Maximum Renyi Entropy Principle for Systems with Power-Law Hamiltonians

A. G. Bashkirov*

Institute Geospheres Dynamics of Russian Academy of Sciences, Moscow, Russia

(Received 16 February 2004; published 20 September 2004)

The Renyi distribution ensuring the maximum of Renyi entropy is investigated for a particular case of a power-law Hamiltonian. Both Lagrange parameters α and β can be eliminated. It is found that β does not depend on a Renyi parameter q and can be expressed in terms of an exponent κ of the power-law Hamiltonian and an average energy U. The Renyi entropy for the resulting Renyi distribution reaches its maximal value at $q = 1/(1+\kappa)$ that can be considered as the most probable value of q when we have no additional information on the behavior of the stochastic process. The Renyi distribution for such q becomes a power-law distribution with the exponent $-(\kappa+1)$. When $q = 1/(1+\kappa) + \epsilon$ ($0 < \epsilon \ll 1$) there appears a horizontal head part of the Renyi distribution that precedes the power-law part. Such a picture corresponds to some observed phenomena.

DOI: 10.1103/PhysRevLett.93.130601 PACS numbers: 05.10.Gg, 05.20.Gg, 05.40.–a

Numerous examples of power-law distributions (PLDs) are well known in different fields of science and human activity [1]. Power laws are considered [2] as one of the signatures of complex self-organizing systems. They are sometimes called the Zipf-Pareto law or fractal distributions. We can mention here the Zipf-Pareto law in linguistics [3], in economy [4], and in the science of sciences [5], geophysics [6], critical phenomena [7], models of granulated media [8], the impact fragmentation [9,10], etc.

Graphically, PLD is presented by a linear graph in a double logarithmic plot of frequency or cumulative number as a function of size. It should be noticed here that, in general, double logarithmic plots of data from phenomena in nature or economy often exhibit a limited linear regime preceded by a near-horizontal "head" part and followed by a tail of significant curvature. The latter deviation from a power-law description can be explained by a finite-size effect. In reality, for instance for the impact fragmentation, extrapolation of the PLD to infinite fragment masses would predict masses surpassing a mass of the target. This effect will not be considered here.

Successful derivations of PLD with the head part are based on Renyi or Tsallis distributions ensuring maximums of Renyi or Tsallis entropies correspondingly (see, e.g., [11,12]). However, the q parameter and Lagrange multiplier β remains undetermined there.

Here, this problem will be discussed on the basis of a generalization of the maximum entropy principle (MEP) for the Renyi entropy taking into account a variation of the q parameter. In the special case that a Hamiltonian of the system is a power-law function of a variable of the system the Lagrange multiplier β will be expressed in terms of an average energy of the system and exponent of the power-law Hamiltonian, and the q parameter will be determined uniquely.

According to MEP developed by Jaynes [13] for a Boltzmann-Gibbs statistics an equilibrium distribution of probabilities $p = \{p_i\}$ must provide a maximum of the Boltzmann information entropy $S_B(p) = -k_B \sum_i p_i \ln p_i$ upon additional conditions of normalization $\sum_i p_i = 1$ and a fixed average energy $U = \langle H \rangle_p \equiv \sum_i H_i p_i$. Then, the distribution $\{p_i\}$ is determined from the extremum of the functional

$$L_G(p) = -\sum_i^W p_i \ln p_i - \alpha_0 \sum_i^W p_i - \beta_0 \sum_i^W H_i p_i, \quad (1)$$

where α_0 and β_0 are the Lagrange multipliers. Its extremum is ensured by the Gibbs canonical distribution, in which β_0 is determined by condition of correspondence between Gibbs thermostatistics and classical thermodynamics as $\beta_0 = 1/k_B T$ where T is the thermodynamic temperature.

On the contrary, in applications of MEP to the Renyi entropy, the multiplier β may depend on q and its physical meaning is not evident; hence, it is to be determined.

If the Renyi entropy (RE)

$$S_R(p) = \frac{k_B}{1-q} \ln \sum_i p_i^q \quad (2)$$

is used instead of the Boltzmann entropy, the equilibrium distribution must provide the maximum of the functional

$$L_R(p) = \frac{1}{1-q} \ln \sum_i^W p_i^q - \alpha \sum_i^W p_i - \beta \sum_i^W H_i p_i, \quad (3)$$

where α and β are Lagrange multipliers. It can be noted that $L_R(p)$ passes to $L_G(p)$ in the $q \to 1$ limit.

