of temperature as follows from thermodynamic laws.

Besides, is it always correct to say that non-Maxwellian distributions mean absence of thermal equilibrium?
If the system is thermodynamically stabilized into a stationary state, then any non-Maxwellian distribution means: the distribution can be written as a combination or superposition of kappa distributions, the existence of correlations, and the system is not at thermal equilibrium.

Isn’t it possible that a system reaches thermal equilibrium (in the sense that there is no further flow of energy between its components and with its surroundings) while having correlations?
That is actually what happens when the system is reaching a stationary state. The classic understanding of thermal equilibrium is a stationary state, but not all stationary states are in thermal equilibrium. The special about it is no correlations and Maxwellian behavior. All the \( \kappa \) (or \( q \)) stationary states could have similar characteristics with thermal equilibrium but with the highlighting property of local correlations among the particles.

3. Page 2. Line 34. The absence of collisions is not necessary to preserve correlations. On the contrary, they may well be the reason to preserve them.

   Indeed! While thermal collisions destroy correlations, organized collisions may not. However


   In the presence of collisions, energy may be conserved or not. But in the absence of collisions it can be certainly be conserved, unless other particle interactions are taken place.

5. Page 3. Line 1. In principle, energy can be conserved in systems where the energy cannot be separated as the sum of individual particle energies. Please explain.

   You are right! The text has been revised. Thank you!

6. The main claim in the paper seems to be supported by the final paragraphs in Sec. 2 (pages 4 and 5). Although the Tsallis entropy is in general non-additive, there are certain correlations, depending on a certain function \( g(x) \), which make it additive again. And then an expression for the function \( g \) is given, which leads to the Tsallis entropy. However, this leaves the impression that this is a particular case, which allows to recover the Tsallis entropy, but does not help to understand why the Tsallis entropy should be the "right" form to describe correlated systems. Can all physically acceptable correlations be expressed as \( p_{A+Bij} = g^{-1}[g(p_{AI}) + g(p_{BJ})] \)? And is there any argument why one would expect that this holds specifically for plasmas? (Or space plasmas, which is the subject of this paper.) One could make an a posteriori argument, since Tsallis entropy is known to lead to kappa distributions naturally, but the paper seems to make more general claims. Given the above, some of the sentences in the conclusions are not clear. It says that "The paper resolved a basic problem about the origin of the distributions... in space plasmas", and that the q-entropy can be derived by considering additive energy and entropy. However, this seems to be true as long as the assumptions on the correlations [Eq. (21)] are correct, and there is no argument on its validity either for general systems or for plasmas in particular. Thus, it is not clear if the paper resolves a basic problem. Please make more explicit arguments for these statements.

   - The BG entropy and the produced Maxwell-Boltzmann distribution follow a certain correlation type, that is, no correlation, as given by Eq.(11): \[ \ln[p_{y}^{A+B}] = \ln[p_{y}^{A}] + \ln[p_{y}^{B}] \Rightarrow p_{y}^{A+B} = p_{y}^{A} \cdot p_{y}^{B} \]. Section 2 is based on one of the following three equivalent assumptions: (i) BG entropic formulation; (ii) Maxwell-Boltzmann distributions in the canonical ensemble; (iii) no correlation or independence, as given by Eq.(11).

   - The Tsallis entropy and the produced (canonical ensemble) q-exponential or kappa distributions follow a certain correlation type, that is the one given in Eq.(21): \[ \ln_{q}[(p_{y}^{A+B})^{-1}] = \ln_{q}[(p_{y}^{A})^{-1}] + \ln_{q}[(p_{y}^{B})^{-1}] \Rightarrow (p_{y}^{A+B})^{q^{-1}} = (p_{y}^{A})^{q^{-1}} + (p_{y}^{B})^{q^{-1}} - 1 \]. Section 3 is based on one of the following three equivalent assumptions in the canonical ensemble: (i) Tsallis entropic formulation; (ii) q-exponential or kappa distributions; (iii) Special Correlation type called also q-independence, as given by Eq.(21).

   - There is no assumption on correlations to produce the main claim of the paper. The only assumptions are the additivity of the entropy and energy. Section 4 is based on these assumptions to produce that Tsallis entropy describes particle systems such as space plasmas; thus, kappa distributions too; and thus, the correlations of Eq.(21), too. In addition, superposition of stationary states can be described by more complicated distributions or correlations (e.g., see Livadiotis & McComas 2013a; Livadiotis 2017a, Chapter 4).

