

Beyond the Shannon-Khinchin Formulation: The Composability Axiom and the Universal Group Entropy

Piergiulio Tempesta^{a,b}

^a*Departamento de Física Teórica II (Métodos Matemáticos de la física), Facultad de
Físicas, Universidad Complutense de Madrid, 28040 – Madrid, Spain*

and

^b*Instituto de Ciencias Matemáticas, C/ Nicolás Cabrera, No 13–15, 28049 Madrid, Spain*

Abstract

The notion of entropy is ubiquitous both in natural and social sciences. In the last two decades, a considerable effort has been devoted to the study of new entropic forms, which generalize the standard Boltzmann-Gibbs (BG) entropy and are widely applicable in thermodynamics, quantum mechanics and information theory. In [25], by extending previous ideas of Shannon [40, 41], Khinchin proposed a characterization of the BG entropy, based on four requirements, nowadays known as the Shannon-Khinchin (SK) axioms.

The purpose of this paper is twofold. First, we show that there exists an *intrinsic group-theoretical structure* behind the notion of entropy. It comes from the requirement of composability of an entropy with respect to the union of two statistically independent systems, that we propose in an axiomatic formulation. Second, we show that there exists a simple universal family of trace-form entropies. This class contains many well known examples of entropies and infinitely many new ones, a priori multi-parametric. Due to its specific relation with Lazard’s universal formal group of algebraic topology, the new general entropy introduced in this work will be called the universal-group entropy. A new example of multi-parametric entropy is explicitly constructed.

Email address: p.tempesta@fis.ucm.es, piergiulio.tempesta@icmat.es (Piergiulio Tempesta)

Contents

1	Introduction	2
2	The group-theoretical content of the notion of entropy	5
2.1	Strong and weak composability	5
2.2	On the relation between the SK axioms, composability and admissible trace-form entropies	7
3	Formal group laws	7
3.1	The Lazard formal group	8
4	The universal-group entropy.	9
4.1	The universal-group entropy	9
5	Main properties of S_U entropy	10
5.1	Composability	10
5.2	Extensivity	12
5.3	Relation with other entropies	12
6	A new three-parametric group entropy: the $S_{q,\alpha,\beta}$ entropy	18
7	Distribution functions and thermodynamic properties	19
8	On the asymptotic behaviour of generalized entropies	20
9	Open problems and future perspectives	21
Appendix A	The Shannon-Khinchin axioms	21

1. Introduction

Entropy is a fundamental notion, at the heart of modern science. In the second half of the twentieth century, its range of applicability has been extended from the traditional context of classical thermodynamics to new areas such social sciences, economics, biology, quantum information theory, linguistics, etc. More recently, the role of entropy in the theory of complex systems has been actively investigated. From one side, several studies were devoted to axiomatic formulations, aiming at clarifying the foundational aspects of the notion of entropy. From the other side, many researchers pursued the idea of generalizing the classical Boltzmann-Gibbs statistical mechanics. Consequently, a plethora of new entropic forms, designed for extending the applicability of BG entropy to new contexts, was introduced.

The first research line was started by the seminal works by Shannon [40, 41] and Khinchin [25]. A set of axioms, nowadays called the SK axioms, characterizing uniquely the BG entropy, was introduced (we shall make reference to the formulation of the axioms reported in Appendix A). The axioms (SK1)-(SK3)

represent natural requirements (continuity, maximum principle, independence from zero probability events), that should be satisfied by any functional playing the role of an entropy. Instead, the axiom (SK4) simply characterizes the behaviour of an entropy with respect to the composition of two subsystems, which reduces to additivity in the case of statistical independence of the subsystems.

For long time, additivity was interpreted as the property that ensures *extensivity*, i.e. the linear dependence of entropy on the number of particles of a system. According to Clausius, extensivity is crucial for an entropy to be thermodynamically admissible. Surprisingly, the two concepts are completely independent: additivity does not imply, nor is implied by extensivity. In addition, no entropy, irrespectively of being additive or nonadditive, can be extensive in any dynamical regime. For instance, if $W(N)$ is the total number of states of a complex system as a function of the number of its particles N , it turns out that a (sufficient) condition for the BG entropy to be extensive is that $W(N) \sim k^N$, with $k \in \mathbb{R}_+$; however, if $W(N) \sim N^k$, it is not.

The second research line, i.e. the study of generalized entropies and thermostatics, in which the additivity axiom is explicitly violated, has become an extremely active research area in the last three decades. Since the work of Tsallis [50], many new entropic functionals have been proposed in the literature (see e.g. [1], [4], [6], [9], [15], [17, 18], [31], [32], [39], [45], [53]). From the point of view of statistical mechanics, they may generalize the BG entropy in weakly chaotic regimes, when the ergodicity hypothesis is violated and the correlation functions exhibit a non-exponential decay, typically a power-law one. [49]. In particular, Tsallis entropy, which is nonadditive, is extensive for special values of the parameter q in regimes where the BG entropy is not [54].

Another source of nonadditive entropies is Information Theory. In this context, generalized entropies can provide more refined versions of Kullback-Leibler-type divergences [26], useful for constructing comparative tests of sets of data.

In the study of entanglement, nonadditive entropies arise as necessary alternatives to the von Neumann entropy [56], [15]. As usual, a system composed by two subsystems A and B is said to be entangled or separated if its density operator ρ cannot be written as a convex combination of uncorrelated densities, i.e. $\rho \neq \sum_k q_k \rho_k^A \otimes \rho_k^B$. As shown in [21], the direct maximization of the von Neumann entropy $S = -\text{Tr} \rho \ln \rho$ does not avoid the detection of fake entanglement even for the case of two spin $\frac{1}{2}$ systems. The need for generalized nonadditive entropies in order to design efficient criteria for separability has been advocated in [15]. Recently, the relevance of generalized entropies in processes with quantum memory has been recognized [7].

In the present analysis of the mathematical foundations of the concept of entropy, we wish to point out the centrality of the notion of *composability*. We shall say that an entropy is composable if the following requirements are satisfied. First, given two statistically independent systems A and B , the entropy of the composed system $A \cup B$ depends on the entropies of the two systems $S(A)$ and $S(B)$ *only* (apart possibly a set of parameters). This is the original formulation of the concept, as in [52]. In addition, we require further properties, as the symmetry of a given entropy in the composition of the systems

A and B , the stability of total entropy if one of the two systems is in a state of zero entropy, and the associativity of the composition of three independent systems (see axioms (C1)-(C4) below). With these assumptions, the composability property, in this new, broader sense, is equivalent to the existence of a *group-theoretical structure underlying the notion of entropy*, which guarantees the physical plausibility of the composition process.

All the previous properties ensure that a given entropy can be expressed just in terms of *macroscopic configurations* of a system, without the need for a microscopic description of the associated dynamics. An entropy should be *coarse-grained* [16]; if not so, the concept of entropy would be simply empty. Even the second law of thermodynamics loses any meaning if not referred to the evolution of macroscopic subsystems: the entropy would stay invariant if defined on microscopic configurations.

From the previous discussion, it emerges that keeping only with the first three SK axioms is a too weak requirement from a thermodynamical perspective: the fourth axiom cannot be simply dropped away. This situation closely reminds the role of Euclids' fifth postulate of geometry. The replacement (not the mere exclusion) of this postulate with different axiomatic formulations paved the way to non-Euclidean geometries.

Our proposal is *to assume the property of composability in its complete group-theoretical formulation as the generalized form of the fourth axiom*. This assumption is not at all obvious. From this point of view, there are several different perspectives.

a) If we require composability in a strict sense, i.e. for any possible choice of the probability distributions of the given systems, it emerges that the Boltzmann-Gibbs entropy and the Tsallis entropy are the only known cases of trace-form composable entropies (at the best of our knowledge).

b) At the same time, the notion of composability can be formulated in a weak sense. Indeed, we can impose that the composability axiom be satisfied at least on the *uniform distribution*. This requirement, in thermodynamics, is a fundamental one. Indeed, the uniform distribution emerges when one deals with isolated physical systems at the equilibrium (microcanonical ensemble), or in contact with a thermostat at very high temperature (canonical ensemble). For instance, this is the physical situation occurring when considering high-temperature astrophysical objects. It turns out that a large set of entropies introduced in the last decades are weakly composable, although not all of them. Again, our formulation unravels the intrinsic group-theoretical content of the concept of entropy: in all cases, the composability requirement amounts to the existence of a *group law* for the composition process, defined either over the whole set of probability distributions (strict composability) or just over the uniform one (weak composability).

