Nonextensive statistical mechanics and high energy physics

Constantino Tsallis\textsuperscript{1,2,a} and Zochil Gonzalez Arenas\textsuperscript{1,b}

\textsuperscript{1}Centro Brasileiro de Pesquisas Físicas, and National Institute of Science and Technology for Complex Systems - Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro-RJ, Brazil
\textsuperscript{2}Santa Fe Institute - 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Abstract. The use of the celebrated Boltzmann-Gibbs entropy and statistical mechanics is justified for ergodic-like systems. In contrast, complex systems typically require more powerful theories. We will provide a brief introduction to nonadditive entropies (characterized by indices like \(q\), which, in the \(q \to 1\) limit, recovers the standard Boltzmann-Gibbs entropy) and associated nonextensive statistical mechanics. We then present some recent applications to systems such as high-energy collisions, black holes and others. In addition to that, we clarify and illustrate the neat distinction that exists between Lévy distributions and \(q\)-exponential ones, a point which occasionally causes some confusion in the literature, very particularly in the LHC literature.

1 Introduction

Boltzmann-Gibbs (BG) statistical mechanics constitutes one of the pillars of contemporary physics, together with Newtonian, quantum and relativistic mechanics, and Maxwell’s electromagnetism. Central points of any physical theory are when and why it works, and when and why it fails. This is so for any human intellectual construct, hence for the BG theory as well. We intend to briefly discuss here when we must (or must not) use the BG entropy

\[
S_{BG} = k \sum_{i=1}^{W} p_i \ln \left( \frac{1}{p_i} \right) \left( \sum_{i=1}^{W} p_i = 1 \right),
\]

where \(k\) is a constant (either taken equal to Boltzmann constant \(k_B\), or to unity). For equal probabilities, i.e., \(p_i = 1/W\), we obtain the celebrated expression carved on Boltzmann’s grave in Vienna:

\[
S_{BG} = k \ln W.
\]

There are basic thermodynamic (see [1] and references therein) as well as large-deviation-theory (see [41] and references therein) reasons for which we expect the thermodynamical entropy \(S\) to always be extensive, i.e., \(S(N) \propto N\) (\(N \to \infty\)), where \(N\) is the number of elements of the system (\(N\) being in turn proportional to \(L^d\), where \(L\) is a linear size of the system, and \(d\) its euclidean or fractal dimension). It is then obvious that, for systems such that \(W(N) \propto \mu^N\) (\(\mu > 1\)), we must use the BG entropic functional

\[\text{e-mail: tsallis@cbpf.br}\]
\[\text{b-e-mail: zochil@cbpf.br}\]
Table 1. In order to satisfy extensivity for the thermodynamic entropy we must use $S_{BG}$, $S_q$ and $S_\delta$ for the exponential [$W(N) \propto \mu^N$ with $\mu > 1$], power-law [$W(N) \propto N^\rho$ with $\rho > 0$] and stretched-exponential [$W(N) \propto N^{\gamma N}$ with $\gamma > 1$] classes respectively.

<table>
<thead>
<tr>
<th>$S_{BG}$ (additive)</th>
<th>$S_q$ (nonadditive for $q \neq 1$)</th>
<th>$S_\delta$ (nonadditive for $\delta \neq 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>extensive</td>
<td>nonextensive</td>
<td>nonextensive</td>
</tr>
<tr>
<td>nonextensive</td>
<td>extensive for $q = 1 - \frac{1}{\rho}$</td>
<td>nonextensive</td>
</tr>
<tr>
<td>nonextensive</td>
<td>nonextensive</td>
<td>extensive for $\delta = \frac{1}{\gamma}$</td>
</tr>
</tbody>
</table>

(1), hence (2), since it yields $S_{BG}(N) \propto N$. This class of systems corresponds to the existence of weak correlations (or no correlations at all) between the elements of the system.