7. Page 9. Line 11. Entropy is stated to be symmetric on probabilities, arguing that "none of the probability components should have special effect on the entropy". This is too vague and should be rephrased. In the canonical formalism some states are more probable that others. What does "special effect" or "equal weights"
means, then? Maybe it is an argument on the states rather than on the probabilities: relabeling the states does not change the entropy?

Correct, it is an argument on the states, and the text has now been rephrased. Thanks!

8. Tsallis entropy was proposed in the late 80s, and it was always proposed as a way to model the ubiquity of power-law/kappa distributions, understanding that they are “equilibrium” (maximum entropy) configurations, but for an entropy different to the Boltzmann’s one. It is thus not clear why the paper says in the conclusions that “it was just in the last decade that was completely understood that the statistical origin of these distributions is not the Boltzmann-Gibbs’ classical statistical mechanics.” Please explain.

Empirical Kappa distributions have been used without any statistical framework for almost half a century. The connection of these distributions with the statistical framework of non-extensive statistical mechanics is the one that has been completed about a decade ago. Corrections have been made to resolve the confusion. Thanks!

A few additional formal issues:
1. Page 5, line 4: “both the formalisms”.
2. Page 9, line 5: “these component”

All corrected. Thanks!
MS No.: npg-2017-54
Title: “Derivation of the entropic formula for the statistical mechanics of space plasmas”
Author: George Livadiotis

Respond to Referee #3

Thank you for your valuable comments that have improved this work. Thank you also for finding that definitely deserves publication. The revised version has been prepared by taking care all of your comments and corrections.

Together with this response, we have included a “track changes” version as well as a clean version of the revised manuscript. In the “track changes” version we mark all the changes-corrections in regards to all referees’ comments.

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

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

Correspondence to: George Livadiotis (glivadiotis@swri.edu)

Abstract. Kappa distributions describe velocities and energies of plasma populations in space plasmas. The statistical origin of these distributions is the non-extensive statistical mechanics. Indeed, the kappa distribution is derived by maximizing the $q$-entropy of Tsallis under the constraints of canonical ensemble. However, there remains the question what is the physical origin of this entropic formulation. 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 to a collective behavior among particles (e.g., see Jund et al. 1995; Salazar & Toral 1999; Villain 2008; Tsallis 2009; Grassi 2010; Tirnakli & 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 re-stabilized to other stationary states out of thermal equilibrium described by kappa distributions.

Kappa distributions describe numerous space plasma populations. Several examples are the following: (i) 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., Dzifčáková and Dudík 2013; Dzifčáková et al. 2015), 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 solar flares related (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., Ogasawara et al. 2015), magnetotail (e.g., Grabbe 2000), ring current (e.g., Pisarenko et al. 2002), plasma sheet (e.g., Christon 1987; Wang et al. 2003; Kletzing et al. 2003), magnetospheric substorms (e.g., Hapgood et al. 2011), Aurora (e.g., Ogasawara et al. 2017), magnetospheres of giant planets, such as Jovian (e.g., Collier and Hamilton 1995; Mauk et al. 2004), Saturnian (e.g., Dialynas et al. 2009; Liv et al.
Empirical kappa distributions have been introduced in mid-60’s by Binsack (1966), Olbert (1968), and Vasyliūnas (1968), while their connection with statistical mechanics was shown and studied in detail in about half 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 non-extensive statistical mechanics (Tsallis 2009). This is a consistent generalization of the classical statistical mechanics, which is based on a mono-parametric ($q$-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 with the empirical kappa distribution; the two distributions are identical under the transformation of their characteristic indices ($q = 1 + 1/\kappa$).

Having attained a consistent connection of the mathematical model of kappa distributions with the physical means of entropy maximization does not precisely answer the main question regarding the origin of these distributions. We have 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 non-extensive statistical mechanics. Thus, once a kappa
distribution is generated and stabilized in a plasma population, the whole “tool package” of non-extensive 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 about whichever mechanisms generate kappa distributions in space plasmas, but for 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. The collisionless environment is preserving the energy. Moreover, weakly coupled plasmas (mutual electron and ion potential energy is small compared to the average kinetic energy) can be described as an ideal gas. 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 involved one-particle states. 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 Murtaza 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 involved one-particle states.