In this work, we present two main results. First, we encode composability into a general axiomatic formulation, required for an entropy to be considered admissible.

Second, we introduce a very general trace-form entropic form. We shall call it the *universal-group entropy*, due to its relation with Lazard's universal formal

group. This entropy offers infinitely new cases of trace-form weakly composable entropies, therefore satisfying the first and third SK axioms. Also, all known entropies are directly related with our new entropic form.

The universal group entropy depends a priori on a (possibly) infinite number of independent parameters. Special attention will be paid to the recently introduced $S_{c,d}$ entropy [17] and to group entropies [45], respectively. We will show that both cases are related to the universal-group theoretical framework in a very direct way. Besides, multi-parametric entropies can be introduced, not related to any of the known entropies. As an interesting example, the three-parametric $S_{\alpha,\beta,q}$ entropy is proposed.

The paper is organized as follows. In Section 2, the group structure behind the notion of entropy is introduced. In Section 3, the theory of formal groups, and in particular that of the universal formal group, is sketched. In Section 4, the universal-group entropy is defined and its main properties are studied in Section 5. In Section 6, a new example of three-parametric entropy is proposed. In Section 7, the Legendre-type structure and some thermodynamical aspects of the theory are discussed. In Section 8, the concept of thermodynamic limit of an entropy is discussed. In Section 9, some open problems and possible related research lines are briefly proposed. In the Appendix A, the formulation of the SK axioms adopted in the paper is reported.

We expect that the present general formulation of the notion of entropy, apart its intrinsic theoretical interest, can have further use both in classical and quantum information theory and beyond.

2. The group-theoretical content of the notion of entropy

The purpose of this section is to establish the existence of a group-theoretical structure underlying the concept of (generalized) entropy. To this aim, we propose first a new formulation of the notion of composability. We shall assume that the entropies we consider are sufficiently regular functions $S(p_1, \dots, p_W)$ defined for any integer $W \geq 1$ in the set \mathcal{P}_W of all discrete probability distributions with W entries, and taking values in $\mathbb{R}^+ \cup \{0\}$.

2.1. Strong and weak composability

Definition 1. *An entropy S is strongly (or strictly) composable if there exists a smooth function of two real variables $\Phi(x, y)$ such that*

$$(C1) \quad S(A \cup B) = \Phi(S(A), S(B); \{\eta\}) \quad (1)$$

where $A \subset X$ and $B \subset X$ are two statistically independent systems, each defined in terms of an arbitrary probability distribution $\{p_i\}_{i=1}^W$, and $\{\eta\}$ is a possible set of real continuous parameters. In addition, $\Phi(x, y)$ satisfies the following properties:

$$(C2) \text{ Symmetry:} \quad \Phi(x, y) = \Phi(y, x); \quad (2)$$

(C3) *Associativity*:

$$\Phi(x, \Phi(y, z)) = \Phi(\Phi(x, y), z); \quad (3)$$

(C4) *Null-composability*:

$$\Phi(x, 0) = x. \quad (4)$$

The symmetry property is an obvious requirement. The null-composability ensures that the composition of a system with another system in a state of zero entropy cannot affect thermodynamics. The associativity property is a new, essential point: it guarantees the composability of more than two systems.

Remark 1. Notice that, assuming $\Phi(x, y) = x + y + \sum_{kl} c_{kl} x^k y^l$, the existence of a power series $\varphi(x)$ such that $\Phi(x, \varphi(x)) = 0$, i.e. playing the role of an inverse, is a direct consequence of the previous axioms (this observation is also valid when working with formal power series only). The previous requirements amount to say that, from a mathematical point of view, $\Phi(x, y)$ defines a formal *group law* over the reals. For instance, in the case of the Boltzmann entropy it is nothing but the additive group over \mathbb{R} .

However, when we restrict to values $x, y \in \mathbb{R}^+ \cup \{0\}$, as is the case when $\Phi(x, y)$ is computed over standard entropies, we do not have an inverse.

Remark 2. Condition (1) can be formulated in more general terms. Consider the case of a composite system $A \cup B$ arising from two systems not statistically independent, with a conditional probability distribution $p_{ij}(B | A) := p_{ij}(A \cup B)/p_i(A)$. Here $p_{ij}(A \cup B)$, $i = 1, \dots, W_A$, $j = 1, \dots, W_B$ denotes the joint probability distribution for the composite system $A \cup B$, and $p_i(A)$ is the marginal probability distribution $p_i(A) = \sum_{j=1}^{W_B} p_{ij}(A, B)$. In this context we postulate the relation

$$S(A \cup B) = \Phi(S(A), S(B | A); \{\eta\}), \quad (5)$$

where $S(B | A)$ denotes the conditional entropy associated with the conditional distribution $p_{ij}(B | A)$. Equation (5) reduces to the relation (1) in the case of statistically independent systems. The relation (5) generalizes the axiom (SK4).

We can also propose a *weak formulation* of composability.

Definition 2. *We shall say that an entropy is weakly composable if the properties (C1)–(C3) of Definition 1 are satisfied at least when the probability distributions of the two statistically independent systems A and B are both uniform, and property (C4) holds in general.*

This weak formulation is in fact satisfied by infinitely many generalized entropies, as we shall prove. For instance, the function $\Phi(x, y)$, and consequently its group structure, can be typically constructed starting from generalized logarithms. This second point of view has been first proposed in [45] for a class of logarithms coming from difference operators. Again, the weak composability requirement implies the existence of a group law.

2.2. *On the relation between the SK axioms, composability and admissible trace-form entropies*

The SK axioms and the composability axiom are strictly related. First observe that, for both formulations of composability, the continuity of a given entropy S with respect to its arguments, as required by the axiom (SK1), implies the continuity of $\Phi(x, y)$ at least for $x, y \in \mathbb{R}^+ \cup \{0\}$.

In the strong formulation of the notion of composability, the property (5), valid for the Boltzmann and Tsallis entropies, is equivalent to the axiom (SK4). However, in both formulations, we still have to impose the further requirement of *strict concavity* (to ensure that axiom (SK2) is fulfilled).

Motivated by the previous discussion, we analyze the requisites for an entropy to be considered *admissible*, i.e. relevant both from a physical and information-theoretical point of view. In the recent literature, usually admissible entropies are considered to be those satisfying just the axioms (SK1)–(SK3). Our point of view is therefore more demanding: *A necessary condition for an entropy S to be admissible is that it satisfies the axioms (SK1)–(SK3) and is (at least) weakly composable.*

Indeed, there exist entropies which do satisfy the axioms (SK1)–(SK3), but are not even weakly composable. Needless to say, the strong composability property is much more suitable for thermodynamical purposes than its weak formulation.

The general problem of classifying the admissible entropies will be discussed in the forthcoming sections.

3. Formal group laws

The theory of formal groups [8] offers the natural language for formulating our approach to the theory of generalized entropies. Formal groups have been intensively investigated in the last decades, especially for their prominent role in fields as algebraic topology, the theory of elliptic curves, and arithmetic number theory [20], [44]. Here we briefly recall only some salient aspects, necessary in the subsequent discussion.

