Group entropies, correlation laws, and zeta functions

Piergiulio Tempesta
Departamento de Física Teórica II, Facultad de Físicas, Ciudad Universitaria, Universidad Complutense, E-28040 Madrid, Spain
(Received 15 February 2011; revised manuscript received 3 May 2011; published 10 August 2011)

The notion of group entropy is proposed. It enables the unification and generalization of many different definitions of entropy known in the literature, such as those of Boltzmann-Gibbs, Tsallis, Abe, and Kaniadakis. Other entropic functionals are introduced, related to nontrivial correlation laws characterizing universality classes of systems out of equilibrium when the dynamics is weakly chaotic. The associated thermostatistics are discussed. The mathematical structure underlying our construction is that of formal group theory, which provides the general structure of the correlations among particles and dictates the associated entropic functionals. As an example of application, the role of group entropies in information theory is illustrated and generalizations of the Kullback-Leibler divergence are proposed. A new connection between statistical mechanics and zeta functions is established. In particular, Tsallis entropy is related to the classical Riemann zeta function.

DOI: 10.1103/PhysRevE.84.021121

PACS number(s): 05.20.—y, 05.90.+m, 02.20.—a

I. INTRODUCTION

In many physical contexts, as well as in economy, biology, and social sciences, the nature of the correlations among the parts or subsystems constituting a given system is crucial for the understanding of the underlying dynamics. A possible example is offered by the present international economical context. Indeed, the simplifying hypothesis of microeconomical independence in local or national markets appears to be inadequate to describe the evolution of the present economical scenario: the strong and unexpected correlations that the global crisis has generated among previously unrelated economic and financial entities render inadequate the foresight scenario: the strong and unexpected correlations that the global crisis has generated among previously unrelated economic and financial entities render inadequate the foresight.
and analyze the corresponding generalized nonextensive statistical mechanics. Important physical constraints, like concavity or Lesche stability, are also considered, since they are indispensable in order to have physically satisfactory entropies.

It should be noticed that the meaning of “universality” in our context presents some similarities with the usual meaning, as referred to critical phenomena. In any case, we will not be concerned with critical phenomena in this work.

The physical need for such a family of new entropies is evident, for instance, from the physical example of a system of \( N \) interacting over-damped particles, recently published in Ref. [19]. The entropy for this system, which is neither a Boltzmann nor a Tsallis entropy, falls in the class of universal group entropy and can be easily studied in our formalism.

However, the group entropies are not exclusively designed for physical purposes. Indeed, due to the ubiquitous role of the notion of entropy in modern science [20], the possible applications of the proposed construction are manifold. For instance, another nontrivial application of our entropic functionals emerges in the context of information theory. We will show that these entropies provide a class of information measures, including the Shannon information measure as a particular case. Also, we propose a generalization of the Kullback-Leibler divergence, that allows us to define a new set of statistical tests measuring the difference between two given probability distributions.

A further result is the connection we establish between statistical mechanics and number theory. This fascinating topic has an intriguing history, dating back to the 70s with the works of Montgomery and Odlyzko, relating the Gaussian unitary ensemble with the zeros of the Riemann zeta function \( \xi(s) \) [21]. In Ref. [22], it has been shown that the Riemann zeta function is the canonical partition function of a free bosonic gas. Also, recently the zeros of \( \xi(s) \) have been related to Landau levels [23] for a charged particle on a planar surface in an electric potential and uniform magnetic field. A quantum mechanical model whose spectrum is the sequence of prime numbers has been proposed in [24].

Our construction is quite different from other number-theoretical approaches to statistical mechanics. Under appropriate hypotheses, with a class of universality we can associate a zeta function, constructed by using the same realization of the Lazard group used to define the corresponding entropy. The family of zeta functions considered here has been introduced in [16] and further studied in [17]. The first nontrivial case is the Tsallis class, which corresponds to the celebrated Riemann zeta function.

We mention that other generalizations of nonextensive statistical mechanics are known in the literature, based on different kinds of deformations of logarithmic and exponential functions (see, for instance, [25,26] and chapter 6 of Ref. [7] and references therein). The present one also differs from the superstatistics scenario proposed in [27]. Our approach aims to unravel both the group- and number-theoretical content of the notion of entropy. Also, it is constructive: all the entropic functionals are obtained in an explicit form.

Work is in progress on a quantum version of the theory developed here [28]. In particular, group entropies emerge as natural measures in the Hilbert space of the states of a multipartite system as entanglement measures [29].

