PHYSICAL REVIEW E **67**, 0671XX (2003)

Additive generalization of the Boltzmann entropy

Alexander N. Gorban*
ETH Zürich, Department of Materials, Institute of Polymers, ETH-Zentrum, CH-8092 Zürich, Switzerland and Institute of Computational Modeling RAS, 660036 Krasnoyarsk, Russia

Iliya V. Karlin† and Hans Christian Öttinger‡
ETH Zürich, Department of Materials, Institute of Polymers, ETH-Zentrum, CH-8092 Zürich, Switzerland

(Received 21 June 2002)

There exists a unique extension of the classical Boltzmann entropy functional to a one-parametric family of additive trace-form entropy functionals. We find the analytical solution to the corresponding deformation of the classical ensembles, and present an example of the deformation of the uncorrelated state caused by finiteness of the number of particles.

DOI: 10.1103/PhysRevE.67.0671XX PACS number(s): 05.70.Ln, 05.20.Dd

I. INTRODUCTION

The growing interest in nonclassical entropies in recent years [1,2] is motivated by the fact that they can be used to describe observable statistical effects such as the following.

(i) Nonclassical tails of distribution functions which can deviate significantly from the Gaussian distribution. In particular, this asymptotics can be power law (''long tails'') or, instead, distribution functions can decay in a more rapid fashion (''short tails''), in particular, they can become equal to zero at finite distance (''cut tails'').

(ii) Strong correlations between subsystems in equilibrium and quasiequilibrium states.

The entropy-based description of these effects in the spirit of the Gibbs ensembles is advantageous both in static and dynamic problems. For the latter (dynamic) aspect, we refer here to a vast literature on theories of nonequilibrium statistical thermodynamics (see, e.g., Ref. [3]), as well as entropy-based kinetic modeling [4].

Usually, when one attempts to introduce nonclassical entropies, there is a price to be paid. Nonclassical entropies at use in most of the contemporary studies violate at least one of the following important and familiar properties of the BGS entropy: (i) Additivity—the entropy of the system which is composed of independent subsystems equals the sum of the entropies of the subsystems, (ii) trace form—the entropy is a sum over the states (see below). For example, the Tsallis entropy [1] is not additive, the Rényi entropy [5] is not of the trace form. A useful discussion of various properties of the entropy can be found in Ref. [6].

Recently [7], it was indicated that there exists a one-parametric family of concave entropy functions which satisfy both the conditions (additivity and trace form) simultaneously. This family is essentially the linear convex combination between the Boltzmann entropy and the so-called Burg entropy (cf. Ref. [7]). While the existence of such a family of additive entropies was eventually mentioned some time ago [8], the result of the paper [7] indicates that it can be most pertinent to a unified study of systems out of the strict thermodynamic limit.

The finding of the present paper is that the solution to the maximization problem pertinent to this class of entropies is actually tractable analytically almost as efficiently as the classical Gauss distribution.

The structure of the paper is as follows. In Sec. II, we describe the one-parametric family of additive entropies [7] for the sake of completeness. In Sec. III, we demonstrate that the maximum entropy problem for the family of entropies [7] reduces to studying of one function of one variable. This result enables the analytical formulas for the deformation of the classical ensembles around the thermodynamic (BGS) limit in Sec. IV. An example of such a deformation of the classical uncorrelated ensemble of the configurational N-body distribution function is discussed in Sec. V. Concluding remarks are given in Sec. VI.

II. ADDITIVE TRACE-FORM ENTROPIES

For the sake of presentation, we consider a finite set of states characterized by the probabilities p_i (finiteness and discreteness are by no means the crucial restrictions, and are employed only in order to avoid the convergence questions). We consider systems which allow for a positive equilibrium, $p_i^*>0$ (for infinite systems, it is often advantageous to use unnormalized p^*). Then, any convex function of one variable, $h(x)$, defines the *trace-form* convex function of the probability distribution $H_h(p)$:

$$H_h(p)=\sum_i p_i^* h(p_i/p_i^*). \tag{1}$$

[We consider below the convex H_h functions rather than the concave entropy functions $S_h=-H_h$. The variety of the convex functions (1) was viewed in Refs. [7,8] as a set of Lyapunov functions of the master equation with p^* the equilibrium, but this is not essential to our present discussion.]