Let R be a commutative ring with identity, and $R\{x_1, x_2, \dots\}$ be the ring of formal power series in the variables x_1, x_2, \dots with coefficients in R . A *commutative one-dimensional formal group law* over R [8] is a formal power series in two variables $\Psi(x, y) \in R\{x, y\}$ of the form $\Psi(x, y) = x + y + \text{terms of higher degree}$, such that

$$\begin{aligned} i) \quad & \Psi(x, 0) = \Psi(0, x) = x \\ ii) \quad & \Psi(\Psi(x, y), z) = \Psi(x, \Psi(y, z)). \end{aligned}$$

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

The simplest examples are the additive formal group law $\Psi(x, t) = x + y$ and the multiplicative one $\Psi(x, y) = x + y + xy$.

The previous definition can be naturally extended to the case of n -dimensional formal group laws.

The relevance of formal groups relies first of all on their close connection with group theory. Precisely, a formal group law $\Psi(x, y)$ defines a functor $F : \mathbf{Alg}_R \rightarrow \mathbf{Group}$, where \mathbf{Alg}_R denotes the category of commutative unitary algebras over R and \mathbf{Group} denotes the category of groups [20]. The functor F is by definition the formal group (sometimes called the formal group scheme) associated to the formal group law Ψ .

As is well known, over a field of characteristic zero, there exists an equivalence of categories between Lie algebras and formal groups [38].

Any n -dimensional formal group law defines a n -dimensional Lie algebra over the same ring R by means of the identification

$$[x, y] = \Psi_2(x, y) - \Psi_2(y, x), \quad (6)$$

where $\Psi_2(x, y)$ denotes the quadratic part of the formal group law $\Psi(x, y)$. This equivalence of categories is no longer true in a field of characteristic $p \neq 0$.

3.1. The Lazard formal group

The main algebraic structure we need is provided by the following construction. Let $B = \mathbb{Z}[b_1, b_2, \dots]$ be the ring of integral polynomials in infinitely many variables. We shall consider the series (formal group logarithm)

$$F(s) = \sum_{i=0}^{\infty} b_i \frac{s^{i+1}}{i+1}, \quad (7)$$

with $b_0 = 1$. Let $G(t)$ be its compositional inverse (the formal group exponential):

$$G(t) = \sum_{k=0}^{\infty} a_k \frac{t^{k+1}}{k+1} \quad (8)$$

so that $F(G(t)) = t$. We have $a_0 = 1, a_1 = -b_1, a_2 = \frac{3}{2}b_1^2 - b_2, \dots$. The Lazard formal group law [20] is defined by the formal power series

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

It has a great relevance in many branches of mathematics, as algebraic topology, cobordism theory, etc. [20], [10], [11], [35].

The coefficients of the power series $G(F(s_1) + F(s_2))$ lie in the ring $B \otimes \mathbb{Q}$ and generate over \mathbb{Z} a subring $A \subset B \otimes \mathbb{Q}$, called the Lazard ring L .

The following important results, due to Lazard, hold. First, for any commutative one-dimensional formal group law over any ring R , there exists a unique homomorphism $L \rightarrow R$ under which the Lazard group law is mapped into the given group law (the so called universal property of the Lazard group).

At the same time, for any commutative one-dimensional formal group law $\Psi(x, y)$ over any ring R , there exists a series $\psi(x) \in R[[x]] \otimes \mathbb{Q}$ such that

$$\psi(x) = x + O(x^2), \quad \text{and} \quad \Psi(x, y) = \psi^{-1}(\psi(x) + \psi(y)) \in R[[x, y]] \otimes \mathbb{Q}.$$

In [11] and [33], it has been pointed out that the Lazard universal group law is fundamental in the discussion of both the classical and modern theory of unitary cobordisms. A beautiful connection with combinatorial Hopf algebras and Rota's umbral calculus was established in [13]. A combinatorial approach has been proposed in [5].

In the papers [42], [44] a connection between the Lazard universal formal group and the theory of Dirichlet series has been established. In this context, the *Universal Bernoulli polynomials* were also introduced and their remarkable properties studied in [46] (see also [43]). In [30], the universal Bernoulli polynomials have been related to the theory of hyperfunctions of one variable by means of an extension of the classical Lipschitz summation formula to negative powers.

4. The universal-group entropy.

We are now able to construct a very general family of *trace-form entropies* with relevant thermodynamical properties.

4.1. The universal-group entropy

Definition 3. Let $\{p_i\}_{i=1, \dots, W}$, $W \geq 1$, with $\sum_{i=1}^W p_i = 1$, be a discrete probability distribution. Let

$$G(t) = \sum_{k=0}^{\infty} a_k \frac{t^{k+1}}{k+1} \quad (9)$$

be a real analytic function, where $\{a_k\}_{k \in \mathbb{N}}$ is a sequence of parameters, with $a_0 \neq 0$, such that the function $S_U : \mathcal{P}_W \rightarrow \mathbb{R}^+ \cup \{0\}$, defined by

$$S_U(p_1, \dots, p_W) := k_B \sum_{i=1}^W p_i G\left(\ln \frac{1}{p_i}\right), \quad (10)$$

is a concave one. This function will be called the *universal-group entropy*.

The name of the entropy (10) is obviously reminiscent of its direct connection with Lazard's construction of a universal formal group law. Indeed, $G(t)$ is a group exponential. The function $G(t)$, or equivalently the real sequence $\{a_k\}_{k \in \mathbb{N}}$, encodes all the main features of the entropy (10). A simple way to construct an important subclass of concave entropies is provided by the following observation.

Remark 3. The condition

$$a_k > (k+1)a_{k+1} \quad \forall k \in \mathbb{N} \quad \text{with } \{a_k\}_{k \in \mathbb{N}} \geq 0 \quad (11)$$

is sufficient to ensure that the series $G(t)$ is absolutely and uniformly convergent with a radius $r = \infty$ and that $S_U[p]$ is a strictly concave functional (see the proof of Theorem 2). Although certainly restrictive, condition (11) is satisfied by many of the entropies known in the literature.

Remark 4. The universal-group entropy depends on the infinite set of parameters a_k , which are a priori independent, apart for the existence of specific constraints, as in eq. (11). To recover known cases of one-parametric or two-parametric entropies, depending let's say on the parameters q_1 and q_2 , we shall have that $a_k = f_k(q_1, q_2)$. Also, notice that the (apparently) more general case of an entropy of the form

$$S_U = k_B \sum_{i=1}^W p_i G \left(\ln \frac{1}{p_i^c} \right), \quad (12)$$

where c is a positive constant, can be easily recast in the language of Definition 3. Indeed, it is sufficient to consider a formal group exponential $\tilde{G}(t) = \sum_{k=0}^{\infty} \alpha_k \frac{t^{k+1}}{k+1}$, with $\alpha_k = a_k c^k$ or, equivalently, to require a modified condition (11) of the form $a_k > c(k+1)a_{k+1}$, with $\{a_k\}_{k \in \mathbb{N}} \geq 0$.

Remark 5. The entropy (10) is *trace-form*. Although very general, the family of trace-form entropies does not include other functional forms, like Renyi's entropy [37], very useful in several applications. The fact that the entropy (18) is trace-form, with a group exponential at least piecewise differentiable, ensures *Lesche stability* [29], which is a desirable property of any entropic functional, and more generally of any physical observable. Essentially, Lesche stability is the property of uniform continuity of an entropic functional in the space of probability distributions. It guarantees that a small variation of the set of probabilities produces a small change of entropy: given two distributions $\{p_i\}_{i=1, \dots, W}$ and $\{p'_i\}_{i=1, \dots, W}$ whose values are slightly different,

$$\forall \epsilon \quad \exists \delta > 0 \quad s.t. \quad \|p - p'\|_1 < \delta \implies \left| \frac{S[p] - S[p']}{S_{max}} \right| < \epsilon, \quad (13)$$

where, given a vector x , $\|x\|_1$ denotes the l_1 norm and S_{max} denotes the maximum value of the entropy.

This requirement, crucial for physical applications, is not so relevant in other contexts as, for instance, Information Theory.

5. Main properties of S_U entropy

In the following, we shall focus on the most relevant properties of the entropy (10), directly coming from its group-theoretical formulation, and on its relationship with other entropic functionals.