The paper is organized as follows. In Sec. II, we discuss a family of logarithm-type functions, obtained in the context of finite difference calculus, along with the formulation of G. C. Rota. In Sec. III, the notion of group entropy is introduced, based on the previous construction. In Sec. IV, the universality classes related to group entropies are constructed; their thermodynamic properties are studied in Sec. V. In Sec. VI, we develop a measure-theoretic setting for interpreting the group entropies as information measures. In Sec. VII, the entropies are associated with zeta functions. Some open problems are discussed in Sec. VIII.

II. DIFFERENCE OPERATORS AND LOGARITHMIC FUNCTIONS

In order to construct generalized thermostatistics and the related entropic functionals, we start by introducing a family of logarithmic functions obtained from suitable representations of certain finite-difference operators. For the sake of clarity and to offer a self-contained exposition, some basic concepts are reviewed. The mathematical apparatus is kept to a minimum.

Let \( \mathcal{H} \) a space of sufficiently regular functions of a real variable \( x \); a possible choice is, for instance, a Banach algebra of functions. Let us denote by \( T \) the shift operator whose action on a function \( f \in \mathcal{H} \) is given by \( T f(x) = f(x + \sigma) \). Here \( \sigma \) is a real parameter, whose absolute value can be interpreted as the width of a regular, equally spaced lattice of points \( \mathcal{C} \). We essentially adopt the formalism of [18].

In order to define the class of entropic functionals of interest for this work, we will consider operators expressed as finite Laurent series in shift operators [30]:

\[
\Delta_x = \frac{1}{\sigma} \sum_{n=1}^{m} k_n T^n, \quad l,m \in \mathbb{Z}, \quad l < m, \quad m - l = r, \quad (1)
\]

where \( \{k_n\} \) are real constants such that

\[
\sum_{n=0}^{m} k_n = 0, \quad \sum_{n=0}^{m} n k_n = c, \quad (2)
\]

and \( k_0 \neq 0 \). We choose \( c = 1 \), to reproduce the derivative \( \mathcal{D} \) in the continuum limit, when the lattice spacing \( \sigma \) goes to zero.

**Definition 1.** A difference operator of the form (1), which satisfies equations (2), is said to be a delta operator of order \( r \), if it approximates the continuous derivative up to terms of order \( \sigma^r \).

As Eq. (1) involves \( m - l + 1 \) constants \( k_n \) subject to the two conditions (2), we can fix all constants \( k_n \) by choosing \( m - l - 1 \) further conditions. A possible choice is, for instance, to set

\[
\sum_{n=1}^{m} |n|^{\ell} k_n = K_\ell, \quad \ell = 2, 3, \ldots, m - l, \quad (3)
\]

with \( K_\ell \) suitable real numbers.

The main idea underlying our construction is to represent delta operators in logarithmic form, in terms of a suitable function.
**Definition 2.** We call logarithmic representation of the delta operator (1) the correspondence $T \leftrightarrow x^\sigma$ that defines an isomorphism $\mathcal{T}$ between the space of shift-invariant operators and the space of functions $f \in H$.

The main definition of the section is the following:

**Definition 3.** We call the generalized logarithm the function
\[
\log_G(x) = \frac{1}{\sigma} \sum_{n=0}^m k_n x^{\sigma n}, \quad l, m \in \mathbb{Z}, \quad l < m,
\]
\[
m - l = r, \quad x > 0,
\]
with the constraints (2), (3) [i.e., the image of the operator (1) under the isomorphism $\mathcal{T}$].

The following lemma motivates the choice of the name "logarithm" for function (4).

**Lemma 1.** The following property holds:
\[
\lim_{\sigma \to 0} \log_G(x) = \ln x,
\]
under the conditions (2), (3).

**Proof.** The constraints (2), (3) ensure that, in the limit $\sigma \to 0$, the discrete derivatives of the family (1) tend to the continuous derivative $\partial_x$. If we put $x = e^t$, it implies that the function $\frac{1}{\sigma} \sum_{n=0}^m k_n x^{\sigma n} \equiv \frac{1}{\sigma} \sum_{n=0}^m k_n e^{\sigma n}$ tends to $t = \ln x$ when $\sigma \to 0$.