We equate a functional derivative of $L_R(p)$ to zero, then

$$\frac{\delta L_R(p)}{\delta p_i} = \frac{q}{1-q} \frac{p_i^{q-1}}{\sum_j p_j^q} - \alpha - \beta H_i = 0. \quad (4)$$

To eliminate the parameter α we can multiply this equation by p_i and sum up over i, taking into account the normalization condition $\sum_i p_i = 1$. Then we get

$$\alpha = \frac{q}{1-q} - \beta U \quad (5)$$

and

$$p_i = \left[(1 - \beta \frac{q-1}{q} \Delta H_i \Big) \sum_j^W p_j^q \right]^{1/(q-1)}, \quad (6)$$

where $\Delta H_i = H_i - U$. Using once more the condition $\sum_i p_i = 1$ we get

$$p_i = p_i^{(R)} = Z_R^{-1} \left(1 - \beta \frac{q-1}{q} \Delta H_i \right)^{1/(q-1)}, \quad (7)$$

$$Z_R^{-1} = \sum_i \left(1 - \beta \frac{q-1}{q} \Delta H_i \right)^{1/(q-1)}. \quad (8)$$

This distribution may be called the Renyi distribution (RD). When $q \to 1$ the distribution $\{p_i^{(R)}\}$ becomes the Gibbs canonical distribution and $\beta/q \to \beta_0 = 1/k_B T$. Such behavior is not enough for unique determination of β. In reality, in general, it may be an arbitrary function $\beta(q)$ which becomes β_0 in the limit $q \to 1$.

To find an explicit form of β, we return to the additional condition of the preassigned average energy $U = \sum_i H_i p_i$ and substitute there the RD (7). For definiteness, we will confine the discussion to the particular case of a power-law dependence of the Hamiltonian on a parameter x,

$$H_i = C x_i^\kappa. \quad (9)$$

This type of the Hamiltonian corresponds to an ideal gas model in the Boltzmann-Gibbs thermostatistics and it seems reasonable to say that it may be useful in construction of thermostatistics of complex systems. Moreover, in most social, biological, and humanitarian sciences the system variable x can be considered (with $\kappa = 1$) as a kind of the Hamiltonian (e.g., the size of population of a country, effort of a word pronouncing and understanding, bank capital, etc.).

If the distribution $\{p_i\}$ allows for smoothing over the range much larger an average distance $\Delta x_i = x_i - x_{i+1}$ without sufficient loss of information, we can pass from the discrete variable x_i to the continuous one x. Then the condition of a fixed average energy becomes

$$Z^{-1} \int_0^\infty C x^\kappa \left[1 - \beta \frac{q-1}{q} (C x^\kappa - U) \right]^{1/(q-1)} dx = U \quad (10)$$

or

$$Z^{-1} \int_0^\infty C_u x^\kappa \left[1 - \beta U \frac{q-1}{q} (C_u x^\kappa - 1) \right]^{1/(q-1)} dx = 1, \quad (11)$$

where

$$Z = \int_0^\infty \left[1 - \beta U \frac{q-1}{q} (C_u x^\kappa - 1) \right]^{1/(q-1)} dx, \quad (12)$$

where $C_u = C/U$. Both integrals in these equations may be calculated with the use of a tabulated [14] integral

$$I = \int_0^\infty \frac{x^{\mu-1} dx}{(a + b x^\nu)^\lambda} = \frac{1}{\nu a^\lambda} \left(\frac{a}{b} \right)^{\mu/\nu} \frac{\Gamma[\frac{\mu}{\nu}] \Gamma[\lambda - \frac{\mu}{\nu}]}{\Gamma[\lambda]} \quad (13)$$

under condition of convergence

$$0 < \frac{\mu}{\nu} < \lambda, \qquad (\lambda > 1). \quad (14)$$

For the integrals in Eqs. (11) and (12) we get

$$0 < \frac{1 + \kappa}{\kappa} < \frac{1}{1 - q}. \quad (15)$$

Then, finally, we find from Eqs. (11) and (12) with the use of the relation $\Gamma[1 + z] = z\Gamma[z]$, that

$$\beta U = \frac{1}{\kappa} \quad \text{for all } q. \quad (16)$$

Independent of this relation on q means that it is true, in particular, for the limit case $q = 1$ where the Gibbs distribution takes a place and, therefore,

$$\beta = \beta_0 \equiv 1/k_B T \quad \text{for all } q. \quad (17)$$

When $H = p^2/2m$ (that is, $\kappa = 2$) we get from (16) and (17) that $U = \frac{1}{2} k_B T$, as would be expected for one-dimensional ideal gas.