5.1. Composability

Theorem 1. *The entropy $S_U[p]$ is weakly composable.*

Proof. For sake of clarity, we shall formulate the proof in a general setting. Let $\{p_i^A\}_{i=1}^{W_A}$ and $\{p_j^B\}_{j=1}^{W_B}$ two sets of probabilities associated with two statistically independent systems A and B . The joint probability is given by

$$p_{ij}^{A \cup B} = p_i^A \cdot p_j^B$$

The total number of states of the composed system (including states with possibly null probability), is $W_{AB} = W_A W_B$. We have

$$\begin{aligned}
S_U(A \cup B) &= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_{ij}^{A \cup B} G \left(\ln \frac{1}{p_{ij}^{A \cup B}} \right) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B G \left(\ln \frac{1}{p_i^A} + \ln \frac{1}{p_j^B} \right) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B G(t_1 + t_2) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B G(F(s_1) + F(s_2)) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B \Phi(s_1, s_2) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B \left(s_1 + s_2 + \sum_{k,m=1}^{\infty} c_{km} s_1^k s_2^m \right) = \\
&= k_{\mathcal{B}} \sum_{i=1}^{W_A} \sum_{j=1}^{W_B} p_i^A \cdot p_j^B \left[G \left(\ln \frac{1}{p_i^A} \right) + G \left(\ln \frac{1}{p_j^B} \right) + \dots \right].
\end{aligned} \tag{14}$$

These equalities hold in full generality. Now, consider the Boltzmann and Tsallis composition laws, corresponding to the additive case (i.e. $c_{km} = 0 \forall k, m$) and to the case $c_{11} \neq 0$ and $c_{km} = 0$ for $(k, m) \neq (1, 1)$, respectively. The last expression in formula (14) yields immediately

$$S_U(A + B) = \Phi(S_U(A), S_U(B)). \tag{15}$$

Observe that whenever $c_{km} \neq 0$ for $(k, m) \neq (1, 1)$, a priori there is no guarantee that formula (15) holds in general, i.e. for any possible choice of the probability distributions of the two subsystems.

However, in the case that $\{p_i^A\}$ and $\{p_j^B\}$ are both the *uniform distribution*, formula (15) holds for the whole family of entropies represented by (10).

The formal power series $\phi(x, y) = G(F(x) + F(y))$ for any choice of G defines a formal group law. It verifies automatically the conditions of symmetry, null composability and transitivity. The thesis follows. \square

In the case of an entropy of the form (12), the same analysis holds.

Theorem 2. *The universal group entropy satisfies the first three SK axioms.*

(SK1). The group exponential $G(t)$ is a real analytic function of t . Consequently, the universal group entropy is (at least) a continuous function of its arguments (p_1, \dots, p_w) .

(SK2). The entropy $S_U[p]$ is supposed to be concave by definition. Nevertheless, we wish to prove here that condition (11), which allows to construct a large subclass of entropies of the form (10), is *sufficient* to guarantee concavity. To this aim, consider the quantity $\sum_{k=0}^{\infty} \frac{a_k}{k+1} x \left(\ln \frac{1}{x}\right)^{k+1}$. By imposing strict concavity we get

$$-\frac{1}{x} \left\{ a_0 - a_1 + (a_1 - 2a_2) \ln \frac{1}{x} + (a_2 - 3a_3) \left(\ln \frac{1}{x}\right)^2 + (a_3 - 4a_4) \left(\ln \frac{1}{x}\right)^3 + \dots \right\} < 0. \quad (16)$$

Condition $a_k > (k+1)a_{k+1}$ ensures that the inequality (16) is satisfied. Consequently, the associated entropy (10) is strictly concave in its space of parameters. Other choices of the sequence $\{a_k\}_{k \in \mathbb{N}}$ are clearly possible.

(SK3). Since by construction $G(0) = 0$, and $\lim_{x \rightarrow 0} x \left(\ln \frac{1}{x}\right)^k = 0$, it follows that $S_U(0) = 0$. Similarly, $S_U(1) = 0$.

5.2. Extensivity

We shall study here the extensivity properties of the universal-group entropy. Consider the uniform probability distribution (i.e. $p_i = 1/W$ for all $i = 1, \dots, W$). We have that

$$S_U[W] = k_B G(\ln W) \sim N \iff W(N) \sim \exp(F(N)). \quad (17)$$

The expression for $W(N)$ so determined can be computed explicitly. Indeed, as a formal series $F(s)$ is the compositional inverse of $G(t)$: it can be constructed by means of the Lagrange inversion principle.

From a physical point of view, one should also ensure that $W(N)$ be interpretable as an occupation law. A sufficient condition is that $W(N)$, as a real function, be defined for all $N \in \mathbb{N}$, with $\lim_{N \rightarrow \infty} W(N) = \infty$. These requirements usually restrict the space of allowed parameters. As an example, for the case of Tsallis entropy we have that the corresponding $W(N)$ is a well defined occupation law for the values of the entropic parameter $q < 1$.

Therefore, provided the previous condition is satisfied, when choosing an occupation law $W(N)$ of the form (17), there exists a representation of the universal-group entropy which is extensive for all systems whose phase space grows according with the selected occupation law. This notable fact ensures the large applicability of the entropy (10) in thermodynamical contexts.

5.3. Relation with other entropies

We shall discuss now how the previous formalism applies to the classification of known entropies.

Let us first observe that the universal entropy admits the following formal decomposition

$$S_U[p] = \sum_{k=1}^{\infty} a'_k S_k[p] \quad (18)$$

with $a'_k = a_{k-1}/k$, in terms of a set of elementary functionals $S_k[p] := k_{\mathcal{B}} \sum_{i=1}^W p_i \left(\ln \frac{1}{p_i}\right)^k$.

a) The *Boltzmann-Gibbs entropy*

$$S_{\mathcal{B}}[p] = k_{\mathcal{B}} \sum_{i=1}^W p_i \ln \frac{1}{p_i} \quad (19)$$

is obtained by means of the choice $G(t) = t$, $c = 1$ (i.e. for $a_1 = 1, a_i = 0, \forall i = 2, 3, \dots$).

b) The *Tsallis entropy* [50]

$$S_q[p] = k_{\mathcal{B}} \frac{\sum_{i=1}^W p_i^q - 1}{1 - q} \quad (20)$$

for $q < 1$ corresponds to the formal group exponential $G(t) = \frac{\exp[(1-q)t]-1}{1-q}$, $c = 1$. It can be decomposed as

$$S_q[p] = k_{\mathcal{B}} \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{1}{2}(1-q) \left(\ln \frac{1}{p_i}\right)^2 + \frac{1}{6}(1-q)^2 \left(\ln \frac{1}{p_i}\right)^3 + \dots \right\}. \quad (21)$$

When $q > 1$, $G(t) = \frac{\exp[(q-1)t]-1}{q-1}$ and a very similar expansion holds:

$$S_q[p] = k_{\mathcal{B}} \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{1}{2}(q-1) \left(\ln \frac{1}{p_i}\right)^2 + \frac{1}{6}(q-1)^2 \left(\ln \frac{1}{p_i}\right)^3 + \dots \right\}. \quad (22)$$

c) The *Kaniadakis entropy* [22]

$$S_{\kappa}[p] = k_{\mathcal{B}} \sum_{i=1}^W p_i \frac{p_i^{-\kappa} - p_i^{\kappa}}{2\kappa} \quad (23)$$

is obtained by means of the choice $G(t) = \frac{\exp(\kappa t) - \exp(-\kappa t)}{2\kappa}$, $c = 1$. The decomposition (18) of the Kaniadakis entropy is given by

$$S_{\kappa}[p] = k_{\mathcal{B}} \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{1}{3!} \kappa^2 \left(\ln \frac{1}{p_i}\right)^3 + \frac{1}{5!} \kappa^4 \left(\ln \frac{1}{p_i}\right)^5 + \frac{1}{7!} \kappa^6 \left(\ln \frac{1}{p_i}\right)^7 + \dots \right\} \quad (24)$$

where $-1 < \kappa \leq 1$.

d) Another important case is the $S_{c,d}$ entropy, introduced in [17]. This beautiful entropic form was obtained by taking into account two scaling laws that emerge from the requirement of the first three SK axioms. It reads

$$S_{c,d} = \frac{e}{1 - c + cd} \sum_{i=1}^W \Gamma(1 + d, 1 - c \ln p_i) - \frac{c}{1 - c + cd}. \quad (25)$$

Here $\Gamma(s, x)$ denotes the upper incomplete Gamma function (and $k_B = 1$), $c \in (0, 1]$, $d \in \mathbb{R}$.