Notice also that the first of the two conditions (2) implies that
\[
\log_G(1) = 0.
\]

Several examples of interesting logarithmic functions can be obtained by using the logarithmic representation. For instance, if we use the entropic parameter $q$, via the identification $\sigma = 1 - q$, the Tsallis logarithm \[4,7\] corresponds to the discrete derivative of order $r = 1 \Delta^+ = \frac{T - 1}{\sigma}$:
\[
\log_T(x) = \frac{x^{1-q} - 1}{1 - q}.
\]

The Kaniadakis logarithm \[31\] is the indicator of the order two $\Delta^2 = \frac{T - T^{-1}}{2\sigma}$:
\[
\log_K(x) = \frac{x^\sigma - x^{-\sigma}}{2\kappa}
\]
(we put here $\sigma = \kappa$ in accordance with the standard notation).

**III. GROUP ENTROPIES**

Motivated by the previous construction, we propose one of the central notions of this work.

**Definition 4.** Consider a discrete probability distribution $\{p_i\}_{i=1,...,W}$, normalized as
\[
\sum_{i=1}^W p_i = 1.
\]
We call group entropy the functional
\[
S_G(p) := k \sum_{i=1}^W p_i \log_G \left( \frac{1}{p_i} \right),
\]
where $\log_G$ denotes the generalized logarithm (4) with the constraints (2), (3), and $k \in \mathbb{R}^+$. In physical contexts, we can typically identify $k$ with the Boltzmann constant $k_B$; otherwise, as in information theory, we tacitly assume $k = 1$. The reason for the denomination of group entropy for $S_G$ comes from its connection with the universal formal group: it will determine the corresponding correlations for the class of physical systems under examination. As will be explained below, the freedom in the determination of the set of constants $k_n$, left by the conditions (2) and (3), can be used to impose the requirement of concavity, Lesche stability \[32\], etc., for our group entropies.

The entropies previously defined belong to the class of trace form entropies. This class is very general, but does not include other functional forms like Renyi’s entropy \[33\], which is also interesting in several applications. The classical Boltzmann-Gibbs entropy
\[
S_B = -k \sum_{i=1}^W p_i \ln p_i
\]
is obtained from (10) in the limit $\sigma \to 0$. The Tsallis and the Kaniadakis one correspond to the choices (7) and (8). Let us now construct some new examples. Consider, for instance, the difference operators
\[
\Delta_{II} = \frac{T - 2T^{-1} + T^{-2}}{\sigma},
\]
\[
\Delta_{IV} = \frac{T^2 - 3T + \frac{3}{2} T^{-1} - T^{-2}}{\sigma},
\]
\[
\Delta_V = \frac{T^3 - 2T^2 + 2T - 2 T^{-1} + T^{-2}}{\sigma - \sigma},
\]
and so on. The corresponding logarithms are
\[
\log_{G_{II}}(x) = \frac{1}{\sigma}(x^\sigma - 2x^{-\sigma} + x^{-2\sigma}),
\]
\[
\log_{G_{IV}}(x) = \frac{1}{\sigma}(x^{2\sigma} - 3x^\sigma + \frac{3}{2} x^{-\sigma} - x^{-2\sigma} + x^{-2\sigma}),
\]
\[
\log_{G_V}(x) = \frac{1}{\sigma}(x^{3\sigma} - 2x^{2\sigma} + 2x^\sigma - 2x^{-\sigma} + x^{-2\sigma}).
\]

Consequently, we introduce the entropies
\[
S_{G_{II}}(p) := \frac{k}{\sigma} \sum_{i=1}^N p_i \left(p_i^{2\sigma} - 2p_i^\sigma + p_i^{-\sigma}\right),
\]
\[
S_{G_{IV}}(p) := \frac{k}{\sigma} \sum_{i=1}^N p_i \left(p_i^{3\sigma} - \frac{3}{2} p_i^{-\sigma} + \frac{3}{2} p_i^{-2\sigma} - p_i^{-3\sigma}\right),
\]
\[
S_{G_V}(p) := \frac{k}{\sigma} \sum_{i=1}^N p_i \left(p_i^{3\sigma} - 2p_i^{2\sigma} + 2p_i^{-\sigma} - 2p_i^{-2\sigma} - 2p_i^{-3\sigma}\right),
\]
and so on.