Among the set of the trace-form functions (1), there exists a one-parametric subset of the *additive* functions, H_β, $0\leqslant\beta\leqslant 1$:

*Electronic address: gorban@icm.krasn.ru
†Corresponding author: ikarlin@mat.ethz.ch
‡Electronic address: hco@ifp.mat.ethz.ch

$$H_\beta = \sum_i p_i^* h_\beta(p_i/p_i^*),$$

$$h_\beta(x) = (1-\beta)x \ln x - \beta \ln x. \tag{2}$$

In particular,

$$H_0 = \sum_i p_i \ln(p_i/p_i^*), \tag{3}$$

$$H_1 = -\sum_i p_i^* \ln(p_i/p_i^*). \tag{4}$$

The function H_0 is the BGS entropy (also, in the form presented, sometimes referred to as the Kullback-Leibler entropy for the reference equilibrium explicitly indicated). The function H_1 is the Burg entropy for p^* as the equipartition, in the present form first given in Ref. [8], to the best of our knowledge. Additivity of functions H_β (2) is readily checked [7,8]: *If* $p = p_{ij} = q_i r_j$, and also *if* $p^* = p_{ij}^* = q_i^* r_j^*$, *then*

$$H_\beta(p) = H_\beta(q) + H_\beta(r). \tag{5}$$

In order to avoid a possible confusion caused by a variety of ways, the notion of additivity of the entropy is used in current literature (see, e.g., Ref. [9]), and we note that the additivity of family (2) is understood in the traditional sense, that is, the usual statistical independence (factorization of the distribution) *implies* Eq. (5). Though we do not prove it here rigorously, the argument why family (2) represents all of the additive functions of the trace form (1) (up to a constant factor and adding a constant) is readily available: Treatment of the additivity condition, $H_h(\boldsymbol{qr}) = H_h(\boldsymbol{q}) + H_h(\boldsymbol{r})$ as a functional equation in order to determine the function h results in averaging of the vector function $\ln q_i r_j$; this can be done either with the joint probability $\boldsymbol{qr}$ (which leads to the BGS case) or with the equilibrium joint probability $\boldsymbol{q}^*\boldsymbol{r}^*$ which leads to the Burg case. The rest follows by convexity of their combination. Note that the second possibility (averaging with $\boldsymbol{q}^*\boldsymbol{r}^*$) is not mentioned in many sources (for example, the classical review by Wehrl [6]) because Burg's entropy is not continuous if some of the probabilities tend to zero. This is, however, one of the possibilities to account for finiteness (see below).

III. THE MAXIMUM ENTROPY PROBLEM

Since a factor in front of H_β is irrelevant, it proves convenient to use a different parametrization of family (2), H_α, $\alpha = \beta/(1-\beta)$, $\alpha \geqslant 0$:

$$H_\alpha = \sum_i [p_i \ln(p_i/p_i^*) - \alpha p_i^* \ln(p_i/p_i^*)]. \tag{6}$$

Parametric representation (6) will be used below. The limiting case $\alpha \to \infty$ corresponds to the pure Burg entropy (4), and it should be considered separately.

The major input into all the applications of the entropy functionals in statistical physics is the description of the quasiequilibria. Quasiequilibrium is the probability distribution which brings to maximum the entropy $S(p)$ at fixed values of the slow variables (their choice depends on the physics of a given problem), $M = m(p)$:

$$S(p) \to \max, \quad m(p) = M. \tag{7}$$

In order to address the construction of the quasiequilibrium in a general setting, we assume the macroscopic variables $M = m(p)$, where $M_s = \Sigma_i m_{si} p_i$, and consider problem (7) with $S = -H_\alpha$. The method of Lagrange multipliers implies

$$\frac{\partial H_\alpha}{\partial p_i} = \lambda_0 + \sum_s \lambda_s m_{si}, \tag{8}$$

where Lagrange multiplier λ_0 corresponds to normalization, and λ_s to the rest of the constraints. Let us denote $-\Lambda_i$ the right hand side of Eq. (8). With this, Eq. (8) may be written,

$$\ln(p_i/p_i^*) - \alpha(p_i^*/p_i) = -\Lambda_i. \tag{9}$$

Solution to an equation

$$\ln q - \alpha q^{-1} = -\Lambda \tag{10}$$

may be written as follows:

$$q = e^{-\Lambda} e^{\mathrm{lm}(\alpha e^\Lambda)}, \tag{11}$$

where we have introduced notation lm$a$ (modified logarithm) for the function which is the solution to the transcendent equation,

$$x e^x = a.$$

The function lm satisfies the following identities:

$$\mathrm{lm}\ a = \ln a - \ln \mathrm{lm}\ a, \tag{12}$$

$$\mathrm{lm}\ a = \ln a - \ln\{\ln a - \ln[\ln a - \ln(\cdots)]\cdots\}. \tag{13}$$

