Non-Extensive Entropic Distance Based on Diffusion: Restrictions on Parameters in Entropy Formulae

Tamás Sándor Biró [1],* and Zsolt Schram [2]

[1]   Wigner Research Center for Physics, Konkoly-Thege M. 29-33, Budapest H-1121, Hungary
[2]   Department of Theoretical Physics, University of Debrecen, Bem tér 18/B, Debrecen H-4026, Hungary;
     schram@phys.unideb.hu
*   Correspondence: Biro.Tamas@wigner.mta.hu; Tel.: +36-1-392-2222-3388

Abstract: Based on a diffusion-like master equation we propose a formula using the Bregman divergence for measuring entropic distance in terms of different non-extensive entropy expressions. We obtain the non-extensivity parameter range for a universal approach to the stationary distribution by simple diffusive dynamics for the Tsallis and the Kaniadakis entropies, for the Hanel–Thurner generalization, and finally for a recently suggested log-log type entropy formula which belongs to diverging variance in the inverse temperature superstatistics.

---

1. Introduction

Over the last decades, there have been several suggestions for generalizations of the Boltzmann–Gibbs–Shannon (BGS) entropy formula [1–6]. Most formulas can be grouped into categories either by their mathematical form (trace form or a function of the trace form) [7], or by the scaling properties for large systems, usually providing large entropies, S, even if not necessarily proportional to the logarithm of the number of states, $\ln W$ [8–11]. Such entropy formulas are often presented based on axioms, like the Khinchin or Shannon axioms for the most known logarithmic entropy formula due to Boltzmann and Gibbs. Non-extensive entropy then abandons the additivity axiom and replaces it with a more general assumption. For the group entropy, the assumption of the associativity of the entropy composition rule is used [7]. Another way to obtain a generalized entropy formula starts with physical properties, like the universality of the thermal equilibrium between two systems as described by the zeroth law of thermodynamics [12,13], and by this it also includes the associativity assumption [14]. It is also noteworthy that starting with a general, non-associative composition prescription one arrives asymptotically at an associative one—only by repeating it in small steps and reconstructing the effective composition formula in the continuous scaling limit [15].

Deformations of the original Boltzmann–Gibbs–Shannon entropy formula can also be achieved by tracing it back from the canonical distribution, exponential in the subsystem energy, when parameters of this distribution are treated as stochastic ones, or otherwise randomized, and hence integrated over another—the superstatistical—distribution [16–18]. Instead of treating the temperature or its inverse as a fluctuating quantity, one may also consider the number of degrees of freedom, and hence the dimensionality of phase space, as a fluctuating quantity. In [19–23] it has been shown that simple fluctuation patterns of the particle number, N, influencing the phase space volume Ω occupied by an ideal gas, can lead both to an exponential distribution of $\exp(-\omega/T)$ by a Poisson N distribution,

or to cut power-law like Tsallis–Pareto distribution by a negative binomial N distribution. In general, a trace form entropy,

$$K(S) := \sum_{i=1}^{W} p_i \, K(-\ln p_i) \tag{1}$$

emerges with $K(S)$ satisfying a second order differential equation, the additivity restoration condition (ARC) [24]. Connecting to a more traditional notation, the K function is related to the deformed logarithm function as

$$\ln_{\mathrm{def}}(x) = -K(-\ln x) \tag{2}$$

This form includes the one parameter deformed logarithms as well as the generalizations with more parameters [12,13,25–27].

In the present paper we investigate whether such entropy formulas define an entropic distance between two probability distributions, which has the following useful properties:

1.　it is positive for any two different distributions;
2.　it is zero for comparing any distribution with itself;
3.　it is symmetric.

A symmetric distance measure is most naturally formulated as a sum of relative divergences, $\rho(P,Q) = \sigma(P|Q) + \sigma(Q|P)$. For deformed logarithms two basic forms of suggestions occur in the literature, one based on the ratio P_n/Q_n and one based on the difference in their logarithm, *cf.* refs. [28–31]. Furthermore we are interested in a definition which describes a distance that is always shrinking between a dynamically evolving distribution and the one belonging to maximal entropy under a spontaneous dynamics.

