Postprint of: Czachor M., Naudts J., Thermostatistics based on Kolmogorov–Nagumo averages: unifying framework for extensive and nonextensive generalizations, Physics Letters A, Vol. 298, Iss. 5–6 (2002), pp. 369-374, DOI: 10.1016/S0375-9601(02)00540-6

Thermostatistics based on Kolmogorov-Nagumo averages: Unifying framework for extensive and nonextensive generalizations

Jan Naudts[1] and Marek Czachor[1,2]

[1] *Departement Natuurkunde, Universiteit Antwerpen UIA, Universiteitsplein 1, B2610 Antwerpen, Belgium*

[2] *Katedra Fizyki Teoretycznej i Metod Matematycznych, Politechnika Gdańska, 80-952 Gdańsk, Poland*

E-mail: Jan.Naudts@ua.ac.be and mczachor@pg.gda.pl

We show that extensive thermostatistics based on Rényi entropy and Kolmogorov-Nagumo averages can be expressed in terms of Tsallis non-extensive thermostatistics. We use this correspondence to generalize thermostatistics to a large class of Kolmogorov-Nagumo means and suitably adapted definitions of entropy.

PACS numbers: 05.20.Gg, 05.70.Ce

Generalized averages of the form

$$\langle x\rangle_\phi = \phi^{-1}\Big(\sum_k p_k\phi(x_k)\Big), \tag{1}$$

where ϕ is an arbitrary continuous and strictly monotonic function, were introduced into statistics by Kolmogorov [1] and Nagumo [2], and further generalized by de Finetti [3], Jessen [4], Kitagawa [5], Aczél [6] and many others. Their first applications in information theory can be found in the seminal papers by Rényi [7,8] who employed them to define a one-parameter family of measures of information (α-entropies)

$$I_\alpha = \varphi_\alpha^{-1}\Big(\sum_k p_k\varphi_\alpha(\log_b\frac{1}{p_k})\Big) = \frac{1}{1-\alpha}\log_b\Big(\sum_k p_k^\alpha\Big). \tag{2}$$

The Kolmogorov-Nagumo (KN) function is here $\varphi_\alpha(x) = b^{(1-\alpha)x}$, a choice motivated by a theorem [9] stating that only affine or exponential ϕ satisfy

$$\langle x + C\rangle_\phi = \langle x\rangle_\phi + C \tag{3}$$

where C is a constant. Random variable

$$I_k = -\log_b p_k \tag{4}$$

represents an amount of information received by learning that an event of probability p_k took place [10,11]; b specifies units of information ($b = 2$ corresponds to bits; below we use $b = e$ which is more common in the physics literature). α-entropies were also derived in a purely pragmatic manner in [12] as measures of information for concrete information-theoretic problems.

Rényi's definition becomes more natural if one notices that KN-averages are invariant under $\phi(x)\mapsto A\phi(x)+B$ and one replaces φ_α by

$$\phi_\alpha(x) = \frac{e^{(1-\alpha)x}-1}{1-\alpha} \equiv \ln_\alpha[\exp(x)] \tag{5}$$

where $\ln_\alpha(\cdot)$ is the deformed logarithm [14] ($\ln_1(\cdot) = \ln(\cdot)$).

The above (original) derivation of I_α clearly shows the two elements which led Rényi to the idea of α-entropy: (1) one needs a generalized average and (2) the random variable one averages is the logarithmic measure of information. The latter has a well known heuristic explanation which goes back to Hartley [13]: To uniquely specify a single element of a set containing N numbers one needs $\log_2 N$ bits; but, if one splits the set into n subsets containing, respectively, $N_1,\ldots,N_n$ elements ($\sum_i N_i = N$) then in order to specify only in which set the element of interest is located it is enough to have $\log_2 N - \log_2 N_i = \log_2(N/N_i)$ bits of information. The latter construction ignores the information encoded in correlations between the subsets. For this reason typically one needs less information if such correlations are present. The idea is used in data-compression algorithms and is essential for the argument we will present below.

Although α-entropies are occasionally used in statistical physics [15] it seems the same cannot be said of KN-averages. Thinking of the original motivation behind generalized entropies one may wonder whether this is not logically inconsistent. Constructing statistical physics with α-entropies one should consistently apply KN-averaging to all random variables, internal energy included. Applying the procedure to thermostatistics one may expect to arrive at a one-parameter family of equilibrium states which, in the limit $\alpha\to 1$, reproduce Boltzmann-Gibbs statistics.