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 (Figure 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 on the fact that the kappa 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 Section 2 we describe the physical motive of this paper in detail. In Section 3 we show in detail a similar property for both the entropic formalisms of Boltzmann Gibbs (BGs) 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 Section 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, Section 5 briefly summarizes the conclusions.
Figure 1: The infogram indicates the following triplet of concepts: (i) additive energy, (ii) additive entropy, and (iii) BG or Tsallis entropic formulation. Given any two out of the three features, the third can be derived. It is already known that BG or Tsallis entropic formulations can lead to additive entropy 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).

2 Physical Motive

Classical Boltzmann–Gibbs (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 independent discrete distributions; i.e., labelling the two particles with A and B, $p_{ij}^{AB} = p_i^A \cdot p_j^B$. Hereafter, we consider a particle system described by a discrete energy spectrum $\{\varepsilon_k\}_{k=1}^W$, 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, $\ln p_{ij}^{AB} = \ln p_i^A + \ln p_j^B$, from which we obtain the additivity of entropy, $S^{AB} = S^A + S^B$. For special cases, however, where the independent relation does not apply, $p_{ij}^{AB} \neq p_i^A \cdot p_j^B$, the entropy is non-additive, $S^{AB} \neq S^A + S^B$. The logical reciprocate to the statements above is provided by the uniqueness theorem of Shannon (1948) and Khinchin (1957) that showed that under the assumption of the additivity of entropy (and other basic properties of entropy), the sufficient and necessary entropic form is given by the BG formula.

Non-extensive statistical mechanics characterizes systems with correlations among particles, $p_{ij}^{AB} \neq p_i^A \cdot p_j^B$. For special systems, however, where the independent relation still applies, the entropy is non-additive, $S^{AB} \neq S^A + S^B$; in particular, a square, nonlinear term is added to the summation, $S^{AB} = S^A + S^B + (1-q)S^A S^B$ for some value of the entropic parameter $q$ (where we set the Boltzmann constant $k_B$ to 1). Note that the logical reciprocates exists also for this
case, as shown by dos Santos (1997) and Abe (2000); namely, under the assumption of the mentioned non-additive property (and other basic properties of entropy), the sufficient and necessary entropic form is given by the q-entropic formula of Tsallis (1988).

Another property that is related with the additivity but is even more subtle and difficult to ascertain is the extensivity of the entropy. A non-additive entropy may be assumed to be also non-extensive, but it is the inverse assumption that is always correct, i.e., non-extensivity implies non-additivity). Nevertheless, certain correlations, expressed by the relation $p_{ij}^{A+B} = g^{-1}[g(p_i^A) + g(p_j^B)]$ for some function $g$, can make the Tsallis entropy additive, and thus, recover its extensivity (e.g., Tsallis et al. 2005, Ruseckas 2015). In his book, C. Tsallis (2009) goes to great lengths to show that it is possible to find systems for which the BG entropy is not extensive. On the other hand, he argues that there are certain systems for which the entropic form can be extensive, for certain values of the entropic index $q$. In fact, he mentions in the preface that the term “nonextensive entropy” is somewhat incorrect in this sense, but it stuck for historical reasons.

The two statistical formalisms, classical BG and Tsallis non-extensive, have the common property that their entropy becomes additive for some specific function $g$ in the relation $p_{ij}^{A+B} = g^{-1}[g(p_i^A) + g(p_j^B)]$, that is, $g(x) \propto \ln(x)$ and $g(x) \propto \frac{1}{x^{q-1}} - 1$, respectively; (the latter is related to the $q$-deformed logarithm; see: Silva et al. 1998; Yamano 2002).

It is important that the above probability relation is a characteristic feature of the canonical probability distribution in both the formalisms. In other words, the probability distribution that maximizes the BG entropy under the constraints of the canonical ensemble obeys to correlations expressed by $g(x) \propto \ln(x)$ or $p_{ij}^{A+B} = p_i^A \cdot p_j^B$, which means zero correlation (due to the factorization of the exponentials, Livadiotis and McComas 2011b) that makes the entropy additive. Also, the probability distribution that maximizes the Tsallis entropy under the same constraints obeys to specific correlations expressed by $g(x) \propto \frac{1}{x^{q-1}} - 1$ or $(p_{ij}^{A+B})^{q-1} = (p_i^A)^{q-1} + (p_j^B)^{q-1} - 1$, which makes again the entropy additive. In Section 3 we show in detail this similar property of the two statistical formalisms.