Additionally, the Lagrange parameter β can be eliminated from the RD (7) with the use of Eq. (16) and we have, alternatively,

$$p^R(x|q, \kappa) = Z^{-1} \left[1 - \frac{q-1}{\kappa q} (C_u x^\kappa - 1) \right]^{1/(q-1)} \quad (18)$$

or

$$p_i^R(q, \kappa) = Z^{-1} \left[1 - \frac{q-1}{\kappa q} (C_u x_i^\kappa - 1) \right]^{1/(q-1)}. \quad (19)$$

So, at least for the power-law Hamiltonian, the Lagrange multiplier β does not depend on the Renyi parameter q and coincides with the Gibbs parameter $\beta_0 = 1/k_B T$, and, moreover, can be eliminated completely with the use of the relation (16).

The problem to be solved for a unique definition of the RD is the determination of a value of the Renyi parameter q.

Some successes in this direction were achieved for particular cases of a set of independent harmonic oscillators [15,16] and fractal systems [17].

An excellent example of the solution of this problem for a physical non-Gibbsian system was presented by Wilk and Wlodarczyk [18]. They took into consideration fluctuations of both energy and temperature in contrast to the traditional Gibbs method in which temperature is a constant. As a result, their approach led (see [19]) to the RD with the parameter q expressed via heat capacity C_V of the minor subsystem $q = (C_V - k_B)/C_V$. The approach by Wilk and Wlodarczyk was advanced by Beck [20] and Beck and Cohen [21] who offered for it a new apt term "superstatistics." In the frame of superstatistics, the parameter q is defined by physical properties of a system.

On the other hand, there are many stochastic systems for which we have no information related to a source of fluctuations. In that case the parameter q cannot be determined with the use of the superstatistics.

Here, a further extension of MEP is proposed. It consists of looking for a maximum of the RE in a space of the RDs with different values of q.

The next step consists of substitution of the RD $p^R(x|q, \kappa)$ into the definition of the RE, Eq. (2), and variation of the q parameter. The resultant picture of $S_R[p^R(x|q, \kappa)]$ as a function of q is illustrated in Fig. 1 (top). It is seen that $S_R[p^R(x|q, \kappa)]$ attains its maximum at the minimal possible value of q which fulfills the inequality (15), that is,

$$q_{\min} = \frac{1}{1+\kappa}. \tag{20}$$

For $q < q_{\min}$, the integral (10) diverges and, therefore, the RD does not determine the average value $U = \langle H \rangle_p$, that is a violation of the second condition of MEP [22].

To check self-consistency of the proposed extension of MEP the similar procedure is applied to the Boltzmann-Shannon entropy $S_B(p)$. Substituting there $p = p^R(x|q, \kappa)$ we get the q-dependent function $S_B[p^R(x|q, \kappa)]$ illustrated in Fig. 1 (bottom). As would be expected, the Boltzmann-Shannon entropy $S_B[p^R(x|q, \kappa)]$ attains its maximum value at $q = 1$ where $p^R(x|q, \kappa)$ becomes the Gibbs canonical distribution.

Thus, it is found that the maximum of the RE is realized at $q = q_{\min}$ and it is just the value of the Renyi parameter that should be used for the discussed particular case of the power-law Hamiltonian if we have no additional information on behavior of the stochastic process under consideration.

Substitution of $q = q_{\min}$ into Eq. (18) leads to

$$p \sim x^{-(\kappa+1)}. \tag{21}$$

Thus, for $q = q_{\min}$ the RD for a system with the power-law Hamiltonian becomes a PLD over the whole range of x.

For a particular case of the impact fragmentation where $H \sim m^{2/3}$ the power-law distribution of fragments over their masses m follows from (21) as $p(m) \sim m^{5/3}$ that coincides with results of our previous analysis [10] and experimental observations [9].