If the elements of the system are, in contrast, strongly correlated, it might occur that $W(N) \propto N^\rho$ ($\rho > 0$). In this case, if we use Eq.(2), we shall obtain $S(N) \propto \ln N$, which violates thermodynamics. Now, it was introduced in 1988 [3] the following entropic functional:

$$S_q = k \frac{1 - \sum_{i=1}^{W} p_i^q}{q - 1} = k \sum_{i=1}^{W} p_i \ln_q \frac{1}{p_i} \left( S_1 = S_{BG} \right),$$

hence, for equal probabilities,

$$S_q = k \frac{W^{1-q} - 1}{1 - q} = k \ln_q W,$$

where $\ln_q z = \frac{z^{1-q} - 1}{1-q}$ ($\ln_1 z = \ln z$).

If our system belongs to the class $W(N) \propto N^\rho$, we must use $S_q$ with $q = 1 - 1/\rho$, since $S_{q=1-1/\rho}(N) \propto N$, in conformity with thermodynamics.

Let us remark that, if $A$ and $B$ are two probabilistically independent systems, i.e., $p_{ij}^{A+B} = p_i^A p_j^B$, we verify that

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

Consequently, unless $q = 1$, $S_q$ is nonadditive. Summarizing the situation, remarkably enough, for the class $W(N) \propto N^\rho$, in order to have an extensive thermodynamic entropy we must use a nonadditive entropic functional such as $S_q$ and not the (additive) BG one! At this point let us mention another (quite striking) fact, namely that $S_q$ is directly related to the Riemann zeta function [4].

Another strongly correlated class corresponds to $W(N) \propto N^{\gamma N}$ ($\gamma > 1$). No value of $q$ can produce an extensive thermodynamic entropy in this case. But a different nonadditive entropy (to the best of our knowledge, independently introduced in the footnote of page 69 of [5], and in [6]), namely $S_\delta = k \sum_{i=1}^{W} p_i \left( \ln \frac{1}{p_i} \right)^\delta$, produces precisely that for $\delta = 1/\gamma$. Indeed, $S_{\delta=1/\gamma}(N) \propto N$. The situation is summarized in Table 1.

From the dynamical viewpoint, the entropy $S_{BG}$ must be used for ergodic systems (typically for systems whose maximal Lyapunov exponent is positive), whereas for nonergodic systems (typically for systems whose maximal Lyapunov exponent vanishes), the occupancy of the full phase space typically corresponds to a vanishing Lebesgue measure, and nonadditive entropies such as $S_q$ and $S_\delta$ must be used. In fact, $S_q$ and $S_\delta$ can be unified into a two-indices entropic functional $S_{q,\delta} = k \sum_{i=1}^{W} p_i \left( \ln \frac{1}{p_i} \right)^\delta$ such that $S_{1,1} = S_{BG}$, $S_{q,1} = S_q$, and $S_{1,\delta} = S_\delta$ [1] (see also [7–11]).

What about the stationary distributions associated with these entropies for appropriate simple constraints (such as fixed mean total energy)? In the $S_{BG}$ case, the stationary state (in this case, currently referred to as thermal equilibrium) is given by the celebrated canonical BG factor, namely
\[ p(E_i) \propto e^{-E_i/kT}, \quad \text{where } E_i \text{ denotes the total energy of a given admissible state associated with the Hamiltonian system, and } T \text{ denotes the temperature of the thermostat with which the system is in thermal contact.} \]

In the \( S_q \) case, the stationary state is essentially given \([3]\) (see details in \([12]\) and in \([5]\)) by \( p_q(E_i) \propto e^{-E_i/kT_q} \), where \( T_q \) is the effective temperature of the system, and \( e_q^z \) is the inverse function of \( \ln_q z \), i.e., \( e_q^z \equiv [1 + (1 - q)z]^{\frac{1}{q}} \) (\( e_1^z = e^z \)). In the \( S_\delta \) case, the situation is totally analogous \([13]\).