In the following we show this behavior for the traditional logarithmic entropy formula and a generally state-dependent nearest neighbour master equation (defined below), and then propose a generalization of the symmetrized entropic distance measure based on the deformed logarithm function. We conclude that some proposals encertain a shrinking of the entropic distance during an approach to the stationary distribution only for a restricted range of the non-extensivity parameter(s) used in the entropy formula.

2. Probability Distributions

The Poisson distribution is ubiquitous in statistical phenomena, where the discreteness of the basic variable, n, and its positivity are natural constraints. This distribution,

$$P_n^{\mathrm{POI}} = \frac{a^n}{n!} \, \mathrm{e}^{-a} \tag{3}$$

is parameterized by a single parameter, coinciding with the mean value of the variable $n: a = \langle n \rangle$. The logarithm of its characteristic function,

$$\ln \chi^{\mathrm{POI}} = \ln \sum_n \mathrm{e}^{-nx} \, P_n^{\mathrm{POI}} = - \langle n \rangle \, (1 - e^{-x}) \tag{4}$$

generates all central moments of the distribution. Notably, due to $\delta n^2 = \langle n \rangle$ the Poisson distribution narrows for growing mean value of n.

This distribution plays a central role in different areas of physics: it occurs as the statistics of spontaneous radioactive decays during given time intervals, as the number of photons in a Glauber coherent state, or as the number distribution of randomly produced hadrons in high energy experiments when uniformly filling the available phase space. In the latter case, a given total energy, E, is distributed among n newly made particles in experiment. The statistical weight to find a single particle with individual energy ω among the products is roughly proportional to the relative phase

space described by a ratio of corresponding n-dimensional spheres. Its average over the Poisson distribution is given by the following remarkable formula

$$\left\langle \frac{(E-\omega)^n}{E^n} \right\rangle = \sum_n \left(1 - \frac{\omega}{E}\right)^n P_n^{\mathrm{POI}} = e^{-\langle n \rangle \omega / E} \tag{5}$$

Several authors use the power $n - 1$ instead of n in this formula. The difference traces back to different concepts on the hyperspheric phase space volume *vs.* surface. Since for large n it does not matter, we do not go into details here.

This result allows for a Gibbsean interpretation of this statistical factor with the kinetic temperature $T = E / \langle n \rangle$.

Beyond the Poissonian also sub- and super-Poissonian bosonic states occur with the corresponding Bernoulli or negative binomial multiplicity distributions in experiments. The above ideal phase space ratio averages in those cases to a statistical factor generalizing the exponential to a cut power-law form, also known as a canonical distribution to the Tsallis entropy, e.g.,

$$\left\langle \frac{(E-\omega)^n}{E^n} \right\rangle = \sum_n \left(1 - \frac{\omega}{E}\right)^n P_n^{\mathrm{NBD}} = \left(1 + \frac{1}{k+1} \langle n \rangle \omega / E\right)^{-k-1} \tag{6}$$

for the negative binomial distribution

$$P_n^{\mathrm{NBD}} = \binom{k+n}{n} p^n (1+p)^{-n-k-1} \tag{7}$$

with $\langle n \rangle = p(k+1)$

The above Tsallis–Pareto distribution, Equation (6) can also be obtained as a canonical distribution to the Tsallis entropy formula [3–5],

$$K(S) = \frac{1}{1-q} \sum_i \left(p_i^q - p_i\right) \tag{8}$$

with $q = 1 + 1/(k+1)$ and $T = E / \langle n \rangle$.

These and similar experiences suggest that different entropy formulas can be used as a basis for defining the entropic distance. Earlier attempts mostly defined a divergence formula based on the deformed logarithm, $\ln_{\mathrm{def}}$, of the ratio of the two respective probabilities [31,32]. An alternative approach—which we propose to follow—uses the difference of the deformed logarithms as a definition, it is based on a Bregman type divergence. While for the traditional logarithm function these alternative definitions coincide (due to $\ln(x/y) = \ln x - \ln y$), they do differ appreciably for the deformed logarithm functions in general use.

3. Master Equation

Let us denote the probability of having exactly n quanta in the observed system (part of phase space) at the time t generally by $P_n(t)$. This quantity may depend on further parameters, this is suppressed in the notation by now. However, in any blind choice case it is proportional to the quantity $1/W_n$, to the ratio of one to the number of possible ways to realize the state with n quanta.