Then, we may ask: Is the above described property of BG and Tsallis entropies a general feature of any physically meaningful entropic function? Or, can we reverse the question, and ask which specific entropic function obeys to 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 Section 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 $S(p_1, p_2, ..., p_W)$ under the constraints of canonical ensemble, i.e., (i) normalization $1 = \sum_{k=1}^{W} p_k$, and (ii) fixed internal energy $U = \sum_{k=1}^{W} p_k \epsilon_k$, involves maximizing the functional

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

(1)

Next, we examine the BG and Tsallis entropic formulations.

3.2 BG entropy

First, we start from the classical case of BG entropy

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

(2)

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

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

(3)

we find

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

(4)

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

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

(5)

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

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

(6)

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

$$\epsilon_{ij}^{A+B} = \epsilon_i^A + \epsilon_j^B.$$  

(7)

Hence, by eliminating energies from Eqs.(5,6), we find
\[
\ln p^{A+B}_\theta + (\lambda_1 - 1) = \lambda_2 e^A_i + (\lambda_1 - 1) + \lambda_2 e^B_j + (\lambda_1 - 1) = \ln p^A_i + \ln p^B_j, \quad \text{or} \tag{8}
\]
\[
p^{A+B}_\theta = p^A_i \cdot p^B_j \cdot e^{-(\lambda_1-1)}. \tag{9}
\]

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.(8,9), we obtain

\[
\ln p^{A+B}_\theta = \lambda_2 e^A_{\theta} + (\lambda_1 - 1)^{A+B} = \lambda_2 e^A_i + (\lambda_1 - 1)^A + \lambda_2 e^B_j + (\lambda_1 - 1)^B = \ln p^A_i + \ln p^B_j, \tag{10}
\]

which clearly shows that the canonical probabilities are independent,

\[
\ln(p^{A+B}_\theta) = \ln(p^A_i) + \ln(p^B_j) \Rightarrow p^{A+B}_\theta = p^A_i \cdot p^B_j. \tag{11}
\]

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 \(\sum_{i=1}^W \sum_{j=1}^W p^{A+B}_\theta \times\) on both sides of Eq.(11), we obtain

\[
p^{A+B}_\theta \ln p^{A+B}_\theta = p^A_i \ln p^A_i + p^B_j \ln p^B_j
\]

\[
\Rightarrow -\sum_{i=1}^W \sum_{j=1}^W p^{A+B}_\theta \ln p^{A+B}_\theta = -\sum_{i=1}^W p^A_i \ln(p_i^A) - \sum_{j=1}^W p^B_j \ln(p_j^B). \tag{12}
\]

because \(\sum_{j=1}^W p^{A+B}_\theta = p^A_i, \sum_{i=1}^W p^{A+B}_\theta = p^B_j\). Hence, we arrive at the additivity of the entropy of the system to the entropies of the subsystems,

\[
S^{A+B} = S^A + S^B. \tag{13}
\]

3.3 Tsallis entropy

Next, we continue with the Tsallis \(q\)-entropy,

\[
S(p_1, p_2,..., p_w) = \frac{1}{q-1} \phi(p_1, p_2,..., p_w) = \frac{1}{q-1} \sum_{k=1}^w (p_k - p_k^q) \tag{14}
\]

(e.g., Havrda and Charvát 1967; Tsallis 1988), where the argument \(\phi\) is defined by

\[
\phi(p_1, p_2,..., p_w) = \sum_{k=1}^w p_k^q. \tag{15}
\]

Again, the maximization of the entropy under the constraints of canonical ensemble involves maximizing the functional
\[
G(p_1, p_2, ..., p_W) = \frac{1}{q-1} \sum_{k=1}^{W} (p_k - p_k^q) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k^q \lambda_k \cdot \epsilon_k.
\] (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 the modern nonextensive statistical mechanics (Livadiotis 2017a; Chapter 1). It was 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, \((\partial / \partial p_j) G(p_1, p_2, ..., p_W) = 0\), gives