The \( q \)-generalized statistical mechanics based on the nonadditive entropy \( S_q \) is currently referred to as nonextensive statistical mechanics (the word nonextensive refers here to the typically nonextensive total energy, by no means to the entropy, which remains extensive in all cases). There is a vast literature on the subject (see \([14]\)).

Before focusing, in the next Section, onto applications, it is worthy to stress an important fact. The standard Central Limit Theorem (CLT) shows that the sum of \( N \gg 1 \) independent (or quasi-independent in some sense) random variables with finite variance has an unique attractor in the space of probability distributions. This attractor is, after centering and scaling, a Gaussian distribution. This is widely believed to be the mathematical reason for which so many Gaussians are found in nature. An accompanying theorem, sometimes referred to as Lévy-Gnedenko theorem, still assumes independence, but the variance is supposed to diverge. Under this main hypothesis (and some supplementary mathematical hypothesis) the attractor of the sum is shown to be a Lévy distribution, also called \( \alpha \)-stable distribution. What happens if the hypothesis of independence is violated? The attractors will be different ones and will depend on the class of correlations. If we have a specific, and nevertheless ubiquitous, class of correlations referred to as \( q \)-independence (1-independence corresponds to the usual probabilistic independence), the attractors are shown to be \( q \)-Gaussians if a \( q \)-generalized variance is finite \([15]\), and \( (q, \alpha) \)-stable distributions if this same \( q \)-generalized variance diverges \([16]\). Lévy \( (0 < \alpha < 2) \), \( q \)-Gaussian \( (q > 1) \) and \( (q, \alpha) \)-stable distributions \( (q > 1 \text{ and } 0 < \alpha < 2) \) — as well as infinitely many other distributions! — asymptotically behave as power-laws, i.e., they exhibit fat tails. They are however definitively different, a fact which can be easily checked in the central and intermediate regions of the distributions.

The so \( q \)-generalized CLT (referred to as the \( q \)-CLT) is very important in its consequences, and has been the subject of various papers in the literature (which are out of the scope of the present brief review). Its importance is based on the fact that it strongly suggests that many \( q \)-Gaussians (and consistently many \( q \)-exponentials) should be found in nature. In the next Section we show that it is indeed so.

## 2 Applications

The present theory has received a large number of analytical, experimental, observational and computational predictions, verifications and applications in natural, artificial and social systems. We briefly mention here some selected ones: cold atoms in optical lattices \([17]\), trapped ions \([18]\), asteroid motion and size \([19]\), motion of biological cells \([20]\), edge of chaos \([21–31]\), restricted diffusion \([32]\), defect turbulence \([33]\), solar wind \([34]\), dusty plasma \([35, 36]\), spin-glass \([37]\), overdamped motion of interaction particles \([38]\), tissue radiation \([39]\), nonlinear relativistic and quantum equations \([40]\), large deviation theory \([41]\), long-range-interacting classical systems \([42–46]\), microcalcification detection techniques \([47]\), ozone layer \([48]\), scale-free networks \([49–51]\), among others.

In high-energy physics there is a plethora of applications as well (\([52–65]\) among many others). A remarkable recent example is shown in Fig. 1 from \([66]\), where a satisfactory fitting with experimental data goes along impressive 14 decades of the probability distribution of the hadronic transverse momenta.
Figure 1. From [66] (courtesy of the authors). In all cases, quite satisfactory fittings are obtained along amazingly many decades. The typical value for the power-law index is $n \approx 7$, which corresponds (assuming $\eta = 1$, that is defined later on) to an entropic index $q \approx 1.15$.

All of the above applications concern $S_q$. But there have recently appeared applications also of $S_\delta$, focusing on the thermodynamics of black holes [1, 67].