The authors analyzed essentially all entropies known in the literature that satisfy the axioms (SK1)–(SK3), with the exception of group entropies, and observed that these entropies can be considered as particular cases of the $S_{c,d}$ entropy, for a suitable choice of (c, d) (see also [19] for a recent discussion of the role of the $S_{c,d}$ entropy). These parameters appear as the exponents characterizing the two scaling laws.

Consequently, in this Section, we will not compare exhaustively all the other cases already considered in [17] (as e.g. the Anteneodo-Plastino [4] and Shafee entropy [39]). Instead, we shall discuss directly the relationship between the $S_{c,d}$ entropy and the universal-group entropy. Consider the identity

$$\Gamma(s, x) = \Gamma(s) - \gamma(s, x)$$

where $\gamma(s, x)$ is the lower incomplete gamma function, as well as the series expansion of $\gamma(s, x)$ for real positive values of its arguments. We obtain that

$$\Gamma(1 + d, 1 - c \ln p_i) = \Gamma(1 + d) + t^d \sum_{k=0}^{\infty} \delta_k \frac{t^{k+1}}{k+1}, \quad (26)$$

where $\Gamma(a)$ is the Euler Gamma function,

$$\delta_k = \frac{(-1)^{k+1} (k+1)}{k!(k+d+1)}, \quad t = \left[\ln \frac{e}{p_i^c} \right]. \quad (27)$$

In the subsequent considerations, we shall restrict to the case $d \in \mathbb{N}$. To perform our analysis we shall use the identity, valid for $d \in \mathbb{N}$:

$$\int_K^{\infty} t^d e^{-t} dt = e^{-K} \sum_{n=0}^d \frac{\prod_{j=0}^n (d-j+1)}{d+1} K^{d-n}. \quad (28)$$

This identity (new, to the best of our knowledge) can be proven by a direct computation.

The identity (28) allows to expand the entropy in terms of the set $\{S_k[p]\}$. In general, we have

$$\begin{aligned} S_{c,d} &= \frac{1}{1-c+cd} \sum_{i=1}^W p_i^c \left\{ \sum_{k=0}^d \frac{1}{d+1} \prod_{j=0}^k (d-j+1) \sum_{n=0}^{d-k} \binom{d-k}{n} \left(\ln \frac{1}{p_i^c} \right)^n \right\} \\ &- \frac{c}{1-c+cd}, \quad d \in \mathbb{N}. \end{aligned} \quad (29)$$

Let us see some particular cases of the previous formula. We easily recover the two examples of [17], i.e.

$$S_{1,1}[p] = 1 + \sum_i p_i \ln \frac{1}{p_i}, \quad S_{1,2}[p] = 2 \left(1 + \sum_i p_i \ln \frac{1}{p_i} \right) + \frac{1}{2} \sum_i p_i \left(\ln \frac{1}{p_i} \right)^2.$$

For c arbitrary, let us write explicitly, for instance, the functionals corresponding to $d = 3$:

$$S_{c,3}[p] = \frac{1}{1+2c} \sum_{i=1}^W p_i^c \left\{ 16 + 15 \ln \frac{1}{p_i^c} + 6 \left(\ln \frac{1}{p_i^c} \right)^2 + \left(\ln \frac{1}{p_i^c} \right)^3 \right\} - \frac{c}{1+2c}, \quad (30)$$

and $d = 5$:

$$\begin{aligned} S_{c,5}[p] &= \frac{1}{1+4c} \sum_{i=1}^W p_i^c \left\{ 326 + 325 \ln \frac{1}{p_i^c} + 160 \left(\ln \frac{1}{p_i^c} \right)^2 + 50 \left(\ln \frac{1}{p_i^c} \right)^3 \right. \\ &\quad \left. + 10 \left(\ln \frac{1}{p_i^c} \right)^4 + \left(\ln \frac{1}{p_i^c} \right)^5 \right\} - \frac{c}{1+4c}. \end{aligned} \quad (31)$$

From the previous analysis, it emerges that the $S(c, d)$ entropy fits into the class (10) in the two cases ($c = 1, d \in \mathbb{N}$) (we can get rid of the constant term in the expansion) and ($c > 0, d = 0$).

e) The S_δ entropy has been introduced in [52] and independently in [55], and recently discussed in [53] in relation with black-holes thermodynamics:

$$S_\delta = k_{\mathcal{B}} \sum_{i=1}^W p_i \left(\ln \frac{1}{p_i} \right)^\delta, \quad 0 < \delta \leq (1 + \ln W). \quad (32)$$

The underlying algebraic structure can be analyzed on the uniform distribution (weak composability). We get easily the associated function $\Phi(x, y) = [x^{1/\delta} + y^{1/\delta}]^\delta$. Notice that the function $\Phi(x, y)$ has not an expansion in terms of a formal power series around $(x, y) = (0, 0)$, unless $\delta = 1$, which is the only strictly composable case. Also, $\Phi(x, y)$ does not define a group law over the reals, but simply a monoid, except for $\delta \in \mathbb{N}$, δ odd. A similar analysis can be performed for the case of the entropic functional

$$S_{q,\delta} = k_{\mathcal{B}} \sum_{i=1}^W p_i \left(\ln_q \frac{1}{p_i} \right)^\delta. \quad (33)$$

It reduces to (20) for $q \in \mathbb{R}$ and $\delta = 1$.

f) The *Borges-Roditi entropy* is a two-parametric entropy, introduced in [9]. The associated generalized logarithm reads

$$\text{Log}_{a,b}(x) = \frac{x^a - x^b}{a - b}, \quad (34)$$

which reproduces Abe's entropy [1] for $a = \sigma - 1, b = \sigma^{-1} - 1$. It also generalizes the entropy (23). The entropy (34) for a suitable choice of a, b satisfies the first three SK axioms (however, it is not related to the $S_{c,d}$ entropy). The related group exponential is

$$G_{\mathcal{A}}(t) = \frac{e^{at} - e^{bt}}{a - b}. \quad (35)$$

The formal group corresponding to (35), giving the interaction rule for this class, is known in the literature as the *Abel formal group*, defined by [10]

$$\Phi_{\mathcal{A}}(x, y) = x + y + \beta_1 xy + \sum_{j>i} \beta_i (xy^i - x^i y). \quad (36)$$

The coefficients β_n in (36) can be expressed as polynomials in a and b (see Proposition 3.1 of [10]):

$$\beta_1 = a + b, \quad \beta_n = \frac{(-1)^{n-1}}{n!(n-1)} \prod_{\substack{i+j=n-1 \\ i,j \geq 0}} (ia + jb), \quad n > 1.$$

The Borges-Roditi entropy possesses the following expansion:

$$S_{BR}[p] = k_B \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{1}{2}(a+b) \left(\ln \frac{1}{p_i} \right)^2 + \frac{1}{6}(a^2 + ab + b^2) \left(\ln \frac{1}{p_i} \right)^3 \dots \right\}. \quad (37)$$

g) The original notion of *group entropies* was introduced in [45] to describe an infinite family of weakly composable entropies. They are defined by

$$S_G(p) := k_B \sum_{i=1}^W p_i \text{Log}_G \left(\frac{1}{p_i} \right). \quad (38)$$

Here Log_G denotes the generalized logarithm

$$\text{Log}_G(x) = \frac{1}{\sigma} \sum_{n=l}^m k_n x^{\sigma n}, \quad l, m \in \mathbb{Z}, \quad m - l = r > 0, \quad x > 0 \quad (39)$$

where k_n are real constants such that

$$\sum_{n=l}^m k_n = 0, \quad \sum_{n=l}^m n k_n = 1, \quad (40)$$

and $k_m \neq 0, k_l \neq 0$. Conditions (40) are sufficient to ensure that $\lim_{\sigma \rightarrow 0} \text{Log}_G(x) = \ln x$. The class (38) is a subclass of (10), corresponding to the choice

$$G(t) = \frac{1}{\sigma} \sum_{n=l}^m k_n \exp(n\sigma t), \quad l, m \in \mathbb{Z}, \quad m - l = r > 0, \quad (41)$$

with the constraints (40).