A wide range of phenomena can be described by a simplified dynamics, assuming that only one quantum can be exchanged in a time instant. In this case, the evolution of $P_n(t)$ and W_n depends only on the state probabilities of having one more or one less quantum. We consider here only the linearized version of this dynamics, the so called *master equation*:

$$\dot{P}_n = (\lambda P)_{n+1} - (\lambda P)_n + (\mu P)_{n-1} - (\mu P)_n \tag{9}$$

Here, λ_n is the transition (decay) rate from a state with n quanta to the state $n - 1$, μ_n is the corresponding reverse (growth) rate from n to $n + 1$. In this case the occurrence of state n in a huge parallel ensemble (Gibbs ensemble) of systems is fed by both the $(n + 1) \to n$ and $(n - 1) \to n$ processes and it is diminished by the reverse processes. It is of special interest to investigate processes when λ and μ are related by symmetry principles, like time reversal invariance or subsystem—reservoir homogeneity.

We shall quote the general detailed balance solution of Equation (9), $P_n^{eq} = Q_n$, as the *equilibrium distribution*. Since the equation is homogeneous and linear in the Q_n-s the overall normalization is not fixed by it. Being the equilibrium distribution, all $\dot{Q}_n$-s should be zero. That condition is fulfilled only if

$$(\lambda Q)_{n+1} = (\mu Q)_n \tag{10}$$

from which it follows that also

$$(\lambda Q)_n = (\mu Q)_{n-1} \tag{11}$$

holds, annullating all evolution. Based on this observation, the detailed balance distribution satisfies

$$Q_n = \frac{\prod\limits_{i=0}^{n-1} \mu_i}{\prod\limits_{j=1}^{n} \lambda_j} Q_0 \tag{12}$$

where Q_0 can be obtained from the normalization condition

$$\sum_n Q_n = 1 \tag{13}$$

This is kept during the evolution, since $\sum_n \dot{P}_n = 0$ upon Equation (9).

4. Entropic Distance

Based on the Boltzmann-Gibbs entropy,

$$S = -\sum_n P_n \ln P_n \tag{14}$$

a distance measure from a reference distribution, Q_n is based on the quantity

$$\sigma(P|Q) = \sum_n P_n \ln \frac{P_n}{Q_n} \tag{15}$$

By this definition $\sigma(Q|Q) = 0$, and it gives a positive Kullback–Leibler divergence [33–35] between the normalized distributions P_n and Q_n. This can be easily proved by using the Jensen inequality, $\prod A_i^{P_i} \le \sum P_i A_i$ applied to $A_i = Q_i / P_i$. The symmetric combination, based on a sum of Bregman type divergences [28–31],

$$\rho(P, Q) = \sigma(P|Q) + \sigma(Q|P) = \sum_n (P_n - Q_n) \ln \frac{P_n}{Q_n} \tag{16}$$

is non-negative then *term by term*. For a fixed, time-independent Q_n the master Equation (9) causes $\sigma(P|Q)$ to decrease:

$$\dot{\sigma}(P|Q) = \sum_n \left[(\lambda P)_{n+1} - (\lambda P)_n + (\mu P)_{n-1} - (\mu P)_n \right] \ln \frac{P_n}{Q_n} \tag{17}$$

The sum can be re-arranged by summation index re-definition to

$$\dot{\sigma}(P|Q) = \sum_n [(\lambda P)_{n+1} - (\mu P)_n] \ln \frac{P_n Q_{n+1}}{Q_n P_{n+1}} \tag{18}$$

This expression is non-positive for arbitrary $P_n(t)$ values only if

$$\lambda_{n+1} Q_{n+1} = \mu_n Q_n \tag{19}$$

Namely, one arrives solely in this case at the $(1 - x) \ln x$ type expression under the sum. But this means that the reference distribution, to which any initial distribution approaches if evolving according to the master equation, is the detailed balance distribution! It can be shown that by the virtue of the master equation also the distance $\sigma(Q|P)$ is decreasing term by term. The symmetrically summed Kullback–Leibler distance to the detailed balance distribution is also reduced: $\dot{\rho}(P, Q) \leq 0$, due to

$$\dot{\sigma}(Q|P) = \sum_n (P_n - Q_n) \frac{\dot{P}_n}{P_n} = \sum_n -(\lambda P)_{n+1} \frac{Q_n}{P_n} (1 - x)^2 \leq 0 \tag{20}$$

(Here $x = (\mu P)_n / (\lambda P)_{n+1}$).