\[
p_j(\epsilon_j) = \left[ 1 + (1 - q^{-1}) \cdot (\lambda_1 - 1) \right]^{q^{-1}} \cdot \left[ 1 + (1 - q^{-1}) \cdot \frac{\lambda_j \epsilon_j}{1 + (1 - q^{-1}) \cdot (\lambda_1 - 1)} \right]^{q^{-1}}.
\] (17a)
or

\[
p_j(\epsilon_j) = \exp_{q^{-1}}(\lambda_1 - 1) \cdot \exp_{q^{-1}} \left[ \frac{\lambda_j \epsilon_j}{1 + (1 - q^{-1}) \cdot (\lambda_1 - 1)} \right],
\] (17b)

where reflects a generalization of Eq.(4). We used the \(Q\)-deformed exponential function, and its inverse, the \(Q\)-logarithm (Silva et al. 1998; Yamano 2002), defined by

\[
\exp_Q(x) = \left[ 1 + (1 - Q) \cdot x \right]^{1/Q}, \quad \ln_Q(x) = \frac{1 - x^{-Q}}{Q-1}.
\] (18a)

We also used the \(Q\)-deformed “unity function” (Livadiotis and McComas 2009), defined by

\[
1_Q(x) = \left[ 1 + (1 - Q) \cdot x \right],
\] (18b)

The subscript “+” in \([...]+\) denotes the cut-off condition, where \(\exp_Q(x)\) becomes zero if its base \([...]\) is non-positive. Therefore, Eq.(17b) leads to

\[
p_j^{-1} = 1 + (1 - q^{-1}) \cdot (\lambda_1 - 1) + (1 - q^{-1}) \cdot \lambda_j \epsilon_j,
\] (19)

\[
1 - p_j^{-1} = \ln_q(p_j^{-1}) = -q \cdot \ln_q(p_j^{-1}) = \lambda_j \epsilon_j + (\lambda_1 - 1),
\] (20)

Dividing again the whole system in two subsystems A and B, using the additivity of energy, and setting \(\lambda_1 = 1\), the independence relation (11) is generalized to

\[
\ln_q[(p_j^{A+B})^{-1}] = \ln_q[(p_j^A)^{-1}] + \ln_q[(p_j^B)^{-1}] \Rightarrow (p_j^{A+B})^{-1} = (p_j^A)^{-1} + (p_j^B)^{-1} - 1,
\] (21)

which is sometimes called \(q\)-independence relation (Umarov et al. 2008). Then, we apply the operator \(\sum_{i=1}^{W} \sum_{j=1}^{W} p_j^{A+B} \times \),

\[
\sum_{i=1}^{W} \sum_{j=1}^{W} (p_j^{A+B})^q = \sum_{i=1}^{W} (p_i^{A})^q + \sum_{j=1}^{W} (p_j^B)^q - 1 \Rightarrow \phi^{A+B} = \phi^A + \phi^B - 1,
\] (22)
and using the entropic formula (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 some function \( f \) is expressed by \( f(A + B) = f(A) + f(B) \), or considering \( N \) different subsystems,

\[
f\left( \bigcup_{n=1}^{N} A_n \right) = \sum_{n=1}^{N} f(A_n) ,
\]

while the extensivity is expressed by

\[
f\left( \bigcup_{n=1}^{N} A_n \right) = N \cdot f(A_n) .
\]

Therefore, the canonical probability distribution, the one that maximizes the entropy under the constraints of 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 \( \lambda_1=1 \) (independent of the probability distribution). This is true for both the entropic formulation of classical BG and Tsallis nonextensive statistical mechanics.

Next, we will try to 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, let’s say the \( i \)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 2nd constraint (fixed internal energy) of the canonical ensemble connects the \( i \)th entropic derivative to some function \( h_i \) of the \( i \)th energy, \( \varepsilon_i \), namely, \( \partial S / \partial p_i = h_i(\varepsilon_i) \). On the other hand, the canonical probability distribution derived from the entropy maximization constitutes an expression of the \( i \)th probability component with some invertible function \( g \) of the \( i \)th energy, \( p_i = g(\varepsilon_i) \). 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

\[
\begin{align*}
\text{we set} & \quad f_i(p_i) = \int F_i(p_i) dp_i . \\
\text{Finally, we consider that none of the probability components states} \ (k=1, \ldots, W) & \text{should have special effect on the entropy, i.e., the whole distribution each state “weights” in the same way, so that the entropic functional} \\
S = S(\{p_k\}_{k=1}^{W}) & \text{should be symmetric to any permutation of each components, e.g.,} \\
S = S(\ldots, p_k, \ldots, p_i, \ldots) & = S(\ldots, p_i, \ldots, p_k, \ldots) \ (i.e., \text{the entropy is invariant under any relabelling of the states}) . \\
\end{align*}
\]