Notice that there is much freedom in choosing the coefficients k_n , since only two conditions are required to guarantee the correct limit.

Due to the specific choice of the formal group structure, the generalized logarithm (39) is intimately related to a family of discrete derivatives. Indeed, let us denote by T the shift operator, acting on a function f as $Tf(x) = f(x+\sigma)$.

The discrete derivatives of order r associated to the generalized logarithms (39) are defined to be

$$\Delta_r(\sigma) = \frac{1}{\sigma} \sum_{n=l}^m k_n T^n, \quad r = m - l.$$

In the limit $\sigma \rightarrow 0$, $\Delta_r(\sigma) = \partial_x$. For a suitable choice of the values of the parameter σ , each of the obtained entropies is concave, and weakly composable. The relation between group entropies and the class $S_{c,d}$ was an open question. Actually, the two classes are simply *different*. In other words, the entropies (38), with the exception of Tsallis entropy, cannot be obtained from $S_{c,d}$ entropy by specializing the coefficients (c, d) .

The first representative of the group-entropy class is indeed the Tsallis entropy, obtained for $\text{Log}_G(x) = \frac{x^\sigma - 1}{\sigma} = \text{Log}_T(x) = \frac{x^{1-q} - 1}{1-q}$ (here $\sigma = 1 - q$). The second representative of the class, the Kaniadakis entropy, is not in the $S_{c,d}$ family.

Consider the logarithm $\text{Log}_{III}(x) = \frac{x^{1-q} - 2x^{-(1-q)} + x^{-2(1-q)}}{1-q}$ (associated to a third order discrete derivative). The corresponding entropy is

$$S_{III} := \frac{k_B}{1-q} \sum_{i=1}^W p_i \left(p_i^{2(1-q)} - 2p_i^{(1-q)} + p_i^{-(1-q)} \right) \quad (42)$$

This entropy is concave for $2/3 < q < 1$ (here we prefer to use the parameter $q = 1 - \sigma$). In the limit $q \rightarrow 1$, the entropy S_{III} reduces to the Boltzmann-Gibbs entropy. Also, the entropy S_{III} is weakly composable, since it corresponds to $G(t) = \frac{e^{(1-q)t} - 2e^{-(1-q)t} + e^{-2(1-q)t}}{1-q}$. However, it is not obtainable neither from Borges-Roditi's nor from $S_{c,d}$ entropy. Its expansion is given by

$$S_{III}[p] = k_B \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{3}{2}(1-q) \left(\ln \frac{1}{p_i} \right)^2 - \frac{5}{6}(1-q)^2 \left(\ln \frac{1}{p_i} \right)^3 \dots \right\}$$

Remark 6. Entropy (42) provides an interesting example of a functional which is in the universal class (10), but it does not satisfy condition (11). Indeed, this condition, although sufficient, is not necessary to define an entropic functional with good thermodynamic properties.

Another entropy, with very similar properties, is the following one, corresponding to a fourth order discrete derivative:

$$S_{IV} := \frac{k_B}{1-q} \sum_{i=1}^W p_i \left(p_i^{-2(1-q)} - \frac{3}{2}p_i^{-(1-q)} + \frac{3}{2}p_i^{(1-q)} - p_i^{2(1-q)} \right).$$

By using the construction proposed in [45], infinitely many new entropies, each of them depending on a parameter q , and defined in a specific interval of values of q , can be constructed. They are not particular cases of any of the previously discussed entropies. At the same time, being trace-form and satisfying the

axioms (SK1)–(SK3), they comply with the two scaling laws postulated in [17] (as is also easy to prove directly).

Both families, i.e. group entropies and the $S_{c,d}$ entropies can be interpreted in the universal group-theoretical framework. Indeed, the $S_{c,d}$ entropy for specific choices of (c, d) is expressible in terms of the entropy (10) by means of the general formula (29) and is consequently weakly composable. The group entropies (38) by construction are a subfamily of the entropy (10).

6. A new three-parametric group entropy: the $S_{q,\alpha,\beta}$ entropy

To illustrate the potential richness of the theory previously developed, we wish to present here a new nontrivial multi-parametric entropic functional, obtained as a special case of the construction sketched above. It reads

$$S_{\alpha,\beta,q}[p] := \frac{k_B}{1-q} \sum_{i=1}^W p_i \left(\alpha p_i^{-2(1-q)} + \frac{1}{2} (1-3\alpha+\beta) p_i^{-(1-q)} + \frac{1}{2} (\alpha-1-3\beta) p_i^{(1-q)} + \beta p_i^{2(1-q)} \right). \quad (43)$$

The main properties of this entropy are the following.

i)

$$\lim_{q \rightarrow 1} S_{\alpha,\beta,q}[p] = S_{BG}[p]$$

irrespectively of the choice of the parameters.

ii) The entropy (43) is trace-form.

iii) It satisfies the first three SK axioms; in particular, it is concave, for instance, for $1/2 < q < 3/2$, $0 < \alpha < 1/4$, $-1/4 < \beta < 0$.

iv) The parameters α, β, q are independent; consequently, the entropy (43) is not reducible to the previously discussed two-parametric classes.

v) It is weakly composable, with expansion

$$S_{\alpha,\beta,q}[p] := k_B \sum_{i=1}^W p_i \left\{ \ln \frac{1}{p_i} + \frac{3}{2} (\alpha + \beta) (1-q) \left(\ln \frac{1}{p_i} \right)^2 + \frac{1}{6} (1 + 6\alpha - 6\beta) (1-q)^2 \left(\ln \frac{1}{p_i} \right)^3 + \dots \right\} \quad (44)$$

The entropy (44) is in the group entropy class, but it was not considered explicitly in [45]. Also, it is not a particular case of $S_{c,d}$ or S_{BR} . This example is just a representative of a large class of entropic functionals, each of them multiparametric, that can be constructed by using the group-theoretical framework previously discussed. However, the complete analysis of this class, which is a specific realization of the notion of universal-group entropy, is outside the scopes of this paper.

7. Distribution functions and thermodynamic properties

We shall discuss the maximization of group entropies under appropriate constraints: we adopt a generalized maximum entropy principle (see e.g. [23], [2]). We will see that the Legendre structure of classical thermodynamics, at least in some aspects, is preserved in our group-theoretical framework.

Precisely, let

$$\text{Log}_U[\epsilon] = G(\ln \epsilon)$$

where $G(t)$ is the power series (9), with the constraint (11). Consider an isolated system in a stationary state (*microcanonical ensemble*). The optimization of S_U leads to the equal probability case, i.e. $p_i = 1/W, \forall i$. Therefore, we have

$$S_U[p] = k_B \text{Log}_U W, \quad (45)$$

which reduces to the celebrated Boltzmann formula $S_{BG} = k_B \ln W$ in the case of uncorrelated particles.

Let us consider a system in thermal contact with a reservoir (*canonical ensemble*). We introduce the numbers ϵ_i , interpreted as the values of a physically relevant observable, typically the value of the energy of the system in its i th state. Assume that $p_i(\epsilon_i)$ is a normalized and monotonically decreasing distribution function of ϵ_i . The *internal energy* \mathcal{V} in a given state is defined as $\mathcal{V} = \sum_{i=1}^W \epsilon_i p_i(\epsilon_i)$.