For another particular case, $\kappa = 1$, PLD is $p \sim x^{-2}$. Such a form of the Zipf-Pareto law is the most useful in social, biological, and humanitarian sciences. The same exponent of PLD was demonstrated [24] for energy spectra of particles from atmospheric cascades in cosmic ray physics and for distribution of users among the web sites [25].

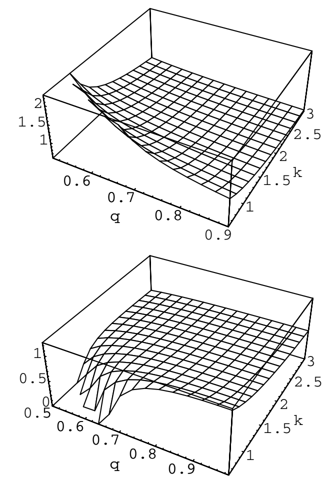

FIG. 1. The entropies $S_R[p^R(x|q, \kappa)]/k_B$ (top) and $S_B[p^R(x|q, \kappa)]/k_B$ (bottom) for the power-law Hamiltonian with the exponent κ within the range $3 > \kappa > 0.5$ and $q > 1/(1+\kappa)$.

It is necessary to notice here that inequalities (15) suggest in fact $q > q_{\min}$, that is, $q = q_{\min} + \epsilon$, where ϵ is a positive infinitesimal value. It is clear that ϵ ($\ll 1$) should be a finite constant at physical realizations. Taking into account the finite ϵ gives rise to the RD in the form

$$p^R(x) = Z^{-1}(C_u x)^{-(\kappa+1)\{1+\epsilon[(\kappa+1)/\kappa]\}}[1 - \epsilon(\kappa+1)^2(1 - C_u x^{-\kappa})]^{-[(\kappa+1)/\kappa]\{1+\epsilon[(\kappa+1)/\kappa]\}}. \tag{22}$$

For sufficiently great x's this RD passes to PLD where all terms with ϵ can be neglected.

On the other hand, for sufficiently small x, only the term $\epsilon(\kappa+1)^2 C_u x^{-\kappa}$ may be accounted for in the expression in the square brackets, so we get

$$p^R(x)|_{x\ll 1} \sim (\epsilon(\kappa+1)^2).^{-(\kappa+1)/\kappa} \tag{23}$$

This equation points to the fact that the asymptote to the RD for small x's is a constant.

The picture of the RD over the whole range of x is illustrated in Fig. 2.

Now there is no method for a unique theoretical determination of ϵ, so it may be considered as a free parameter. It can be estimated for those experimental data where the head part preceding PLD is presented. As an example, for the probability distribution of connections in World Wide Web network [26] where ϵ is estimated as $\sim 10^{-4}$.

In summary, the maximum entropy principle applied to the RE gives rise to a RD that depends on the Renyi parameter q and two Lagrange multipliers α and β. The multiplier α corresponds to the condition of normalization of the distribution and may be eliminated with ease. The second Lagrange multiplier β corresponds to the condition of a fixed average energy $U = \langle H \rangle_p$ just as

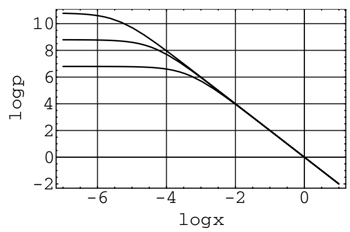

FIG. 2. Log-log graph of the RDs $p^R(x)$ (non-normalized) for the power-law Hamiltonian $H \sim x^\kappa$ ($\kappa = 1$) and different values $\epsilon = 10^{-6}, 10^{-5}, 10^{-4}$ from top to bottom.

$\beta_0 = 1/k_B T$ in the Gibbs distribution function. The connection of β_0 with the thermodynamic temperature obviates the necessity to eliminate the second Lagrange multiplier from the Gibbs distribution. It is not so for β, at least until the new Renyi thermostatistics is constructed and β obtains a physical meaning.

It is shown here that for the particular case of a power-law Hamiltonian $H_i = Cx^\kappa$ the Lagrange multiplier β does not depend on the Renyi parameter q and coincides with β_0. Moreover, it can be expressed in terms of U and κ and thus eliminated completely from the RD function.

In the absence of any additional information on a nature of the stochastic process, the q parameter can be determined with the further use of MEP in the space of the q-dependent RDs. Maximum of the RE is found at $q = 1/(1 + \kappa)$. The RD for such q becomes a power-law distribution with the exponent $-(1 + \kappa)$ that agrees with observable data for stochastic systems.