During the past ten years it became quite clear that there is a need for some generalization of standard thermostatistics, as exemplified by the unquestionable success of Tsallis' q-thermodynamics [16]. Systems with long-range correlations, memory effects or fractal boundaries are well described by $q\neq 1$ Tsallis-type equilibria. Gradual development of this theory allowed to understand that there is indeed a link between generalized entropies and generalized averages. However, the averages one uses in Tsallis' statistics are the standard linear ones but expressed in terms of the so-called escort probabilities. So there is no direct link to KN-averages.

In what follows we present a thermostatistical theory based on KN-averages. It deals with the problem of maximizing average information under the constraint that the average of some energy function has a given value. As we shall see there *is* a link between such a theory and Tsallis' thermostatistics. Actually, many technical developments obtained within the Tsallis scheme have a straightforward application in the new framework. An important difference with respect to the Tsallis theory is that we can obtain both non-extensive and extensive generalizations so that one may expect the formalism will have still wider scope of applications.

We begin with the KN-average depending on parameters p_k which we shall later identify with escort probabilities. α-entropy defined with the help of the modified KN-function (5) is

$$I_\alpha = \phi_\alpha^{-1}\Big(\sum_k p_k \phi_\alpha(I_k)\Big) = \phi_\alpha^{-1}\Big(\sum_k p_k \phi_\alpha(-\ln p_k)\Big) \quad (6)$$

$$= \phi_\alpha^{-1}\Big(\frac{\sum_k p_k^\alpha - 1}{1-\alpha}\Big) = \phi_\alpha^{-1}\Big(\sum_k p_k \ln_\alpha(1/p_k)\Big). \quad (7)$$

It is interesting that in the course of calculation of I_α the expression for the Daróczy-Tsallis entropy [15–17] arises

$$S_\alpha(p) = \frac{\sum_k p_k^\alpha - 1}{1-\alpha}. \quad (8)$$

This shows that in the context of KN-means there is an intrinsic relation between I_α and S_α:

$$\phi_\alpha(I_\alpha) = S_\alpha. \quad (9)$$

Let us note that the formula

$$\phi_\alpha(I_k) = \ln_\alpha(1/p_k) \quad (10)$$

may hold also for other pairs (ϕ_α, I_k), with ϕ_α not given by (5), and be valid even for measures of information different from the Hartley-Shannon-Wiener random variable $I_k = -\ln p_k$. The key assumption of the present paper is that the generalized theory is characterized by the properties (9) and (10). One can see (10) as a definition of I_k in case ϕ_α is given, or as a constraint on ϕ_α if I_k is given.

The generalized thermodynamics is obtained by maximizing I_α under the constraint of fixed internal energy

$$\langle \beta_0 H\rangle_{\phi_\alpha} = \phi_\alpha^{-1}\Big(\sum_k p_k \phi_\alpha(\beta_0 E_k)\Big) = \beta_0 U \quad (11)$$

where β_0 is a constant needed to make the averaged energy dimensionless. Equivalently, the problem may be reformulated as maximizing (8) under the constraint

$$\sum_k p_k \phi_\alpha(\beta_0 E_k) = \phi_\alpha(\beta_0 U). \quad (12)$$

This problem is of the type originally considered by Tsallis [16]. However, since then the formalism of non-extensive thermostatistics has evolved. In particular, one has learned [18] that the optimization problem should be reparametrized using the so called *escort* probabilities. The reason why one should do so is the following. The standard thermodynamical relation for temperature T is

$$\frac{1}{T} = \frac{\mathrm{d}S}{\mathrm{d}U} \quad (13)$$

with S and U, respectively, entropy and energy calculated using the equilibrium averages. In generalized thermostatistics this definition of temperature is not necessarily correct. Recently has been shown [28,29] that (13) is valid if the entropy is additive and must be modified in all other cases. The reparametrization of non-extensive thermostatistics, by introduction of escort probabilities, is such that energy U becomes generically an increasing function of some (unphysical) temperature T^* (see e.g. Prop. 3.5 of [24]), which is then related to physical temperature T.