As usual, we shall study the variational problem of the existence of a stationary distribution $\tilde{p}_i(\epsilon)$. However, this analysis can not be performed in direct analogy with the standard case. We introduce the functional

$$L = S_G[p] - \alpha \left[\sum_i p(\epsilon_i) - 1 \right] - \beta \left[\sum_{i=1}^W \epsilon_i p_i(\epsilon_i) - \mathcal{V} \right], \quad (46)$$

where α and β are Lagrange multipliers. The vanishing of the variational derivative of this functional with respect to the distribution p_i provides the stationary solution

$$\tilde{p}_i = \frac{E(-\alpha - \beta \epsilon_i)}{Z}, \quad (47)$$

with $Z = \sum_{i=1}^W E(-\alpha - \beta \epsilon_i)$, and $E(\cdot)$ is an invertible function.

However, as already pointed out in [45], only in particular cases we are able to identify E with the inverse of a generalized logarithm Log_G .

In [23], [24], the class of entropies allowing a treatment with the variational approach described above has been determined. In the Legendre-Massieu framework, one can derive the interesting relation

$$\text{Log}_G(Z) + \beta \mathcal{V} = S_G. \quad (48)$$

The previous equation can be used to introduce a thermodynamic observable T , which has the interpretation of a local temperature for a non-equilibrium metastable state. Precisely, we can define it from $\frac{\partial S_U}{\partial \mathcal{V}} = \frac{1}{T}$. Analogously, a

generalized free energy can be introduced according to $\mathcal{F} = \mathcal{V} - TS_U$. However, in the context of Tsallis entropy, the variational approach leading to the definition of an effective temperature is based on more specific constraints (escort distributions [52]).

The determination of a group-theoretical procedure to construct the appropriate constraints for an entropy of the class (10) is an open problem.

8. On the asymptotic behaviour of generalized entropies

From a mathematical point of view, the study of the asymptotic behaviour of a given entropy, in the limit of large size systems, is not a well defined task, even for the standard Boltzmann-Gibbs entropy.

Let W be the number of microscopic states admitted by a system. We shall focus on the case of large size systems (under the hypothesis that $W \rightarrow \infty$ for $N \rightarrow \infty$). An entropy is then a functional $S = S[p]$ defined on the space \mathcal{P}_∞ .

A simple argument proves that, depending on the choice of the distribution in \mathcal{P}_∞ , we can get infinitely many different limits even for the S_{BG} case. Indeed, consider the probability distribution $\{p\} = (1, 0, 0, \dots)$ with infinitely many entries. Then $S_{BG}[p] = 0$. If $\{p\} = (1/2, 1/2, 0, 0, \dots)$, then $S_{BG}[p] = \ln 2$ (in units of k_B). Let $\mathcal{N} = 10^{80}$ (i.e., the estimated number of atoms in the observable Universe), and $\{p\} = (\underbrace{1/\mathcal{N}, 1/\mathcal{N}, \dots, 1/\mathcal{N}}_{\mathcal{N}\text{-times}}, 0, 0, \dots)$, then $S_{BG}[p] =$

$80 \cdot \ln 10 \simeq 184.2$. From the point of view of probability and information theory, there is no way to have uniqueness of the limit on the full space \mathcal{P}_∞ .

In the domain of classical thermodynamics, a priori the same objection applies. However, if we accept to restrict to the important case of the uniform distribution (which is not the only physically interesting case), then one can define properly a *thermodynamic limit of an entropy on the uniform distribution*. From a physical perspective, this would correspond to the case of a double limit, both of large system size and of large times.

The only meaningful way to compare the behaviour of different entropies is to compute them all *in the same regime*. Once we accept to restrict to uniform probabilities, then the known entropies assume the form of the following asymptotic functions: either $(\ln W)^a$ or W^b (possibly multiplied by parameters), or the product of these forms. Presently, we cannot exclude that other forms could also be possible.

All this is a consequence of the analysis of scaling laws performed in [17]. However, it must be noticed that these simple functions of W are not entropic functionals, and they cannot be obtained from a known entropy by specializing its parameters.

The problem of “comparing” entropies in the large size limit finds its most simple answer in the regime, if exists, where they are extensive. In classical thermodynamics, the only reason to consider generalized entropies is the fact that they can be *extensive* in regimes where the S_{BG} entropy is not. This is the truly important limiting property.

What is crucial is that *all admissible entropies have the same behavior* in the regime where they are extensive, i.e. proportional to the number N of particles of the system.

Therefore, it is not surprising that different entropies could share the same asymptotic behavior.

When we are interested in more general contexts as probability theory and information theory, then all of the infinitely many possible distributions are a priori relevant. Therefore, according to the previous discussion, the comparison among asymptotic behaviors of different entropies loses its meaning.

9. Open problems and future perspectives

The main message of this paper is that *the notion of entropy is intimately related with group theory*. Indeed, admissible entropies have associated a group law, controlling crucially their properties.

The universal-group entropy is a flexible tool, providing new insight in different contexts of the theory of complex systems. For instance, it could offer a more general approach to the description of the evolution of biological complexity [3]. The role of the universal-group entropy in geometric information theory should be properly clarified. New examples of Kullback-Leibler divergences are presently under investigation. They could provide refined tests of similarity, for instance, in the analysis of genomic sequences. A nonadditive entropy taken from the class (10) a priori could be more appropriate than the Shannon one, in order to take into account the so-called *epistatic correlations* between sites in a genomic chain.

An open problem is to find constructively classes of new entropies obtained as special cases of the general definition (10). New results along these lines are contained in [47].

Another important aspect that deserves to be clarified is the possible number-theoretical content of the notion of universal group entropy. In [45], a connection between a set of entropic functionals and a class of Dirichlet series has been established. It has been proved that the formal group exponentials defining a specific family of entropies can be used to define L -functions. The most paradigmatic case is offered by Tsallis entropy, directly related to the Riemann zeta function. A construction of universal congruences and generalized Bernoulli polynomials connected with formal groups has been provided in [46] and in [43]. It would be interesting to realize explicitly this picture in the general group-theoretical framework developed in this work.

Appendix A. The Shannon-Khinchin axioms

In the original formulation due to Khinchin [25], the Shannon-Khinchin axioms were called “properties” that guarantee Khinchin’s uniqueness theorem for the Boltzmann-Gibbs entropy S . In particular, the properties of continuity and maximum were unified into a unique statement. Here we propose a modern version, in terms of four requirements.

(SK1) (Continuity). The function $S(p_1, \dots, p_W)$ is continuous with respect to all its arguments

(SK2) (Maximum principle). The function $S(p_1, \dots, p_W)$ takes its maximum value for the uniform distribution $p_i = 1/W, i = 1, \dots, W$.

(SK3) (Expansibility). Adding an impossible event to a probability distribution does not change its entropy: $S(p_1, \dots, p_W, 0) = S(p_1, \dots, p_W)$.

(SK4) (Additivity). Given two subsystems A, B of a statistical system,

$$S(A \cup B) = S(A) + S(B | A);$$

here $S(B | A)$ denotes the conditional entropy associated with the conditional distribution $p_{ij}(B | A)$.

Acknowledgments. I wish to thank heartily prof. C. Tsallis for a careful reading of the manuscript and many useful discussions, and prof. G. Parisi for reading the manuscript and for encouragement. Interesting discussions with prof. A. González-López, R. A. Leo and M. A. Rodríguez are also gratefully acknowledged.

This work has been supported by the research project FIS2011–22566, Ministerio de Ciencia e Innovación, Spain.