This leads to \( f_k = f \); hence, considering (1) Entropy maximization, (2) No weighting, we obtain

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

For example, in the cases of Boltzmann (Eq.(2)) and Tsallis (Eq.(14)) entropies, function \( f \) is respectively given by:
The maximization of entropy under the constraints of canonical ensemble, i.e.,

\[ 1 = \sum_{k=1}^{W} p_k \quad \text{and} \quad U = \sum_{k=1}^{W} p_k \varepsilon_k, \]

involves maximizing the functional

\[ G(p_k) = \sum_{k=1}^{W} f(p_k) + \lambda_1 \sum_{k=1}^{W} p_k + \lambda_2 \sum_{k=1}^{W} p_k \varepsilon_k. \]

Hence, setting \( \partial G(\{p_k\})/\partial p_i = 0 \), we obtain

\[ F(p_i) + \lambda_1 + \lambda_2 \varepsilon_i = 0, \]

or \( p_i(x) = F^{-1}(\varepsilon_i), \) with \( F(x) = f'(x). \)

We now consider two systems A and B, with respective energy spectra \( \varepsilon_A^i \) and \( \varepsilon_B^j \), associated with the discrete probability distributions \( W_{A(i)}^i \) and \( W_{B(j)}^j \). The total system A+B has energy spectrum \( \varepsilon_{A+B}^{ij} \), associated with the joint probability distribution \( W_{A+B}^{ij} \). The probability distributions \( W_{A(i)}^i \) and \( W_{B(j)}^j \) are marginal of the joint distribution, i.e.,

\[ \sum_{j} W_{A+B}^{ij} = W_A^i \quad \text{and} \quad \sum_{i} W_{A+B}^{ij} = W_B^j. \]

As we will find further below, the joint probability can be expressed as a function of the marginal probabilities, \( p_{A+B}^{ij} = H(p_A^i, p_B^j) \). On the other hand, the relation between the joint energies \( \varepsilon_{A+B}^{ij} \) is rather trivial to be derived: particles in A with energy \( \varepsilon_A^i \) and particles in B with energy \( \varepsilon_B^j \) ensemble the particles in A+B with energy \( \varepsilon_{A+B}^{ij} = \varepsilon_A^i + \varepsilon_B^j \). Trivially, the same additivity holds for their mean values – the internal energies,

\[ U_{A+B} = \sum_{ij} p_{A+B}^{ij} \varepsilon_A^i + \sum_{ij} p_{A+B}^{ij} \varepsilon_B^j = \sum_i p_A^i \varepsilon_A^i + \sum_j p_B^j \varepsilon_B^j. \]

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

\[ F(p_A^i) + \lambda_1 + \lambda_2 \varepsilon_A = 0, \]

and due to the additivity of energies, we obtain

\[ F(p_{A+B}^{ij}) - \lambda_1 = F(p_A^i) + F(p_B^j). \]

Again, the Lagrange constants \( \lambda_1 \) and \( \lambda_2 \), are considered to be constant independent of the probability distribution. Setting

\[ \bar{F} \equiv -\frac{1}{\lambda_1} F, \]

Eq.(28) becomes

\[ [\bar{F}(p_{A+B}^{ij}) - 1] = [\bar{F}(p_A^i) - 1] + [\bar{F}(p_B^j) - 1], \]

or,

\[ p_{A+B}^{ij} = H(p_A^i, p_B^j), \quad \text{with} \quad H(x, y) \equiv \bar{F}^{-1} - 1. \]

Then, we apply \( \sum_{i} \sum_{j} p_{A+B}^{ij} \times \) in both sides of Eq.(29),
\[
\sum_{i,j}^W [\bar{F}(p_{ij}^{A+B}) - 1] p_{ij}^{A+B} = \sum_{i=1}^W [\bar{F}(p_i^A) - 1] \sum_{j=1}^W p_{ij}^{A+B} + \sum_{j=1}^W [\bar{F}(p_j^B) - 1] \sum_{i=1}^W p_{ij}^{A+B}, \quad \text{or}
\]
\[
\sum_{i,j}^W [\bar{F}(p_{ij}^{A+B}) - 1] p_{ij}^{A+B} = \sum_i^W [\bar{F}(p_i^A) - 1] p_i^A + \sum_j^W [\bar{F}(p_j^B) - 1] p_j^B.
\]
(31)