3 Lévy distributions versus $q$-exponential distributions

3.1 Lévy distributions

The normalized alpha-stable distribution $L(x; \alpha, \beta, c, \mu)$ is a four-parameter family of continuous distributions parametrized by location and scale (or width) parameters $\mu \in \mathbb{R}$ and $c > 0$, respectively, and two shape parameters $\beta \in [-1, 1]$ and $\alpha \in (0, 2]$, which measure asymmetry (or skewness) and concentration, respectively. These distributions are usually referred to as the Lévy alpha-stable ones and include, as particular cases, the normal ($\alpha = 2$ and $\beta = 0$), the Cauchy-Lorentz ($\alpha = 1$ and $\beta = 0$) and the inverse-gamma ($\alpha = 1/2$ and $\beta = 1$) distributions.

With the exception of the above three cases, the distribution function of an alpha-stable random variable cannot be given in closed form. The family of distributions is defined through its characteristic $\varphi(t)$ [68, 69]. For $\alpha \neq 1$, we have

$$\varphi(t; \alpha, \beta, c, \mu) = \exp \left[ i \mu t - c^\alpha |t|^{\alpha} (1 - i \beta \text{sign}(t) \tan(\frac{\pi \alpha}{2})) \right],$$

and, for the case $\alpha = 1$, we have

$$\varphi(t; \alpha, \beta, c, \mu) = \exp \left[ i \mu t - c^\alpha |t|^\alpha (1 + i \beta \text{sign}(t) \frac{2}{\pi} \log |t|) \right].$$
Figure 2. Typical symmetric Lévy distributions.

The distributions are therefore given by

$$L(x; \alpha, \beta, c, \mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \ e^{-itx} \varphi(t; \alpha, \beta, c, \mu)$$  \hspace{1cm} (8)

Notice the following symmetry: $L(x; \alpha, \beta, c, \mu) = L(-x; \alpha, -\beta, c, \mu)$. Also, it can be verified that, in the $x \to \infty$ limit,

$$L(x; \alpha, \beta, c, \mu) \propto \frac{1}{x^{1+\alpha}} \ (0 < \alpha < 2; \ 0 \leq \beta \leq 1; \ \forall c; \ \forall \mu).$$ \hspace{1cm} (9)

Typical examples of the Lévy family are indicated in Figs. 2 and 3. Finally, let us write down the only three Lévy distributions which admit a closed form:

(i) Gaussian (or normal) distribution:

$$L(x;2,\beta,c,\mu) = \frac{1}{\sqrt{4\pi c}} \exp\left(-\frac{(x-\mu)^2}{4c^2}\right) \ (-\infty < x < \infty; \ \forall \beta).$$ \hspace{1cm} (10)

(ii) Cauchy-Lorentz distribution:

$$L(x;1,0,c,\mu) = \frac{c}{\pi c^2 + (x-\mu)^2} \ (-\infty < x < \infty).$$ \hspace{1cm} (11)

(iii) Inverse-gamma distribution:

$$L(x;1/2,1,c,\mu) = \sqrt{c} \frac{1}{2\pi (x-\mu)^{3/2}} \exp\left(-\frac{c}{2(x-\mu)}\right) \ (\mu < x < \infty).$$ \hspace{1cm} (12)
**Figure 3.** Typical Lévy distributions for various values of $\beta$. *Top:* Linear-linear representation; *Bottom:* Log-log representation. The power-law index $n = 1 + \alpha$ satisfies $1 < n < 3$ since necessarily $0 < \alpha < 2$. Indeed, $\alpha \geq 2$ corresponds to *finite* variance, and therefore, in the sense of the Central Limit Theorem, it yields a Gaussian attractor, and *not* a power-law-like Lévy distribution.
Figure 4. Typical $q$-exponential distributions (with $\eta > 0$). The power-law index $n = \frac{\eta}{q-1}$ satisfies $1 < n < \infty$ since $q$ must satisfy $q < 1 + \eta$ in order to be normalizable, and must satisfy $q > 1$ in order to asymptotically be a power-law.