References

- [1] S. Abe, Generalized entropy optimized by a given arbitrary distribution, J. Phys. A: Math. Gen. **36** 8733–8738 (2003).
- [2] S. Abe and S. Thurner, Robustness of the second law of thermodynamics under generalizations of the maximum entropy method, Europhys. Lett **81** (2008) 1004.
- [3] C. Adami, C. Ofria and T. C. Collier, Evolution of biological complexity, Proc. Natl. Acad. Sci. USA **97** (2000) 4463–4468.
- [4] C. Anteneodo and A. R. Plastino, Maximum entropy approach to stretched exponential probability distributions, J. Phys. A: Math. Gen. **32**, 1089 (1999).
- [5] A. Baker, Combinatorial and arithmetic identities based on formal group laws, Lect. Notes in Math. **1298**, 17–34, Springer (1987).
- [6] C. Beck and E. D. G. Cohen, Superstatistics, Physica A **322** 267–275 (2003).
- [7] M. Berta, M. Christandl, R. Colbeck, J. M. Renes and R. Renner, The uncertainty principle in the presence of quantum memory, Nature Phys., vol. 6, 659–662 (2010).
- [8] S. Bochner, Formal Lie groups, Ann. Math. **47**, 192–201 (1946).
- [9] E. P. Borges and I. Roditi, A family of nonextensive entropies, Phys. Lett. A **246** 399–402 (1998).

- [10] V. M. Bukhshtaber and A. N. Kholodov, *Formal groups, functional equations and generalized cohomology theories*, Math. USSR Sbornik **181** (1990) 75–94, English transl. Math. S. B. **69** (1991), 77–97.
- [11] V. M. Bukhshtaber, A. S. Mishchenko and S. P. Novikov, Formal groups and their role in the apparatus of algebraic topology, Uspehi Mat. Nauk **26** (1971), 2, 161–154, transl. Russ. Math. Surv. **26**, 63–90 (1971).
- [12] V. M. Bukhshtaber and S. P. Novikov, Formal groups, power systems and Adams operators, Math. Sb. **84**, 116–153 (1971).
- [13] A. Baker, F. Clarke, N. Ray, L. Schwartz, On the Kummer congruences and the stable homotopy of BU, Trans. Amer. Math. Soc. **316**, 385–432 (1989).
- [14] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, II edition, John Wiley and Sons (1985).
- [15] N. Canosa and R. Rossignoli, Generalized nonadditive entropies and quantum entanglement, Phys. Rev. Lett. **88** 170401 (2002).
- [16] M. Gell-Mann, *The Quark and the Jaguar: Adventures in the Simple and the Complex*, Macmillan (1995).
- [17] R. Hanel and S. Thurner, A comprehensive classification of complex statistical systems and an axiomatic derivation of their entropy and distribution functions, Europhys. Lett. **93** 20006 (2011).
- [18] R. Hanel, S. Thurner, M. Gell-Mann, Generalized entropies and logarithms and their duality relations, Proc. Nat. Acad. Sci. USA **108**, 6390–6394 (2012).
- [19] R. Hanel, S. Thurner and M. Gell-Mann, How multiplicity determines entropy: derivation of the maximum entropy principle for complex systems, PNAS **111** (19), 6905–6910 (2014).
- [20] M. Hazewinkel, *Formal Groups and Applications*, Academic Press, New York, 1978
- [21] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum Entanglement, Rev. Mod. Phys. **81**, 865 (2009).
- [22] G. Kaniadakis, Statistical mechanics in the context of special relativity, Phys. Rev. E **66** 056125 (2002).
- [23] G. Kaniadakis, Maximum entropy principle and power-law tailed distributions, Eur. Phys. J. B **70** 3–13 (2009).
- [24] G. Kaniadakis, M. Lissia and A. M. Scarfone, Physica A **340** (2004) 41–49; Phys. Rev. E **71** (2005) 046128.
- [25] A. I. Khinchin, *Mathematical Foundations of Information Theory*, Dover, New York, 1957.

- [26] S. Kullback and R. A. Leibler, On Information and Sufficiency, *Ann. of Math. Stat.* **22**, 79–86 (1951).
- [27] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* **30** 175–193 (1906).
- [28] P. T. Landsberg, *Thermodynamics*, Interscience, New York, 1961.
- [29] B. Lesche, Instabilities of Rényi entropies, *J. Stat. Phys.* **27**, 419 (1982).
- [30] S. Marmi and P. Tempesta, Hyperfunctions, formal groups and generalized Lipschitz summation formulas, *Nonlinear Analysis* **75**, 1768–1777 (2012).
- [31] J. Naudts, Deformed exponentials and logarithms in generalized thermo-statistics, *Physica A*, **316**, 323–334 (2002).
- [32] J. Naudts, Generalised exponential families and associated entropy functions, *Entropy*, **10**, 131–149 (2008).
- [33] S. P. Novikov, The methods of algebraic topology from the point of view of cobordism theory, *Izv. Akad. Nauk SSSR Ser. Mat.* **31** (1967), 885–951, transl. *Math. SSR–Izv.* **1** (1967), 827–913.
- [34] N. A. Peters, T.-C. Wei, P. G. Kwiat, Mixed state sensitivity of several quantum information benchmarks, *Physical Review A* **70** (5), 052309 (2004).
- [35] D. Quillen, On the formal group laws of unoriented and complex cobordism theory, *Bull. Amer. Math. Soc.* **75**, 1293–1298 (1969).
- [36] N. Ray, Stirling and Bernoulli numbers for complex oriented homology theory, in *Algebraic Topology*, Lecture Notes in Math. 1370, pp. 362–373, G. Carlsson, R. L. Cohen, H. R. Miller and D. C. Ravenel (Eds.), Springer–Verlag, 1986.
- [37] A. Rényi, *Probability Theory*, North–Holland, Amsterdam, 1970.
- [38] J.–P. Serre, *Lie algebras and Lie groups*, Lecture Notes in Mathematics, 1500 Springer–Verlag (1992).
- [39] F. Shafee, Lambert function and a new nonextensive form of entropy, *IMA J. Appl. Math.* **72**, 785 (2007).
- [40] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27** (1948) 379–423, **27** 623–653 (1948).
- [41] C. Shannon, W. Weaver, *The mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.
- [42] P. Tempesta, Formal groups, Bernoulli–type polynomials and L –series, *C. R. Math. Acad. Sci. Paris, Ser. I* **345**, 303–306 (2007).

- [43] P. Tempesta, On Appell sequences of polynomials of Bernoulli and Euler type, *J. Math. Anal. Appl.* **341**, 1295–1310 (2008).
- [44] P. Tempesta, L-series and Hurwitz zeta functions associated with the universal formal group, *Annali Sc. Normale Superiore, Classe di Scienze*, IX, 1–12 (2010).
- [45] P. Tempesta, Group entropies, correlation laws and zeta functions, *Phys. Rev. E* **84**, 021121 (2011).
- [46] P. Tempesta, The Lazard formal group, universal congruences and special values of zeta functions, *Transactions of the American Mathematical Society*, **367**, 7015–7028 (2015).
- [47] P. Tempesta, A theorem on the existence of generalized trace-form entropies, preprint (2015).
- [48] P. Tempesta, A new family of composable entropies from group theory: the Z-entropies, arxiv: 1507.07436 (2015).
- [49] <http://tsallis.cat.cbpf.br/TEMUCO.pdf>
- [50] C. Tsallis, Possible generalization of the Boltzmann–Gibbs statistics, *J. Stat. Phys.* **52**, Nos. 1/2, 479–487 (1988).
- [51] C. Tsallis, Generalized entropy–based criterion for consistent testing, *Phys. Rev. E* **58**, 1442–1445 (1998).
- [52] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics—Approaching a Complex World*, Springer, Berlin (2009).
- [53] C. Tsallis, L. Cirto, Black hole thermodynamical entropy, *Eur. Phys. J. C* **73**, 2487 (2013).
- [54] C. Tsallis, M. Gell–Mann and Y. Sato, Asymptotically scale–invariant occupancy of phase space makes the entropy S_q extensive, *Proc. Nat. Acad. Sci. USA*, **102**, 15377–15382 (2005).
- [55] M. R. Ubriaco, Entropies based on fractional calculus, *Phys. Lett. A* **373**, 2516–2519 2009.
- [56] A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50**, 221 (1978).