The reparametrization is done by means of $q \leftrightarrow 1/q$ duality [18,19]. Let

$$\rho_k = \frac{p_k^\alpha}{\sum_k p_k^\alpha} \quad (14)$$

Then one has clearly also the inverse relation

$$p_k = \frac{\rho_k^q}{\sum_k \rho_k^q} \quad (15)$$

with $q = 1/\alpha$. The above optimization problem is now equivalent to maximizing $S_q(\rho)$ under the constraint

$$\frac{\sum_k \rho_k^q \phi_{1/q}(\beta_0 E_k)}{\sum_k \rho_k^q} = \phi_{1/q}(\beta_0 U) \quad (16)$$

This is so because $S_q(\rho)$ is maximal if and only if $S_{1/q}(p)$ is maximal (see [19]). The latter optimization problem is of the type studied in the new style non-extensive thermostatistics [18].

We are now ready to solve the optimization problem. The free energy F is defined by

$$\beta_0 F = \frac{\sum_k \rho_k^q \phi_{1/q}(\beta_0 E_k)}{\sum_k \rho_k^q} - \beta_0 T^* S_q(\rho) \quad (17)$$

Minima of $\beta_0 F$, if they exist [24,27], are realized for distributions of the form [18]

$$\rho_k \sim \frac{1}{[1 + a x_k]^{1/(q-1)}} \qquad \text{if } 1 < q \quad (18)$$

or

$$\rho_k \sim [1 - a x_k]_+^{1/(1-q)} \qquad \text{if } 0 < q < 1 \quad (19)$$

Here $x_k = \phi_{1/q}(\beta_0 E_k)$ and $[x]_+$ equals x if x is positive, zero otherwise. Expression (18), with $1/(q-1)$ replaced by $1+\kappa$, is called the kappa-distribution or generalized Lorentzian distribution [25]. There are several reasons why this distribution is of interest. In the first place, the Gibbs distribution, which determines the equilibrium average in the standard setting of thermodynamics [26], is obtained in the limit $\kappa \to +\infty$, or $q \to 1$. The kappa-distribution is frequently used. For example in plasma physics it is used to describe an excess of highly energetic particles [21]. Typical for distribution (19) is that the probabilities p_k are identically zero whenever $aE_k \geq 1$. This cut-off for high values of E_k is of interest in many areas of physics. In astrophysics it has been used [20] to describe stellar systems with finite average mass. A statistical description of an electron captured in a Coulomb potential requires the cut-off to mask scattering states [22,19]. In standard statistical mechanics the treatment of vanishing probabilities requires infinite energies which lead to ambiguities. These can be avoided if distributions of the type (19) are used.

The formulas that follow are based on results already found in literature at many places, e.g. in [24].

Assume first that $\alpha = 1/q$, $0 < \alpha < 1$. Then the equilibrium average is the KN-average with p_k given by

$$p_k = \frac{1}{Z_1} \frac{1}{[1+ax_k]^{1/(1-\alpha)}}. \tag{20}$$

Z_1 is given by

$$Z_1 = \sum_k \frac{1}{[1+ax_k]^{1/(1-\alpha)}}, \tag{21}$$

and $x_k = \phi_\alpha(\beta_0 E_k)$. The unknown parameter a has to be fixed in such a way that (12) holds. This condition can be written as

$$\phi_\alpha(\beta_0 U) = \frac{1}{a}\left(\frac{Z_0}{Z_1} - 1\right) \tag{22}$$

with Z_0 given by

$$Z_0 = \sum_k \frac{1}{[1+ax_k]^{\alpha/(1-\alpha)}} \tag{23}$$

The entropy I_α follows from (8) with (20). One obtains

$$\phi_\alpha(I_\alpha) = \frac{1}{1-\alpha}\left(\frac{Z_0}{Z_1^\alpha} - 1\right). \tag{24}$$

Temperature T^* is given by (cf. Eq. (14) in [24])

$$\frac{1}{\beta_0 T^*} = \frac{a\alpha}{1-\alpha} \frac{Z_1^2}{Z_0^{(1+\alpha)/\alpha}}. \tag{25}$$