Here the roman subindices are used in order to distinguish the logarithms and the associated entropies according to the order of the discrete operator they come from. The entropic forms (12)–(14) at the best of our knowledge are introduced herein. It would be desirable to produce an axiomatic formulation of the notion of group entropy along the lines of the
Shannon-Khinchin-Abe approach [34,35]. Observe that group entropies can be easily constructed in order to fulfill the first three of the classical Khinchin axioms [36]. Indeed, for a suitable choice of the set of coefficients \( \{ \kappa_i \} \) and at least in a suitable range of values of \( \sigma \), the general entropy (10) (i) is continuous; (ii) reaches its maximum for the equiprobability distribution \( p_i = 1/W, i = 1, \ldots, W \); and (iii) satisfies the condition \( S_G(p_1, p_2, \ldots, p_W, 0) = S_G(p_1, p_2, \ldots, p_W) \), which amounts to saying that the addition of an event of zero probability does not affect the value of the entropy. It should be noticed that, although the entropy (10) is certainly continuous in the case of a finite number of microstates \( W \), in general it is no longer continuous for \( W = \infty \), when the Boltzmann-Gibbs entropy is not continuous either [37]. The fourth Khinchin axiom, regarding additivity, obviously does not hold in the present nonextensive context. Precisely, in our approach, the formula (20) below, concerning the “composability” of a system, is no longer an axiom but a property, coming naturally from our group-theoretical construction.

Along these lines, in the recent work [38], a classification of trace-form entropies has been recently proposed. By relaxing the fourth Shannon-Kinchin axiom, universality classes of entropies are introduced and an explicit expression of them in terms of the incomplete Gamma function (Gamma entropies) is provided.

Notice that there is a simple alternative way to derive the proposed entropies and group theory will be clarified. In the following section, the intimate connection between the trace-form entropies will be derived.

In the next section, the intimate connection between the proposed entropies and group theory will be clarified. In particular, the correlation laws underlying the proposed group entropies will be derived.

IV. UNIVERSALITY CLASSES AND CORRELATION LAWS

A. Main definitions

We wish to propose a classification of statistical systems in terms of universality classes related to group entropies. Given a commutative ring \( R \) with identity and the ring \( R \{ x_1, x_2, \ldots \} \) of a formal power series in the variables \( x_1, x_2, \ldots \), with coefficients in \( R \), a commutative one-dimensional formal group over \( R \) is a formal power series \( \Phi(x,y) \in R \{ x, y \} \) such that [9]

\[
\begin{align*}
(1) \quad \Phi(x,0) &= \Phi(0,x) = x, \\
(2) \quad \Phi(\Phi(x,y),z) &= \Phi(x,\Phi(y,z)).
\end{align*}
\]

When \( \Phi(x,y) = \Phi(y,x) \), the formal group is said to be commutative. The existence of an inverse formal series \( \psi(x) \in R \{ x \} \) such that \( \Phi(x, \psi(x)) = 0 \) follows from the previous definition.

Perhaps the most general definition of trace form entropy comes from the theory of formal groups.

Definition 5. Consider the formal power series over the polynomial ring \( \mathbb{Q}[c_1, c_2, \ldots] \) defined by

\[
G(t) = \sum_{i=0}^{\infty} c_i t^{i+1},
\]

with \( c_0 = 1 \), called the formal group exponential. Let \( F = G^{-1} \) be the compositional inverse of (17):

\[
F(s) = \sum_{i=0}^{\infty} s_i t^{i+1},
\]

so that \( G(F(t)) = t \). We have \( c_1 = 1, c_2 = -c_1, c_2 = \sum_{i} c_i^2 - c_2, \ldots \). The Lazard universal formal group law [10] is defined by the formal power series

\[
\Phi(s_1, s_2) = G(F(s_1) + F(s_2)).
\]

Let us analyze the structure of the correlation law among subsystems for the family of entropies (10) (i.e., their “composability” property). Assume that \( S \) is an abstract statistical system (physical, biological, etc.), composed of two independent subsystems \( A \subset S \) and \( B \subset S \). A general property of the group entropy (10) of \( S \) is nonadditivity. Precisely, the following result holds:

Theorem 1. The group entropy (10) satisfies the following nonadditive property:

\[
S_G(A + B) = \Phi(S_G(A), S_G(B)).
\]

Proof. It is easy to prove that Eq. (20) is true for the group entropy \( S_G \) if and only if it holds for the group logarithm \( \log_G(x) \) as well. If we use the exponential representation \( x \leftrightarrow e^x \), we obtain that \( \log_G(e^x) \equiv G(t) \). Since \( t = G^{-1}(s) =: F(S) \), the relation \( G(t_1 + t_2) = G(F(s_1) + F(s_2)) \) holds, which is the thesis.

As a natural consequence of the previous result, a huge class of deformed algebraic structures and calculi closely connected both with the classical one (i.e., the algebra of real numbers) and the standard \( q \) calculus can be derived from formal group laws.

Definition 6. The generalized sum of two numbers \( x \) and \( y \) is given by the universal formal group law (19)

\[
x \oplus_q y = \Phi(x, y).
\]

Here \( q \) is the set of parameters appearing in (18).