In principle different distance measures (based on deformed entropy formulas) can also be used to investigate this property of a linear master equation. If one considers an entropy defined by a deformed logarithm function,

$$S = - \sum_n P_n \ln_{\text{def}} P_n \tag{21}$$

then the Kullback–Leibler divergence and the definition of the distance between probability distributions can be modified accordingly as

$$\rho(P, Q) = \sum_n (P_n - Q_n) \left(\ln_{\text{def}} P_n - \ln_{\text{def}} Q_n \right) = \sum_n R_n \tag{22}$$

This measure is zero only if $P_n = Q_n$ for all n; otherwise it is positive if $\ln_{\text{def}} P$ is a strictly increasing function, *i.e.*, $\ln'_{\text{def}} P > 0$ on the whole interval $(0, 1)$.

The condition of the convergence to the $P_n = Q_n$ detailed balance solution then reads as

$$\dot{\rho} = \sum_n \dot{P}_n R'_n \leq 0 \tag{23}$$

with

$$R'_n = \frac{\partial R_n}{\partial P_n} = \ln_{\text{def}} P_n - \ln_{\text{def}} Q_n + (P_n - Q_n) \ln'_{\text{def}} P_n \tag{24}$$

On the other hand, from the master equation one obtains

$$\dot{P}_n = \Gamma_n (x_{n+1} - x_n) - \Gamma_{n-1} (x_n - x_{n-1}) \tag{25}$$

with $x_n = P_n / Q_n$ and $\Gamma_n = (\lambda Q)_{n+1} = (\mu Q)_n$. Substituting Equation (25) into Equation (23) and re-arranging the summation indices finally we arrive at the requirement

$$\dot{\rho} = \sum_n \Gamma_n (x_{n+1} - x_n) \left[R'_n - R'_{n+1} \right] \leq 0 \tag{26}$$

This is satisfied if R'_n is strictly increasing, which means that $R''_n > 0$ for all possible P_n and Q_n values; $\dot{\rho} = 0$ being realized only if $x_{n+1} = x_n = 1$ (the last equality is due to the normalization

$\sum_n P_n = \sum_n Q_n x_n = 1$ while $\sum_n Q_n = 1$). The condition for approaching the detailed balance distribution then finally reads as

$$\frac{dR'_n}{dx_n} = Q_n \left[2 \ln'_{def} P_n + (P_n - Q_n) \ln''_{def} P_n \right] = Q_n R''_n > 0 \tag{27}$$

It is clear that this is a more detailed restriction than just the concavity of the entropy formula. The Q-independent part of the expression in the bracket equals $(P \ln_{def} P)''$, whose positivity is the well-known *concavity* requirement for the entropy formula. In cases when also $\ln''_{def} P < 0$, this concaivity suffices for the uniform approach to the stationary distribution. In the opposite case we arrive at the nontrivial constraint $(P \ln_{def} P)'' > \ln''_{def} P$ at $Q_n = 1$.

Let us now test whether some familiar suggestions for deformed logarithms satisfy this constraint. The Tsallis logarithm is defined as

$$\ln_{def} P_n = \frac{P_n^{q-1} - 1}{q - 1} \tag{28}$$

It results in $\ln'_{def} P_n = P_n^{q-2} > 0$ and due to Equation (27) the final constraint is

$$Q_n P_n^{q-3} \left[q P_n + (2 - q) Q_n \right] > 0 \tag{29}$$

This is fulfilled for arbitrary (P, Q) probability distribution pairs if $0 \le q \le 2$; otherwise it might be violated. The classical logarithm is reconstructed when $q = 1$.