### 3.2 $q$-exponential distributions

Let us define the $q$-exponential family of distributions as follows:

\[
(x; q, \eta, \beta_q, u_q) = \frac{e^{-\beta_q |x-u_q|^\eta}}{\int_{-\infty}^{\infty} dy e^{-\beta_q |y-u_q|^\eta}} = \frac{[1 - (1 - q)\beta_q |x-u_q|^\eta]^{\frac{1}{\eta}}}{\int_{-\infty}^{\infty} dy [1 - (1 - q)\beta_q |y-u_q|^\eta]^{\frac{1}{\eta}}}
\]  

(13)

$(q \geq 1, \eta > 0, \beta_q > 0, \infty < u_q < \infty)$.

The reason why we exclude here $q < 1$ comes from the fact that we are focusing on the cases where there are tails up to infinity (in fact, fat tails if $q > 1$). If $x$ has the physical dimensions of an [energy]$^{1/\eta}$, an effective temperature $T_q$ can be defined through $\beta_q = \frac{1}{kT_q}$. If $x$ has the physical dimensions of a [time], a characteristic relaxation time $\tau_q$ can be defined through $\beta_q = \frac{1}{(\tau_q)^{1/\eta}}$ (see [37] for such an example). In the $x \to \infty$ limit, we straightforwardly verify that, for $q > 1$,

\[
p(x; q, \eta, \beta_q, u_q) \propto \frac{1}{x^{n+1}} \quad (q > 1; \, \eta > 0; \, \forall \beta_q; \, \forall u_q).
\]  

(14)

The $q$-exponential distributions can be shown to extremize the entropy $S_q$ under appropriate constraints (basically the mean value of $|x|^\eta$, in addition to the norm constraint). For them to be normalizable, the following bound emerges: $\frac{\eta}{q-1} > 1$, hence $q < 1 + \eta$. The case $\eta = 1$ corresponds to the
current $q$-generalization, within nonextensive statistical mechanics, of the BG factor, and $q < 2$ (the mean value of $x$ is finite if $q < 3/2$, and diverges for $3/2 \leq q < 2$). The case $\eta = 2$ corresponds to the $q$-Gaussians, i.e., the current $q$-generalization of Gaussians, and $q < 3$ (the variance is finite for $q < 5/3$, and diverges for $5/3 \leq q < 3$). In general, the mean value of $|x|^\eta$ is finite for $q < \frac{2\eta+1}{\eta}$, and diverges for $\frac{2\eta+1}{\eta} \leq q < 1 + \eta$. Typical examples of $q$-exponentials are shown in Fig. 4.

We may extend the $q$-exponential class quite naturally by introducing, in definition (13), the analog of a density of states $g(x)$, as usually done in condensed matter physics and elsewhere (see in [70] one such example in economics, for the distribution of stock-market volumes). The more general
\[ p(x; q, \eta, \beta_q, u_q) = \frac{g(x) e^{-\beta_q |x-u_q|^{q}}}{\int_{\infty}^{\infty} dy \, g(y) e^{-\beta_q |y-u_q|^{q}}} \]
\[ = \frac{g(x) [1 - (1 - q)\beta_q |x-u_q|^{q}]^{-\frac{1}{q}}}{\int_{\infty}^{\infty} dy \, g(y) [1 - (1 - q)\beta_q |y-u_q|^{q}]^{-\frac{1}{q}}} \] (15)

where \( g(x) \) is a well behaved function which, in the \( x \to 0 \) limit, asymptotically satisfies \( g(x) \propto |x|^\lambda \) with \( \lambda \in \mathbb{R} \). Obviously the case \( g(x) = constant \) (hence \( \lambda = 0 \)) recovers definition (13). Examples with \( g(x) \neq constant \) are very frequent in the literature. For example, the functions used in [66] belong to this class.

### 3.3 Comparison

Lévy and \( q \)-exponential distributions are definitively different, with only two exceptions, namely Gaussians (which correspond to \( \alpha = 2, \forall \beta \), within the Lévy class, and to \( q = 1 \) and \( \eta = 2 \) within the \( q \)-exponential class), and the Cauchy-Lorentz distributions (which correspond to \( \alpha = 1 \) and \( \beta = 0 \) within the Lévy class, and to \( q = 2 \) and \( \eta = 2 \) within the \( q \)-exponential class).