Remark 1. Definition 4 of Group Entropy can be easily generalized to the case when the group exponential is indeed not simply a function of \( e^x \) (as in Tsallis’ case and the other ones considered above), but a generic formal series of type (17). In this case we will talk about the universal group entropy. In the following, a physical example when this more general situation is contemplated is discussed.

We propose here a definition of universality classes for statistical systems. Essentially, it allows us to identify systems sharing the same correlation law among their subsystems.
Definition 7. A universality class of statistical systems is a set of systems that satisfy the following properties:
(i) The correlation law between two independent subsystems of a given system is expressed by the formal group law (19).
(ii) The associated thermostatistics is governed by the entropy constructed starting from the corresponding group logarithm and exponential.

As a matter of fact, once $F$ and $G$ are known, the entropy associated can be easily deduced, as shown in the subsequent discussion.

B. Some relevant universality classes

The simplest example is the Boltzmann-Gibbs universality class. The entropic functional for this class is defined in Remark 1, when the group logarithm is not of the form (4) but possesses a general expansion of type (18), is Abe’s entropy. It is related to the following logarithmic functional [39]:

$$\log_A(x) = \frac{x^{\sigma - 1} - x^{\sigma^{-1} - 1}}{\sigma - \sigma^{-1}}. \tag{29}$$

By using the previous considerations, we can easily generalize the previous logarithm by introducing the two-parametric functional

$$\log_{a,b}(x) = \frac{x^a - x^b}{a - b}, \tag{30}$$

which reproduces Abe’s one for $a = \sigma - 1, b = \sigma^{-1} - 1$. It has already been considered in the Borges-Roditi construction of an entropic functional [40]. In this case, the related group exponential, obtained by using the isomorphism $\mathcal{I}$, reads

$$G_A(t) = \frac{e^{at} - e^{bt}}{a - b}. \tag{31}$$

The formal group corresponding to (31), giving the interaction rule for this class, is known in the literature as the Abel formal group, defined by [41]

$$\Phi_A(x, y) = x + y + \beta_1 xy + \sum_{j>1} \beta_j (xy^j - x^j y).$$

The coefficients $\beta_n$ in (32) can be expressed as polynomials in $a$ and $b$ (see Proposition 3.1 of [41]):

$$\beta_n = \frac{(-1)^{n-1}}{n!(n-1)} \prod_{i,j>0} (ia + jb). \tag{33}$$

Concerning the algebraic structure of the theory, observe that we can introduce a new multiplication law, inspired by the previous construction.

Definition 8. Given two real numbers $x$ and $y$, the product $x \otimes y$ is defined by the relation

$$\log_G(x \otimes y) = \log_G x + \log_G y. \tag{34}$$

We recover easily the known cases:
(i) For the case of the Boltzmann class, $x \otimes y = xy$, i.e., the multiplication (34) reduces to the standard pointwise multiplications of real numbers.
(ii) For the Tsallis class, we have [42]

$$x \otimes_q y = [x^{1-q} + y^{1-q} + 1]^{\frac{1}{1-q}}. \tag{35}$$

(iii) For the Kaniadakis class, we obtain [26]

$$x \otimes_k y = \frac{1}{k} \sinh \left[ \frac{1}{k} \arcsinh(kx) \arcsinh(xy) \right]. \tag{36}$$

However, with these composition laws we do not get a priori standard algebraic structures: as has already been noticed for the specific example of the Tsallis class, there is no guarantee that the distributivity property is satisfied.

C. A physical example

Very recently, in [19], a system of $N$ over-damped interacting particles moving in a narrow channel has been studied. The equations for the velocities of the particles are

$$\vec{v}_i = \sum_{j \neq i} J (\vec{r}_i - \vec{r}_j) + \vec{F} (\vec{r}_i) + \eta_i (\vec{r}_i, t), \tag{37}$$

where $\vec{v}_i$ is the velocity of the $i$th particle, $J (\vec{r}_i - \vec{r}_j)$ is a short-range repulsive particle-particle interaction, $\mu$ is the effective viscosity of the medium, $\vec{F} (\vec{r}_i)$ is an external force,
Consider the variational problem of the existence of a stationary distribution $\tilde{\rho}(\epsilon)$. We introduce the functional

$$L = S_G[p] - \alpha \left[ \sum_i p(\epsilon_i) - 1 \right] - \beta \left[ \sum_{i=1}^{W} \epsilon_i p_i(\epsilon_i) - U \right],$$