The kappa-exponential, promoted by Kaniadakis [36], belongs to the deformed logarithm

$$\ln_{def} P_n = \frac{P_n^\kappa - P_n^{-\kappa}}{2\kappa} \tag{30}$$

In this case one obtains

$$(1 + \kappa) \left[P_n^{1+\kappa} + Q_n P_n^{-\kappa} \right] + (1 - \kappa) \left[P_n^{1-\kappa} + Q_n P_n^\kappa \right] > 0 \tag{31}$$

as the condition for the correct evolution towards the detailed balance solution by the linear master equation. This is fulfilled universally (*i.e.*, for arbitrary P_n and Q_n in the physical interval of $[0, 1]$) as long as $\kappa^2 \le 1$. One obtains this result as follows. Both square bracket expressions in Equation (31) are non-negative. Their sum is also non-negative if $-1 \le \kappa \le +1$. Otherwise there are P_n values near to zero for which one of the summands, either due to $(1 - \kappa) < 0$ or due to $(1 + \kappa) < 0$, diverge to negative infinity, spoiling the inequality in this way. The classical logarithm is reconstructed for $\kappa = 0$ and in order to have the same power-law tail in canonical distributions $\kappa = q - 1$. The condition for universal evolution translates to $0 \le q \le 2$.

By now there is a much bigger variety of suggested entropy formulas, many of them having the trace form using a deformed logarithm function [37–39]. Here we analyse one of them, the (c, d) entropy suggested by Hanel and Thurner [8], since it includes a number of different entropy formulas suggested earlier as particular cases. It is given as

$$S_{c,d} = \sum_{i=1}^{W} \left[A \, \Gamma(d + 1, 1 - c \ln p_i) - B p_i \right] \tag{32}$$

with

$$\Gamma(d + 1, z) = \int_z^\infty t^d e^{-t} dt \tag{33}$$

being the incomplete gamma function. The coefficients A and B can be determined by considering the equiprobability case, $p_i = 1/W$, resulting in $S_{c,d}^{\text{eq.prob.}} = K(\ln W)$, and applying the conditions

$\ln_{\text{def}} 1 = -K(0) = 0$ and $\ln'_{\text{def}} 1 = K'(0) = 1$. By this procedure one obtains $A = 1/(\Gamma_1 + c\Gamma'_1)$ and $B = A\Gamma_1$, with $\Gamma_1 = \Gamma(d+1,1)$ and $\Gamma'_1 = -1/e$.

The corresponding deformed logarithm of the probability, based on the form Equation (21), is

$$\ln_{\text{def}} P_n = -\frac{A}{P_n} \Gamma(d+1, 1 - c \ln P_n) + B \tag{34}$$

Here, for instance, the $c = q$, $d = 0$ choice gives the Tsallis logarithm, Equation (28) with $A = e/(1-c)$ and $B = 1/(1-c)$.

The first and second derivative of the deformed logarithm are given as

$$\begin{aligned} \ln'_{\text{def}} P_n &= \frac{A}{P_n^2} \left(\Gamma + c\Gamma'\right), \\ \ln''_{\text{def}} P_n &= -\frac{A}{P_n^3} \left(2\Gamma + 3c\Gamma' + c^2\Gamma''\right) \end{aligned} \tag{35}$$

with

$$\Gamma' = -\frac{P_n^c}{e} \left(1 - c \ln P_n\right)^d \tag{36}$$

and

$$\Gamma'' = \frac{P_n^c}{e} \left(1 - c \ln P_n\right)^{d-1} \left[1 - d - c \ln P_n\right] \tag{37}$$

Here we used the abbreviations Γ, Γ' and Γ'' without explicitly writing out the arguments $d+1$ and $1 - c \ln P_n$.

All this leads to the following criterion for the term by term approach to the stationary distribution Q_n by a general P_n via the diffusion like master equation:

$$A \left[Q_n \left(2\Gamma + 3c\Gamma' + c^2\Gamma''\right) - P_n \left(c\Gamma' + c^2\Gamma''\right) \right] > 0 \tag{38}$$

Suppose now that $A > 0$. This requires $e\Gamma(d+1,1) > c$. (In the above $d = 0$ example that translates to $c < 1$.) Furthermore, considering the extreme case $Q_n = 0$ (equivalent to the concavity condition) we deal with

$$\left(c\Gamma' + c^2\Gamma''\right) = \frac{c}{e} P_n^c (1 - c \ln P_n)^{d-1} \left[(c-1)(1 - c \ln P_n) - cd\right] < 0 \tag{39}$$

This is fulfilled for $0 < c < 1$ and $d > 1 - 1/c$. Once this is so, observing that the lhs expression of the inequality Equation (39) is linear in Q_n, the condition has to be checked only at the other endpoint of the Q_n interval. At $Q_n = 1$ we have:

$$\left(2\Gamma + 3c\Gamma' + c^2\Gamma''\right) > P_n \left(c\Gamma' + c^2\Gamma''\right) \tag{40}$$

Although this criterion looks involved, we note, that at $P_n = 1$ it simplifies to $1/A > 0$, *i.e.*, $\Gamma(d+1,1) > c/e$ as supposed above. Similar analysis can be done for the opposite, $A < 0$ case.