Generic Lévy \((\alpha < 2)\) and \( q \)-exponential \((q > 1)\) distributions asymptotically decay, as infinitely many other distributions, as power laws. Typical comparative examples are shown in Fig. 5 by imposing, by construction, that both Lévy and \( q \)-exponential distributions exhibit the same exponent in the respective asymptotic power-laws. The only region where this can be done corresponds to \( 1 < n < 3 \). Now, the empiric curves of hadronic transverse momenta distributions in LHC experiments exhibit \( n \approx 7 \), which is definitively out of that interval. Therefore, there is no way for those empiric distributions to belong to the Lévy class. In contrast, since the distributions of the \( q \)-exponential class admit any value of \( n \) above unity, this class becomes, as the quality of the fittings strongly suggest, a very plausible candidate. Consequently, expressions such as Lévy-Tsallis distribution (or function) that, by inadvertence, have repeatedly appeared in the LHC literature (for example, see [60, 64, 65]) should be avoided. Indeed, Lévy distributions do exist, and \( q \)-exponential distributions (sometimes very kindly referred to as “Tsallis distributions”) do exist, but Lévy–\( q \)-exponential distributions (or “Lévy-Tsallis distributions”) have never been mathematically defined.

### 4 Conclusions

We have briefly shown why Boltzmann-Gibbs entropy, and its associated statistical mechanics, need, in a wide class of complex systems, to be generalized, and how this can be done, namely through non-additive entropies such as \( S_q \) and \( S_\delta \). Entropic indices such as \( q \) and \( \delta \) (and other related indices) are to be calculated from first principles, namely from mechanics (classical, quantum, relativistic, field theory, quantum chromodynamics). This is a highly nontrivial mathematical task for most systems, but, nevertheless, it has been successfully accomplished occasionally (e.g., [72]). However, very frequently mathematical intractability forces one to also include some level of fitting (with experimental, observational or computational data) by using the appropriate, analytically determined, functional forms that the theory mandates, such as \( q \)-exponentials and \( q \)-Gaussians. This path has shown to be fruitfully applicable in a variety of natural, artificial and social systems.
In particular, in what concerns high-energy collisional experiments such as those at LHC/CERN (ALICE, CMS, ATLAS, LHCb) and at RHIC/Brookhaven (STAR, PHENIX), interesting results have been found in quantities such as the distributions of hadronic jet transverse momenta and rapidity, among others. For example, the results so far strongly suggest that the index $q$ for transverse momenta distributions slowly increases from 1 (i.e., BG thermal equilibrium) to about say 1.2 while the collision energy increases from low values to values above those presently achieved at LHC. For say 7 $\text{TeV}$, a typical value for $q$ is always close to 1.15 (assuming that $\eta = 1$), for many different hadrons. This fact appears to indicate that the system that evolves during hadronization lives longstandingly in a nonergodic quasi-stationary state, possibly due to memory effects or to long-range interactions (strong confinement, for instance), or both. Studies developing in parallel nonextensive statistical mechanics and quantum chromodynamics (or a similar theory) might be very enlightening for understanding the intimate nature of phenomena occurring at that very tiny scale. One pioneering such study was advanced more than one decade ago in [73].

Acknowledgments

One of us (CT) is indebted to F. Antinori, L.J.L. Cirto, E.M.F. Curado, Y. Foka, M. Jauregui, F.D. Nobre and U. Wiedemann for fruitful discussions. We have benefitted from partial financial support from CNPq, FAPERJ and CAPES (Brazilian agencies). In particular, ZGA is supported by a PDJ Fellowship from CNPq.

References


[69] J. P. Nolan, Stable Distributions - Models for Heavy Tailed Data, Birkhauser, Boston (2013). In progress, Chapter 1 online at academic2.american.edu/~jpnolan.