(43)

where $\alpha$ and $\beta$ are Lagrange multipliers. By imposing the vanishing of the variational derivative of this functional with respect to the distribution $p_i$, we get the stationary solution

$$\tilde{\rho}_i = \frac{E(-\alpha - \beta(\epsilon_i - \tilde{U}))}{Z},$$

(44)

with $Z = \sum_{i=1}^{W} E(-\alpha - \beta(\epsilon_i - \tilde{U}))$ and $E(\cdot)$ is an invertible function. However, only in particular cases (e.g., for the Boltzmann, Tsallis, Kaniadakis, and Borges-Roditi entropies) this function can be identified with the inverse of $\log_G$, according to the analysis of the role of generalized exponentials in thermodynamics performed in [45]. Nevertheless, for generalized entropies constructed as a realization of the universal formal group that no longer belong to the trace-form class, one can contemplate further cases in which $E(\cdot)$ can be set to be the inverse of $\log_G$.

Notice that a closed expression for the inverse of $\log_G$ can be analytically determined in very specific situations, since the inversion of a formal group logarithm would involve the solution of polynomial equations of high degree. Although we will keep the discussion at a formal level, we can, however assume that at least a numerical interpolating solution, in several cases and for specific subsets of the space of parameters, is available.

As usual, the parameter $\alpha$ can be eliminated by means of the constraint $\sum_i p_i(\epsilon_i) \frac{dL}{d\rho_i} |_{p=\tilde{\rho}} = 0$. If we perform the Legendre transform of $\log_G(Z)$, we get the interesting relation

$$\log_G(Z) + \beta U = S_G,$$

(45)

from which we deduce immediately the relation

$$\frac{\partial S_G}{\partial U} = \frac{1}{T},$$

(46)

with $T \equiv 1/\beta$. In this context, $T$ plays the role of a local temperature, a priori a function of space and time. In a nonequilibrium stationary state, the usual definition of temperature, as given by the zeroth law, does not apply. The previous discussion also implies the generalized form of the free energy:

$$\mathcal{F} = U - T S_G.$$

(47)

A natural question is whether the group entropy satisfies an analog of the $H$ theorem. As has been shown in [46], it is possible to prove the validity of the $H$ theorem for a large class of entropic forms. Precisely, assume that the group logarithm (4) is an invertible function and piecewise monotonic. Consider the master equation

$$\frac{dp_i}{dt} = \sum_{j=1}^{W} (A_{ij} p_j - A_{ji} p_i),$$

(48)
where $A_{ij}$ denotes the transition probability per unit time from state $j$ to $i$. If we assume the principle of microscopic reversibility [47], which implies the relation

$$A_{ij} = A_{ji},$$  

(49)

then, under the previous assumptions, we obtain

$$
\frac{dS[p]}{dt} \geq 0.
$$  

(50)

In order for an entropic functional to satisfy a maximum entropy principle, it is sufficient to ascertain that it verifies the following concavity property. Consider two sets $\{p_i\}$ and $\{p'_i\}$ of probabilities taken from the same set of $W$ possibilities. Let $\lambda$ be a real parameter, satisfying $0 < \lambda < 1$. According to [4], define the intermediate probability law

$$p'_i = \lambda p_i + (1 - \lambda)p'_i.$$

(51)

A direct calculation shows that the entropies (12)-(14), in a suitable interval of values of $\sigma$, satisfy the condition

$$S((p'_i)) \geq \lambda S((p_i)) + (1 - \lambda)S((p'_i)).$$

(52)

Another crucial aspect, for thermodynamical applications, is the experimental robustness of the group entropy. It means that, given two distributions $\{p_i\}_{i=1,...,w}$ and $\{p'_i\}_{i=1,...,w}$ whose values are slightly different,

$$\forall \epsilon \exists \delta > 0 \text{ so that } \|p - p'\| < \delta \implies \frac{|S[p] - S[p']|}{S_{\text{max}}} < \epsilon,$$

(53)

where, given a vector $x$, $\|x\|$ denotes the $L_1$ norm $\sum_{i=1}^{w} |x_i|$, and $S_{\text{max}}$ denotes the maximum value of the entropy. This property, also called Lesche stability [32], is ensured under the hypothesis that the group exponential be at least piecewise differentiable. We point out that this requirement is especially important for physical applications (for a recent discussion, see [48]); however, it can be disregarded in other contexts, such as, for instance, information theory.