Now assume $\alpha > 1$. Then the formulas become

$$p_k = \frac{1}{Z_1}[1 - ax_k]_+^{1/(\alpha-1)} \tag{26}$$

with Z_1 given by

$$Z_1 = \sum_k [1 - ax_k]_+^{1/(\alpha-1)}. \tag{27}$$

The expressions for energy and entropy are

$$\phi_\alpha(\beta_0 U) = \frac{1}{a}\left(1 - \frac{Z_0}{Z_1}\right) \tag{28}$$

and

$$\phi_\alpha(I_\alpha) = \frac{1}{\alpha-1}\left(1 - \frac{Z_0}{Z_1^\alpha}\right) \tag{29}$$

with

$$Z_0 = \sum_k [1 - ax_k]_+^{\alpha/(\alpha-1)} \tag{30}$$

Temperature T^* is given by

$$\frac{1}{\beta_0 T^*} = \frac{a\alpha}{\alpha-1} \frac{Z_1^2}{Z_0^{(1+\alpha)/\alpha}}. \tag{31}$$

Let us finally return to the specific case of Rényi's entropy, i.e. I_k and ϕ_α given, respectively, by (4) and (5). This choice is particularly interesting since only then the following three conditions are satisfied

$$\langle \beta_0 H + \beta_0 E \rangle_{\phi_\alpha} = \langle \beta_0 H \rangle_{\phi_\alpha} + \beta_0 E \tag{32}$$

$$\langle \beta_0 H_{A+B} \rangle_{\phi_\alpha} = \langle \beta_0 H_A \rangle_{\phi_\alpha} + \langle \beta_0 H_B \rangle_{\phi_\alpha} \tag{33}$$

$$I_\alpha(A+B) = I_\alpha(A) + I_\alpha(B) \tag{34}$$

where A and B are two uncorrelated noninteracting systems. Condition (32) when combined with the explicit form of equilibrium state means that equilibrium does not depend on the origin of the energy scale. The remaining two conditions imply that we have a one-parameter family of *extensive* generalizations of the Boltzmann-Gibbs statistics, the latter being recovered in the limit $\alpha \to 1$. For $\alpha = q^{-1} \neq 1$ we obtain the well known Tsallis-type kappa-distributions but with energies βE_k replaced by $\phi_\alpha(\beta_0 E_k)$.

In general, the equilibrium probabilities are not of the product form (there is one exception — see below). The product form is of course also absent in the standard formalism when there are correlations between subsystems. Nevertheless, if the correlations are not too strong, then the system in equilibrium is still extensive. This is expressed by stating that the so-called thermodynamic limit exists. We expect that also in the present formalism the thermodynamic limit exists, but this point has still to be studied.

Consider now the case $0 < \alpha < 1$ and $a = 1 - \alpha$. This is a remarkable case because the equilibrium distribution (20) becomes exponential. Indeed, one verifies that

$$p_k = \frac{1}{Z_1} e^{-\beta_0 E_k} \qquad \text{with } Z_1 = \sum_k e^{-\beta_0 E_k}. \tag{35}$$

Internal energy equals

$$\beta_0 U = \frac{1}{1-\alpha} \ln\left(\frac{1}{Z_1}\sum_k e^{-\alpha\beta_0 E_k}\right) = I_\alpha - \ln Z_1. \tag{36}$$

This means that for each system there exists a particular temperature where the equilibrium state is factorizable.

Let us summarize the results.

The formalism of thermostatistics based on KN-averages simultaneously generalizes Boltzmann-Gibbs and Tsallis theories. As opposed to the Tsallis case, which is always nonextensive, the KN-approach allows for a family of extensive generalizations. On the other hand, the family of extensive theories leads to equilibrium states which share many properties with Tsallis $q \neq 1$ distributions. Tsallis formalism enters the KN-formulation also via the relation between I_α and S_α. What is surprising is that one should not simply identify α with q. The correct relation $\alpha = 1/q$ is suggested by the fact that the probabilities p_k are interpreted as escort probabilities. In the present paper the function ϕ_α is kept constant while the probabilities p_k are varied. It could be interesting to consider also the case where ϕ_α is varied as well.

Of particular interest is the choice $\phi_\alpha(x) = \ln_\alpha(\exp x)$ since then the average information coincides with Rényi's entropy. As proved by Rényi, his entropy, together with that of Shannon [10], are the only additive entropies. As shown in [28,29] additivity of entropy is a requirement for physical temperature T to be defined by the usual thermodynamic relation (13). The formalism generalizes to other non-exponential choices of ϕ provided the information measure is adapted in such a way that (9) and (10) still hold. In this more general context entropy is no longer additive and the definition of physical temperature T by means of (13) becomes problematic. A correct definition could be derived along the lines of [28] or [29]. This problem requires further study.