Identity (13) is the recurrent application of identity (12). A different representation of solution (11) reads

$$q = \frac{\alpha}{\mathrm{lm}(\alpha e^\Lambda)}. \tag{14}$$

From representation (11), the asymptotics at $\alpha \to 0$, and fixed Λ, is obvious: $q \to e^{-\Lambda}$, and which corresponds to the usual Boltzmann distribution. On the other hand, representation (14) reveals the asymptotics at $\Lambda \to \infty$:

$$q \sim \frac{\alpha}{\ln \alpha + \Lambda}. \tag{15}$$

For a symmetric distribution on the axis, and for $\Lambda = \lambda_0 + \lambda_2 x^2$, the first of the limits just mentioned gives the Gaussian distribution, while the second limit gives the Cauchy distribution. The corresponding distribution function for the limiting case H_∞ is simply the Cauchy distribution on the axis. We note it in passing that among nonsymmetric Cauchy distributions of the form $p=(\lambda_0+\lambda_1 x+\lambda_2 x^2)^{-1}$, there are distinguished cases with a twice degenerated zero in the denominator: $p=[\lambda(x-a)]^{-2}$. When one attempts to normalize this distribution by choosing a convergent sequence of functions, one gets a Dirac δ function $\delta(x-a)$ which can be interpreted as a microcanonical ensemble.

Thus, the quasiequilibrium distribution has the form

$$p=p^* e^{-\Lambda} e^{\mathrm{lm}(\alpha e^{\Lambda})}=\frac{\alpha p^*}{\mathrm{lm}(\alpha e^{\Lambda})}. \tag{16}$$

[We have omitted indices of states in p, p^*, and Λ.] Analytical formula (16) is the main result of this paper. Several remarks are in order.

(i) We have worked out the deformation of the quasiequilibrium ensembles using the dual variable Λ while the dependence $\Lambda(M)$ has been kept implicit. In general, manifolds of quasiequilibrium states are well defined in terms of dual variables (Lagrange multipliers). What is not always well defined for the distributions with long tails is the moment chart of these manifolds (Lagrange multipliers cannot be expressed in terms of moments if the latter do not exist). For example, the manifold of Cauchy distributions mentioned above is parametrized by Lagrange multipliers, whereas the parametrization in terms of the second moment does not exist. In these cases, a regularization is required, which assumes taking into account finiteness of the physical phase space (cf. Ref. [3]). Other possibilities to parametrize quasiequilibrium manifolds were worked out in applications to the Tsallis entropy where one uses nonlinear functionals of the distribution function (so-called escort probabilities, Refs. [2,10]). The use of these nonlinear parametrizations, however, leads to inconsistencies when dynamic problems are addressed (cf. Ref. [11]).

(ii) Let us indicate a remarkable formal extension of result (16) to $\alpha<0$ [or, alternatively, to $\beta<0$ in representation (2)] when the entropy function (6) loses convexity. Function $\mathrm{lm}a$ is defined, and is continuous, for $a\geqslant -e^{-1}$ ($\mathrm{lm}a\geqslant -1$). At $a\to -e^{-1}$, we have the limit $d\mathrm{lm}a/da\to\infty$. If we formally extend $\mathrm{lm}a=-\infty$ for $a<-e^{-1}$, then Eq. (16) is a distribution with a compact support ("cut tail"). With this, there will be defined a nonzero ratio p/p^*:

$$\inf\{p/p^*|p\neq 0\}\geqslant|\alpha|>0, \tag{17}$$

that is, either $p\geqslant|\alpha|p^*$ or $p=0$. This is similar to a Maxwell construction of a stretched spinodal (the cut at the inflection point), and not to the global maximum of the entropy. Whereas such constructions are always necessary when working with nonconvex thermodynamic potentials, we will not further discuss the case $\alpha<0$ in this paper.