VI. GROUP ENTROPIES AS GENERALIZED KULLBACK-LEIBLER DIVERGENCES

Following the classical approach by Shannon [34,35], we show here that group entropies can naturally be interpreted as generalized information measures. This entails a proper definition of the functionals of the class (10) on a continuous setting. As an important application, we introduce a generalization of the Kullback-Leibler relative entropy. This provides a very large set of measures of divergence between two different probability distributions, or alternatively of tests allowing the discrimination between two different hypotheses. Here, as in [49], we propose a formulation of our approach in a measure-theoretical setting.

Precisely, let $(X, \Sigma, \mu)$ be a measure space, where as usual $X$ is a set, $\Sigma$ a $\sigma$-algebra over $X$, and $\mu : \Sigma \to \mathbb{R}^+$ a measure. In the following, we will assume that $X, \Sigma, \mu$ is a $\sigma$-finite measure space, and that our probability measures are absolutely continuous with respect to $\mu$. A $\Sigma$-measurable function $p : X \to \mathbb{R}^+$ will be called a probability distribution function (PDF) if

$$
\int_X p \, d\mu = 1.
$$

(54)

The probability measure induced by a PDF $p$ is defined by

$$P(E) = \int_E p(x) \, d\mu(x), \quad \forall E \in \Sigma.$$

(55)

As a consequence of the classical Radon–Nikodym theorem, we will discuss on the same footing information measures coming from PDF’s and from probability measures, since we can identify them up to a $\mu$-null set, with $P \ll \mu$.

We can now propose the main definition of this section.

Definition 9. Given a $\sigma$-finite measure space $(X, \Sigma, \mu)$, the group entropy of a PDF $p$ is given by

$$S_G(p) := \int_X p(x) \log_G \frac{1}{p(x)} \, d\mu(x),$$

(56)

or equivalently in terms of a probability measure $P$ absolutely continuous with respect to $\mu$,

$$S_G(p) := \int_X \log_G \left( \frac{dP}{d\mu} \right)^{-1} \, dP,$$

(57)

provided the integral on the right exists.

Here $\frac{dP}{d\mu}$ denotes the Radon-Nikodym derivative of the measure $P$ with respect to $\mu$. Observe that we put $k = 1$, as customary in information theory. The Shannon entropy of the PDF $p$ is obtained immediately from the previous definition:

$$S(p) = -\int_X p(x) \ln p(x) \, d\mu.$$ 

(58)

In a similar way, we can introduce a generalization of the standard Kullback-Leibler relative entropy (KL) [50], that we call relative group entropy. In information theory, given two different probability distributions $P$ and $Q$, the KL entropy is a measure of the expected number of additional bits required to code samples from $P$ if, instead using a code based on $P$, we use a code based on $Q$.

Definition 10. Let $P$ and $Q$ be two probability measures and let $Q$ be absolutely continuous with respect to $P$. The relative group entropy is defined as

$$D_G(P \parallel Q) = -\int_X \log_G \left( \frac{dQ}{dP} \right)^{-1} \, dP,$$

(59)

or equivalently, if $\mu$ is a measure, as

$$D_G(P \parallel Q) = -\int_X p(x) \log_G \left( \frac{q(x)}{p(x)} \right)^{-1} \, d\mu.$$

(60)

The previous definition implies the useful relation

$$S_G(P) = D_G(P \parallel \mu).$$

(61)

Our definition includes as a special case that proposed by Tsallis in [51]. Also, it differs from that of Yamano [52].

Let us study some properties of the relative group entropy. First, consider a uniform distribution on a compact support of length $W$. We deduce that

$$D_G(p, 1/W) = \log_G(W) - W^\sigma S_G(p);$$

(62)
that is, the RGE essentially measures the departure of a given group entropic functional from its value at equiprobability.

If we impose that $x \log_{G}(x)$ is a convex function, the relative group entropies belong to the class of Csiszár $f$-divergence measures [53]. In this case, the main property of the relative group entropy is a generalization of Gibbs’ inequality. For simplicity, we will state it in the discrete case.

**Theorem 3.** Let $\{p_i\}_{i=1}^W$ and $\{q_i\}_{i=1}^W$ be two sets of probabilities. Assume that $x \log_{G}(x)$ is a convex function. Then the relative group entropy satisfies the inequality

$$D_G(P \parallel Q) = \sum_{i=1}^W p_i \log_{G} \left( \frac{p_i}{q_i} \right) \geq 0,$$

and the equality holds if and only if $p_i = q_i$ for all $i = 1, \ldots, W$.