When applying such MEP to the Boltzmann entropy for the q-dependent RD, $S_B[p^R(x|q, \kappa)]$, the maximum is found at $q = 1$ that corresponds to the Gibbs distribution, as would be expected.

It should be noted that all above estimations of the parameters of the RD (7) and the exponent of PLD are true as well for the escort version of the Tsallis' distribution [11]

$$p_i^{(\mathrm{Ts})} = Z_{\mathrm{Ts}}^{-1}(1 - \beta^*(1 - q')\Delta H_i)^{q'/(1-q')} \qquad (24)$$

because of both distributions are identical if $q' = 1/q$. In reality, in this case

$$1 - q' = \frac{q-1}{q}, \qquad \frac{q'}{1-q'} = \frac{1}{q-1}, \qquad (25)$$

and β^* is determined by the same second additional condition of MEP as well as β.

It may be supposed that the parametric generalization of MEP proposed here will be useful for other parametric entropies (see, e.g., [27,28]).

I acknowledge fruitful discussions of the subject with A. Vityazev and S. Abe.

*Electronic address: abas@idg.chph.ras.ru

[1] R. Mantegna, Physica (Amsterdam) **277A**, 136 (2000).
[2] P. Bak, *How Nature Works: The Science of Self-Organized Criticality* (Copernicus, New York, 1996).
[3] G. K. Zipf, *Human Behavior and the Principle of Least Efforts* (Addison-Wesley, Cambridge, 1949).
[4] B. B. Mandelbrot, J. Business **36**, 394 (1963).
[5] D. Price, *Little Science, Big Science.* (Columbia University, New York, 1963).
[6] K. Kasahara, *Earthquake Mechanics* (Cambridge University Press, Cambridge, 1981).
[7] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic, London, 1982).
[8] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
[9] A. Fujiwara, G. Kamimoto, and A. Tsukamoto, Icarus **31**, 277 (1977).
[10] A. G. Bashkirov and A. V. Vityazev, Planet. Space Sci. **44**, 909 (1996).
[11] C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica (Amsterdam) **261A**, 534 (1998).
[12] A. G. Bashkirov and A. V. Vityazev, Physica (Amsterdam) **277A**, 136 (2000).
[13] E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **107**, 171 (1957).
[14] I. S. Gradstein and I. M. Ryzhik, *Tables of Integrals, Sums, Series and Productions* (Academic Press, New York, 1994), 5th ed.
[15] A. R. Plastino and A. Plastino, Phys. Lett. A **193**, 251 (1994).
[16] A. Plastino and A. R. Plastino, Braz. J. Phys. **29**, 50 (1999).
[17] C. Tsallis, Chaos Solitons Fractals **13**, 371 (2002).
[18] G. Wilk and Z. Wlodarczyk, Phys. Rev. Lett. **84**, 2770 (2000).
[19] A. G. Bashkirov and A. D. Sukhanov, Zh. Eksp. Teor. Fiz. **122**, 513 (2002) [Sov. Phys. JETP **95**, 440 (2002)].
[20] C. Beck, Phys. Rev. Lett. **87**, 180601 (2001).
[21] C. Beck and E. G. D. Cohen, Physica (Amsterdam) **322A**, 267 (2003).
[22] While preparing this manuscript for publication the paper [23] appeared where a generalization of MEP including variations of both the distribution function and q parameter was proposed. There, an optimal value of q was looked for by the standard way of equating to zero of q derivative of the variational functional. Such an approach is incorrect here because the optimal value of q lies on the boundary of the range of possible values of q.
[23] A. R. Plastino, H. G. Miller, and A. Plastino, Continuum Mech. Thermodyn. **16**, 269 (2004).
[24] M. Rybczyski, Z. Wlodarczyk, and G. Wilk, Nucl. Phys. B, Proc. Suppl. **97**, 81 (2001).
[25] L. A. Adamic and B. A. Hubermann, Glottometrics **3**, 143 (2002).
[26] G. Wilk and Z. Wlodarczyk, Acta Phys. Pol. B **35**, 871 (2004).
[27] G. Kaniadakis, Phys. Rev. E, **66**, 056125 (2002).
[28] E. P. Borges and I. Roditi, Phys. Lett. A **246**, 399 (1998).