IV. ENSEMBLES NEAR THE BGS LIMIT

Using Eq. (16), it is possible to study perturbatively deformations of quasiequilibrium ensembles near the thermodynamic limit. For the classical BGS entropy ($\alpha=0$), the quasiequilibrium distribution has the form

$$p=p^* e^{-\Lambda}. \tag{18}$$

To first order in α, we get

$$p=p^*(e^{-\Lambda}+\alpha)+o(\alpha). \tag{19}$$

The first-order deformation amounts to just a homogeneous shift of all quasiequilibrium populations (see also example below).

In order to compute the quasiequilibrium to second order in α, we must use the expansion of $\mathrm{lm}a$ to third order,

$$\mathrm{lm}\,a=a-a^2+(3/2)a^3+o(a^3).$$

Then

$$p=p^*\left(e^{-\Lambda}+\alpha-\frac{1}{2}\alpha^2 e^{\Lambda}\right)+o(\alpha^2). \tag{20}$$

Further corrections can also be easily computed using higher-order terms in the expansion of the lm. We now shall consider a specific example of formula (20).

V. CORRELATIONS CAUSED BY FINITENESS

In order to illustrate the effect of second-order deviations from the BGS case, we apply Eq. (20) to the classical quasiequilibrium defined by the one-particle configurational distribution function $f_1(r)$, where r is position variable. Assuming the equipartition for the reference equilibrium, $p^*=1/V^N$, where V is the volume of the system and N is the number of particles, we get $e^{-\Lambda}=e^{\lambda_0}\prod_{i=1}^N\Psi(r_i)$, where the Lagrange multiplier λ_0 is responsible for normalization. Then the N-body quasiequilibrium distribution function to second order in α reads

$$V^N p=e^{\lambda_0}\prod_{i=1}^N \Psi(r_i)+\alpha-\frac{\alpha^2}{2e^{\lambda_0}\prod_{i=1}^N \Psi(r_i)}+o(\alpha^2). \tag{21}$$

Our goal now is to compute the two-body configurational distribution function

$$f_2(r,q)=N(N-1)\int p(r,q,r_3,\dots,r_N)dr_3\cdots dr_N,$$

in quasiequilibrium (21). We recall that the classical result for the BGS entropy gives the uncorrelated two-body distribution, $f_2(r,q)\sim f_1(r)f_1(q)$, which also corresponds to the limit ($\alpha=0$) of Eq. (21). Computation to the order α^2 requires expansion of Lagrange multipliers λ_0 and Ψ to the corresponding order. This computation is straightforward although tedious, thus we give here only the final result: The two-body quasiequilibrium configurational distribution function f_2 reads

$$\frac{N}{N-1}f_2(r,q)=(1+\alpha+\alpha^2)\tilde{f}_1(r)\tilde{f}_1(q)+\alpha n^2 \underline{-\frac{\alpha^2}{2}n^2B^N\varphi_1(r)\varphi_1(q)}+o(\alpha^2), \quad (22)$$

where $n=N/V$ is the average number density, and where we have introduced notations,

$$\tilde{f}_1(r)=f_1(r)-\alpha n, \quad (23)$$

$$\varphi_1(r)=\frac{f_1(r)}{n}-\frac{n}{Bf_1(r)}, \quad (24)$$

$$B=\frac{1}{V}\int_V\frac{n}{f_1(r)}dr. \quad (25)$$

It is readily checked that result (22) gives $f_2=(N-1)N^{-1}f_1f_1$ at $\alpha=0$, which is identical with the classical uncorrelated pair distribution.

The first two terms in Eq. (22) amount again to the uncorrelated state with homogeneously shifted one-particle distributions [$\tilde{f}_1$ (23) instead of f_1, which amounts to a homogeneous subtraction of the average density times α].