Finally, the recently suggested doubly logarithmic entropy formula, designed for extreme large fluctuations in a reservoir by Biro *et al.* [23,24], considers

$$\ln_{\text{def}} P_n = -\ln\left(1 - \ln P_n\right) \tag{41}$$

as the deformed logarithm, leading to

$$\frac{1}{P_n^2 (1 - \ln P_n)^2} \left[2P_n - (P_n + Q_n) \ln P_n\right] > 0 \tag{42}$$

for the evolution condition $\dot{\rho} < 0$. This is fulfilled for any pair of distributions, since $\ln P_n < 0$ is always true.

5. Conclusions

By now the use of non-extensive entropy formulas is clearly numerous. A review of all areas of physics where a power-law tailed canonical distribution occurs can be found in [5]. Furthermore, applications to non-Gaussian velocity distributions in dusty plasmas can be found in [40,41]. In this paper we have concentrated on a mathematical problem in the background testing generalized entropy formulas with respect to whether they serve as a basis for a well-behaving distance measure which describes an approach to the stationary distribution of state-dependent diffusion-like master equations. We have found that, in some parameter range of the suggested modern entropy formulas, the distance measure, based on the difference of deformed logarithms of the respective distributions, behaves well as shrinking uniformly, term by term during the time evolution towards the stationary distribution of general, state-dependent diffusion-like master equations. In particular, while the Tsallis entropy and other generalized entropy formulas (involving further deformed logarithm functions) guarantee the above term by term approach only in a restricted range of parameters, the extreme case belonging to a diverging variance in the temperature superstatistics, with a log-log form entropy formula, behaves universally well in this respect.

Author Contributions: Tamás Sándor Biró and Zsolt Schram performed and controled the calculations and prepared the manuscript alternatingly. Both authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rényi, A. On Measures of Entropy and Information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
2. Rényi, A. *Probability Theory*; North Holland: Amsterdam, The Netherlands, 1970.
3. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
4. Tsallis, C. Nonextensive statistics: Theoretical, experimental and computational evidences and connections. *Braz. J. Phys.* **1999**, *29*, doi:10.1590/S0103-97331999000100002 .
5. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics*; Springer: New York, NY, USA, 2009.
6. Almeida, M.P. Generalized entropies from first principles. *Physica A* **2001**, *300*, 424–432.
7. Tempesta, P. Group entropies, correlation laws and zeta functions. *Phys. Rev. E* **2015**, *84*, 021121.
8. Hanel, R.; Thurner, S. A comprehensive classification of complex statistical systems and axiomatic derivation of their entropy and distribution functions. *Europhys. Lett.* **2011**, *93*, 20006.
9. Hanel, R.; Thurner, S. When do generalised entropies apply? How phase space volume determines entropy. *Europhys. Lett.* **2011**, *96*, 50003.
10. Hanel, R.; Thurner, S.; Gell-Mann, M. How multiplicity determines entropy and derivation of the maximum entropy principle for complex systems. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6905–6910.
11. Hanel, R.; Thurner, S. Generalized (c, d) entropies and aging random walks. *Entropy* **2013**, *15*, 5324–5337.
12. Kaniadakis, G. Maximum entropy principle and power-law tailed distributions. *Eur. Phys. J. B* **2009**, *70*, 3–13.
13. Kaniadakis, G. Theoretical Foundations and Mathematical Formalism of the Power-Law Tailed Statistical Distributions. *Entropy* **2013**, *15*, 3983–4010.
14. Biró, T.S.; Ván, P. Zeroth law compatibility of non-additive thermodynamics. *Phys. Rev. E* **2011**, *83*, 061187.
15. Biró, T.S. Abstract composition rule for relativistic kinetic theory in the thermodynamical limit. *Europhys. Lett.* **2008**, *84*, 56003.
16. Beck, C. Dynamical foundations of nonextensive statistical mechanics. *Phys. Rev. Lett.* **2001**, *87*, 180601.