(Note: The number of allowed states 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 \( \bar{F}(x) \equiv \frac{1}{\lambda_x} f'(x) \), thus, we find
\[
\sum_{i,j}^W \left[ \frac{1}{\lambda_{ij}} f'(p_{ij}^{A+B}) - 1 \right] p_{ij}^{A+B} = \sum_{i=1}^W \left[ \frac{1}{\lambda_i} f'(p_i^A) - 1 \right] p_i^A + \sum_{j=1}^W \left[ \frac{1}{\lambda_j} f'(p_j^B) - 1 \right] p_j^B.
\]
(32)

We compare this property relation with the additivity of entropy
\[
S_{A+B} = \sum_{i,j}^W f(p_{ij}^{A+B}) = \sum_i^W f(p_i^A) + \sum_j^W f(p_j^B) = S_A + S_B.
\]
(33)

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

with solution
\[
f(x) = \lambda_x \cdot \frac{x - \frac{x}{e^{\lambda_x}}} {e^{\lambda_x} - 1} + f(1) \cdot x^{\frac{\lambda_x}{e^{\lambda_x} - 1}}.
\]
(35)

(Note: The selection of proportionality between the two functions \( f(x) \) and \( \left[ \frac{1}{\lambda_x} f'(x) - 1 \right] \cdot x \) 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 mentioned functions.)

A fully organized system has zero entropy, so that \( S(p_j = 1, p_j = 0 \ \forall j : 0, \ldots, 1, \text{with} \ j \neq i) = 0 \). Then, from Eq.(24a) we find \( S = 0 = f(1) + (W - 1)f(0) \). Equation (35) gives \( f(0) = 0 \), hence, we find \( f(1) = 0 \), too. Then, we choose \( f(x - 1) = 0 \), and we set \( q \equiv \frac{\lambda_x}{e} \), where we find
\[
f(x) = \lambda_x \cdot \frac{x - \frac{x}{q}} {q - 1}, \quad \text{(36)}
\]
or, setting also \( \lambda_x = 1 \) (that is, setting the entropic unit \( k_B \) equal to 1).
\[ f(x) = \frac{x - x^q}{q-1}. \quad (37) \]

Therefore, the entropic function \( S = \sum_{k=1}^{W} f(p_k) \) becomes

\[ S = \frac{1}{q-1} \sum_{k=1}^{W} \left(p_k - p_k^q\right), \quad (38) \]

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

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 thereof, can describe the velocities and energies of the plasma populations in space plasmas. While these empirical distributions were used since mid-60’s for modeling space plasma datasets, their physical statistical origin was remaining unknown. It was just the last about a decade ago that the connection of these distributions with the statistical framework of non-extensive statistical mechanics has been completed and understood was completely understood that the statistical origin of these distributions is not the Boltzmann-Gibbs classical statistical mechanics, but the Tsallis non-extensive statistical mechanics (Livadiotis 2017a; Chapter 1). Indeed, the kappa distribution is the outcome of the maximization of the \( q \)-entropy of Tsallis under the constraints of 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 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.

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

References


Binsack, J. H. 1966 Plasma studies with the IMP-2 satellite, Ph.D. Thesis, MIT.


Livadiotis, G. 2015b Kappa distribution in the presence of a potential energy. J. Geophys. Res. 120, 880.

Livadiotis, G. 2015c Kappa and q indices: Dependence on the degrees of freedom. Entropy 17, 2062.

Livadiotis, G. 2016 Curie law for systems described by kappa distributions. EPL 113, 10003.


Livadiotis, G. and McComas, D.J. 2010b Measure of the departure of the q-metastable stationary states from equilibrium. Phys. Scr. 82, 035003.


Livadiotis, G. and McComas, D.J. 2013b Evidence of large scale phase space quantization in plasmas. Entropy 15, 1118-1132.


Tsallis, C., Gell-Mann, M. and Sato, Y. 2005 Asymptotically scale-invariant occupancy of phase space makes the entropy $S_q$ extensive. PNAS 102, 15377–15382.