**Proof.** It suffices to observe that, if the function $\varphi(x) := x \log_{G}(x)$ is convex, we can use the classical Jensen inequality [54]. Indeed, if we have a set of nonnegative real numbers $\{p_1, \ldots, p_W\}$ and $\{q_1, \ldots, q_W\}$, with $\sum p_i = S_1$, $\sum q_i = S_2$, then

$$\sum_{i=1}^W p_i \log_{G} \left( \frac{p_i}{q_i} \right) \geq \frac{S_1}{S_2} \varphi \left( \frac{\sum_{i=1}^W p_i}{\sum_{i=1}^W q_i} \right)$$

Now, if $S_1 = S_2 = 1$, the inequality (63) follows. Due to the condition $\log_{G}(1) = 0$, when $p_i = q_i$ for all $i = 1, \ldots, W$ we have $D_G(P \parallel Q) = 0$.

This result justifies the interpretation of relative group entropies in a measure-theoretic setting, as criteria of information divergence. However, *a priori* we can also construct non Csiszár-type distances; for instance, by relaxing the above convexity condition or, for instance, by constructing generalized Chernoff $\alpha$ distances [55]. Further study is in progress in this direction.

**VII. DIRICHLET ZETA FUNCTIONS ASSOCIATED WITH GROUP ENTROPIES**

In this section, we clarify the number-theoretic content of the previous theory. The main result can be stated as follows: provided some technical hypotheses are satisfied, there exists a Dirichlet zeta function associated with each universality class of a suitable type. In order to make the argument transparent, let us consider first the simplest nontrivial case, which is that of the classical Riemann zeta function.

Clearly, each of the generalized zeta functions (69) is associated to a specific universality class via the corresponding group exponential. Indeed, according to the construction discussed in the previous sections, assigning a group exponential is sufficient for determining completely a universality class.

Observe that the present construction is even more general and can be proposed for logarithms different from that possessing the form (66). In all these cases, under the above regularity hypothesis for $G(t)$, an $L$ function can be constructed, but its representation in terms of a series of the form (69) usually cannot be obtained in a simple way.

We observe that the class of uncorrelated systems does not possesses an associated zeta function: from this point of view, the absence of correlation is trivial. In particular, it would be nice to understand the statistical mechanical interpretation of the zeros of the family of zeta functions described here.

**VIII. CENTRAL LIMIT BEHAVIOR AND OTHER OPEN PROBLEMS**

In this work, we have introduced the notion of group entropy, which emerges naturally as a unifying tool for treating many theoretical aspects of generalized thermostatistics and provides a nontrivial connection with number theory.

A wide spectrum of theoretical perspectives can be explored. Here we suggest some open problems.

A key result in probability theory is the central limit theorem (CLT). In its standard version, it states that the mean
of a large number of independent random variables identically distributed, with finite expectation value and variance, has the normal distribution as an attractor [57]. A central limit theorem for random correlated variables has been established [58, 59]. In the context of the present generalization of statistical mechanics, a natural question arises, which can be stated as follows:

Consider a set of random variables, with any PDF, correlated according to the law (21). Establish whether, under suitable conditions, an analog of the CLT holds for the universality classes described by the group entropy (10). Find the analytic expression of the attractor. An interesting problem is to establish the exact relation between the present approach and the notion of “gamma entropy” recently proposed in [44], that is, under which conditions do the universality classes described by gamma entropies collapse into group entropies and vice versa. In addition, it appears interesting to study the role of general correlation laws in complex networks (especially scale-free networks [60]), in biological complexity [61], in temporal series analysis, and in ecology, in relation with the study of biodiversity [62]. Also, it would be important to deepen the comprehension of the dynamics of financial markets under nontrivial correlations (for a nonextensive approach see, e.g., [63]). An intriguing possibility is to classify economical systems according to the universality class of correlations that govern them.

ACKNOWLEDGMENTS

I am grateful to Professor C. Tsallis for a careful reading of the manuscript, encouragement and many useful discussions. I wish to thank Professor S. Thurner for stimulating discussions at the final stage of this work and, in particular, for drawing to my attention Ref. [44]. I also thank Professor G. Mussardo for interesting discussions. Finally, I wish to thank the referees for useful comments. I wish to thank the Centre of Mathematics for Applications (CMA), University of Oslo, Norway and the IMI-Madrid, Spain for support in the framework of the “NILS mobility project, Abel extraordinary chair”. Part of this work has been carried out during my stay at the CMA. Support from the Spanish Ministerio de Ciencia e Innovación, research project FIS2008-00200 is gratefully acknowledged.


[http://tsallis.cat.cbpf.br/TcUMOC.pdf].