In a natural way Tsallis' entropy appears as a tool for calculating equilibrium averages. This offers the opportunity to reuse the knowledge from Tsallis-like thermostatistics. A tempting question is whether in each of the many applications of Tsallis' thermostatistics one can find a natural KN-average which maps the problem into the present formalism.

In an extended version of this Letter we shall discuss explicit examples. Here we only mention that preliminary results for a two-level system give satisfactory results. In particular, we checked that the $\alpha \to 1$ limit of equilibrium distributions reproduces Boltzmann-Gibbs results, and that the relation between T and T^* was found as in [28]. Of course, more complicated examples should be studied. Examples like the one-dimensional Ising model could clarify the issue of the thermodynamic limit.

For the sake of completeness let us mention that Rényi's entropy has been studied already [23] in relation with the escort probabilities (14). One of the conclusions of that paper is that they obtain the same results as in Tsallis' thermostatistics, which is not a surprise since Rényi's entropy and Tsallis' entropy are monotonic functions of each other. There is nor further relation with the present work.

One of the authors (MC) wishes to thank the NATO for a research fellowship enabling his stay at the Universiteit Antwerpen.

[1] A. Kolmogorov, Atti della R. Accademia Nazionale dei Lincei **12**, 388 (1930).

[2] M. Nagumo, Japan. Journ. Math. **7**, 71 (1930).

[3] B. de Finetti, Giornale di Istituto Italiano del Attuarii **2**, 369 (1931).

[4] B. Jessen, Acta Sci. Math. **5**, 108 (1931).

[5] T. Kitagawa, Proceedings Physico-Mathematical Society of Japan **16**, 117 (1934).

[6] J. Aczél, Bull. Amer. Math. Soc. **54**, 392 (1948).

[7] A. Rényi, MTA III. Oszt. Közl. **10**, 251 (1960); reprinted in [30], pp. 526-552.

[8] A. Rényi, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 547-561, University of California Press, Berkeley (1961); reprinted in [30] pp. 565-580.

[9] Theorem 89 in G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge (1934).

[10] C. E. Shannon, Bell System Technical Journal **27**, 379 (1948); **27**, 623 (1948).

[11] N. Wiener, *Cybernetics*, Wiley, New York (1948).

[12] A. Rényi, Rev. Inst. Internat. Stat. **33**, 1 (1965); reprinted in [30] pp. 304-318.

[13] R.V. Hartley, Bell System Technical Journal, **7**, 535 (1928).

[14] C. Tsallis, Quimica Nova **17**, 468 (1994); E. P. Borges, J. Phys. A **31**, 5281 (1998).

[15] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

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

[17] Z. Daróczy, Inf. Control, **16**, 36 (1970).

[18] C. Tsallis, R.S. Mendes, A.R. Plastino, Physica A**261**, 543 (1998).

[19] J. Naudts, Chaos, Solitons and Fractals **13**(2) (2001), in press.

[20] A.R. Plastino, A. Plastino, Phys. Lett. A**174**, 384 (1993).

[21] N. Meyer-Vernet, M. Moncuquet, and S. Hoang, Icarus **116**, 202 (1995).

[22] L.S. Lucena, L.R. da Silva, C. Tsallis, Phys. Rev. E**51**, 6247 (1995)

[23] E.K. Lenzi, R.S. Mendes, L.R. da Silva, Physica A**280**, 337 (2000)

[24] J. Naudts, Rev. Math. Phys. **12**, 1305 (2000).

[25] A. V. Milovanov and L. M. Zelenyi, Nonlinear Processes in Geophysics, **7**, 211 (2000).

[26] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

[27] J. Naudts and M. Czachor, in *Nonextensive Statistical Mechanics and its Applications,* ed. S. Abe, Y. Okamoto, Lecture Notes in Physics 560 (Springer-Verlag, 2001), p. 243-252.

[28] S. Abe, A. Martinez, F. Pennini, A. Plastino, Phys. Lett. A**281**(2-3), 126 (2001); S. Martinez, F. Pennini, and A. Plastino, Physica A **295**, 246 (2001); **295**, 416 (2001).

[29] R. Toral, cond-mat/0106060.

[30] *Selected Papers of Alfréd Rényi,* vol 2. (Akadémiai Kiadó, Budapest, 1976).