17. Beck, C.; Cohen, E.G.D. Superstatistics. *Physica A* **2003**, *322*, 267–275.

18. Abe, S.; Beck, C.; Cohen, E.G.D. Superstatistics, thermodynamics and fluctuations. *Phys. Rev. E* **2007**, *76*, 031102.

19. Biró, T.S. *Is There a Temperature?* Springer: New York, NY, USA, 2011.

20. Ván, P.; Barnaföldi, G.G.; Biró, T.S.; Ürmössy, K. Nonadditive thermostatistics and thermodynamics. *J. Phys. Conf. Ser.* **2012**, *394*, 012002.

21. Biró, T.S. Ideal gas provides q-entropy. *Physica A* **2013**, *392*, 3132–3139.

22. Biró, T.S.; Ván, P.; Barnaföldi, G.G. Quark-gluon plasma connected to finite heat bath. *Eur. phys. J. A* **2013**, *49*, doi:10.1140/epja/i2013-13110-0.

23. Biró, T.S.; Barnaföldi, G.G.; Ván, P. New Entropy Formula with Fluctuating Reservoir. *Physica A* **2015**, *417*, 215–220.

24. Biró, T.S.; Ván, P.; Barnaföldi, G.G.; Ürmössy, K. Statistical Power Law due to Reservoir Fluctuations and the Universal Thermostat Independence Principle. *Entropy* **2014**, *16*, 6497–6514.

25. Abe, S. A note on the q-deformation-theoretic aspect of the generalized entropies in nonextensive physics. *Phys. Lett. A* **1997**, *224*, 326–330.

26. Kaniadakis, G. Relativistic entropy and related Boltzmann kinetics. *Eur. Phys. J. A* **2009**, *40*, 275–287.

27. Ubriaco, M.R. Entropies based on factional calculus. *Phys. Lett. A* **2009**, *373*, 2516–2519.

28. Naudts, J. Deformed exponentials and logarithms in generalized tehrmostatistics. *Physica A* **2002**, *316*, 323–334.

29. Naudts, J. Continuity of a class of entropies and relative entropies. *Rev. Math. Phys.* **2004**, *16*, 809–822.

30. Naudts, J. Generalized thermostatistics based on deformed exponential and logarithmic functions. *Physica A* **2004**, *340*, 32–40.

31. Naudts, J. Generalised Exponential Families and Associated Entropy Functions. *Entropy* **2008**, *10*, 131–149.

32. Tsallis, C. Generalized entropy-based criterion for consistent testing. *Phys. Rev. E* **1998**, *58*, doi:10.1103/PhysRevE.58.1442.

33. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.

34. Kullback, S. Letter to the Editor: The Kullback–Leibler distance. *Am. Stat.* **1987**, *41*, 340–341.

35. Van Erven, T.; Harremoës, P. Rényi Divergence and Kullback–Leibler Divergence. *IEEE Trans. Inf. Theory* **2014**, *60*, 3797–3820.

36. Kaniadakis, G. Non-linear kinetics underlying generalized statistics. *Physica A* **2001**, *296*, 405–425.

37. Borges, E.P.; Raditi, I. A family of nonextensive entropies. *Phys. Lett. A* **1998**, *246*, 399–402.

38. Schwämmle, V.; Tsallis, C. Two-parameter generalization of the logarithm and exponenetial functions and the Boltzmann–Gibbs–Shannon entropy. *J. Math. Phys.* **2007**, *48*, 113301.

39. Tsallis, C.; Cirto, L.J.L. Black hole thermodynamical entropy. *Eur. Phys. J. C* **2013**, *73*, doi:10.1140/epjc/s10052-013-2487-6.

40. Liu, B.; Goree, J. Superdiffusion and non-Gaussian statistics in a driven-dissipative 2D dusty plasma. *Phys. Rev. Lett.* **2008**, *100*, 055033.

41. Amour, R.; Tribecke, M. Variable charge dust acoustic solitary waves in a dusty plasma with non-extensive electron velocity distribution. *Phys. Plasmas* **2010**, *17*, 063702.