The underlined term (of the order of α^2) is the contribution responsible for correlations caused by finiteness. Note that this extra correlation also has a form of a product, but not of the distribution functions, rather, of functions of one variable (24). In order to see the effect of this term more explicitly, we assume fluctuations around the homogeneous density in the thermodynamic limit,

$$f_1(r)=n[1+\zeta(r)N^{-1/2}], \quad (26)$$

where ζ is a function with zero average, and finite amplitude, $\langle\zeta\rangle=0$, $\langle\zeta^2\rangle=\sigma^2$, where we have introduced notation for averaging over the volume, $\langle h\rangle=V^{-1}\int_V h dr$. Note that the amplitude of the inhomogeneity is realistic, and it scales as $N^{-1/2}$ in full accordance with the classical theory of fluctuations. Assuming large (but finite) number of particles, we find to the leading order in *N*,

$$B=1+\sigma^2N^{-1}+o(N^{-1}), \quad B^N=e^{\sigma^2}+o(1).$$

Thus, specializing to the trial one-body distribution function (26), the deformation to second order of the uncorrelated (in the thermodynamic limit) two-body distribution function reads

$$\frac{N}{N-1}f_2(r,q)\approx(1+\alpha+\alpha^2)\tilde{f}_1(r)\tilde{f}_1(q)+\alpha n^2 -2\alpha^2n^2\sigma^2e^{\sigma^2}N^{-1}\theta(r)\theta(q), \quad (27)$$

where we have denoted $\theta=\sigma^{-1}\zeta$, $\langle\theta^2\rangle=1$.

Note that the correlations induced by finiteness of the system in the present example are highly nonlinear due to the dependence of the function φ_1 on the one-body distribution f_1. For that reason, deformation of the uncorrelated state (22) should be significant, in particular, to the corresponding derivations of the Vlasov mean-field kinetic equation from *N*-particle dynamics. This interesting problem is left for future work.

VI. CONCLUSION

The one-parametric family of the additive trace-form entropy functions considered in this paper is a convex linear combination of the classical Boltzmann entropy and of the Burg entropy, whereas the maximum entropy states are nonlinear combinations of the Gaussian and Cauchy distributions (in the case of the second moment as the macroscopic variable, and generalizations thereof for different macroscopic variables). This feature (trace form and additivity simultaneously) distinguishes the present family of entropies among many suggestions in the recent past. We have found the analytic solution to the maximum entropy problem in terms of one function of one variable, which enables to study perturbations of classical ensembles near the thermodynamic limit. The corresponding deformation of the uncorrelated (in the thermodynamic limit) state is established.

The asymptotic formula (15) reveals that the tail of the distributions in this theory is parameter independent, and is always Cauchy-like when α is away from zero. This is different, in particular, from the Tsallis case which leads to algebraic tails with the power dependence on the Tsallis parameter *q*. The present theory is explicitly focused on studying perturbations to the thermodynamic limit, and this universality of the tails is remarkable. On the other hand, a different general mechanism for nonclassical entropies was also indicated in Ref. [7], and it is related to the incomplete description (akin the Fermi-Dirac entropy of the electron-whole systems). This may affect the behavior at the tail of the distribution. In a subsequent publication [12], we shall study how the present entropy plus the incomplete description perform in fitting the recent turbulence data results of Beck *et al.* [13].

[1] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[2] *Nonextensive Statistical Mechanics and Its Applications*, edited by S. Abe and Y. Okamoto (Springer-Verlag, Berlin, 2001).
[3] A.N. Gorban and I.V. Karlin, in: *Developments of Mathematical and Experimental Physics*, edited by A. Macias *et al.*, Hydrodynamics and Dynamical Systems, Vol. C (Kluwer Academic, Dordrecht/Plenum Publishers, New York, 2003).
[4] S. Succi, I.V. Karlin, and H. Chen, Rev. Mod. Phys. **74**, 1203 (2002).
[5] A. Rényi, *Probability Theory* (North-Holland, Amsterdam, 1970).
[6] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[7] A.N. Gorban and I.V. Karlin, Phys. Rev. E **67**, 016104 (2003).

[8] A.N. Gorban, *Equilibrium Encircling. Equations of Chemical Kinetics and Their Thermodynamic Analysis* (Nauka, Novosibirsk, 1984).
[9] G. Kaniadakis, Phys. Rev. E **66**, 056125 (2002).
[10] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems: An Introduction* (Cambridge University Press, Cambridge, 1993).
[11] I.V. Karlin, M. Grmela, and A.N. Gorban, Phys. Rev. E **65**, 036128 (2002).
[12] P. Ilg and I. V. Karlin (unpublished).
[13] C. Beck, G.S. Lewis, and H.L. Swinney, Phys. Rev. E **63**, 035303(R) (2001).