


Review

# Group Entropies: From Phase Space Geometry to Entropy Functionals via Group Theory

Henrik Jeldtoft Jensen <sup>1,2,\*</sup> and Piergiulio Tempesta <sup>3,4</sup> 

<sup>1</sup> Centre for Complexity Science and Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

<sup>2</sup> Institute of Innovative Research, Tokyo Institute of Technology, 4259, Nagatsuta-cho, Yokohama 226-8502, Japan

<sup>3</sup> Departamento de Física Teórica, Universidad Complutense de Madrid, 28040 Madrid, Spain; p.tempesta@fis.ucm.es

<sup>4</sup> Instituto de Ciencias Matemáticas (ICMAT), 28049 Madrid, Spain

\* Correspondence: h.jensen@imperial.ac.uk; Tel.: +44-(0)-20-759-49853

Received: 12 September 2018; Accepted: 10 October 2018; Published: 19 October 2018



**Abstract:** The entropy of Boltzmann-Gibbs, as proved by Shannon and Khinchin, is based on four axioms, where the fourth one concerns additivity. The group theoretic entropies make use of formal group theory to replace this axiom with a more general composability axiom. As has been pointed out before, generalised entropies crucially depend on the number of allowed degrees of freedom  $N$ . The functional form of group entropies is restricted (though not uniquely determined) by assuming extensivity on the equal probability ensemble, which leads to classes of functionals corresponding to sub-exponential, exponential or super-exponential dependence of the phase space volume  $W$  on  $N$ . We review the ensuing entropies, discuss the composability axiom and explain why group entropies may be particularly relevant from an information-theoretical perspective.

**Keywords:** generalised entropies; formal groups; phase space growth rate

## 1. Introduction

The aim of this paper is to discuss and review the construction of a class of entropies recently introduced, called group entropies [1–4]. We shall make several preliminary observations in order to ensure that our line of thinking is transparent.

In thermodynamics, according to Clausius the entropy  $\Delta S = Q/T$  is defined macroscopically in terms of its change induced by a heat flow  $Q$  at temperature  $T$ . A connection to the microscopic world is obtained in statistical mechanics by Boltzmann's expression

$$S[p] = - \sum_{i=1}^W p_i \ln p_i = \ln W. \quad (1)$$

where the last equality is valid on the equal probability ensemble  $p_i = 1/W$  where  $p_i$  is the probabilistic weight of state  $i$  and  $W$  denotes the number of available states. Hereafter, we assume  $k_B = 1$ .

Jaynes made contact with information theory and pointed out that Boltzmann's microcanonical and canonical ensembles can be viewed as the probabilistic weights that maximise the Boltzmann-Shannon entropy functional in Equation (1) under suitable constraints. The microcanonical ensemble is obtained when only the normalisation constraint is imposed, whereas the canonical ensemble arises when the normalisation and the average energy constraint are both assumed [5].

Here we will think of entropies in the spirit of information theory, i.e., functionals on probability space. Therefore the first three of the four Shannon-Khinchin axioms [6,7] are unobjectionable and the

entropy of a system in the equal probability ensemble  $p_i = 1/W$  will be considered to be a measure of uncertainty. It is then natural to assume that entropy in the equal probability ensemble is extensive (in the limit of a large number of particles). Namely, the more particles the more uncertain is the least biased ansatz  $p_i = 1/W$ . We express this mathematically by saying that an entropy is *extensive* if, in the limit of large  $N$ , the entropy on the equal probability ensemble behaves as  $S(p_i = 1/W) = \lambda N$ ,  $N \gg 1$ . Hence we consider extensivity, defined in this way, to be a required property of the entropies we are going to consider. This is of course also done within the  $q$ -statistics framework [8]. It is also worth to recall that extensivity is a necessary condition for an entropy to play the role as a rate function in large deviation theory [9,10].

To try to keep our use of concepts clear and make a transparent distinction between extensivity and *additivity*, let us immediately mention, though we will elaborate below, that we consider an entropy to be additive, if, for two independent systems the entropy of the two systems considered as a whole is equal to the sum of the parts.

Once having established how the entropy of the entire system in the uniform ensemble scales with the number of particles (degrees of freedom), we need to make an equally important decision about composition of systems. Imagine a system that is obtained by merging two given systems  $A$  and  $B$  and assume that  $A$  and  $B$  are statistically *independent*. We start analysing this case not because we believe that real systems typically can be considered as collections of independent subsystems. Although in classical thermodynamics independence is often an excellent approximation, it is most likely not the case when dealing with complex systems. We consider the independent case for two reasons. First, one can always formally consider independent systems as constituting a whole and the entropy needs to be defined so it can handle this. Secondly, this requirement allows for important mathematical constraints on the entropy and, as explained in Section 2, establishes a link to group theory.

More precisely, since  $A$  and  $B$  are assumed to be independent, we can now either consider the cartesian product  $A \times B$  of the states of the systems  $A$  and  $B$  as one system and compute the entropy of  $S(A \times B)$ , or we may as well first compute the entropy of the parts  $S(A)$  and  $S(B)$  and afterwards decide to consider  $A \times B$  as a whole. We recall that entropies are functionals on probability space, which define a probabilistic weight for each of the microstates of a given system. For the independent combination considered here, we of course have that the microstates of  $A \times B$  are given by the combined states  $(i, j)$  where  $i$  and  $j$  refer to the specific microstates of  $A$  and  $B$ , respectively. The independence ensures that the probability distributions describing  $A$ ,  $B$  and  $A \times B$  are related as  $p_{i,j}^{A \times B} = p_i^A p_j^B$ . So we need to ensure that the entropy functional computed using  $p_{i,j}^{A \times B}$  is consistent with the expression obtained by computing first the functional on  $p_i^A$  and  $p_j^B$  and then combining the result. That is to say, we need a function  $\phi(x, y)$  that takes care of the combination of the two independent systems  $A$  and  $B$  into one whole

$$S(A \times B) = \phi(S(A), S(B)). \quad (2)$$

If the entropy is additive we have  $\phi(x, y) = x + y$ . The relation in Equation (2) is of course basic in as much as it is a formality to consider the cartesian product  $A \times B$  as a whole or as combined of two independent subsystems  $A$  and  $B$ . Equation (2) should therefore be satisfied for all possible choices of  $p_i^A$  and  $p_j^B$ . In Section 2 below we discuss the properties of  $\phi(x, y)$  in more detail. Here we just mention that Equation (2) ensures that in cases when the entire system can be considered to be a collection of subsystems, the entropy of a composed system  $S(A \times B)$  depends on the entropy of  $A$  and the entropy of  $B$  of the component systems only, without the need for a microscopic description of them. Thus, in this way one can associate naturally the notion of entropy to a macroscopic system starting from the knowledge of its constituents. Complex systems are often defined as involving some degree of emergence, which is captured by the famous Aristotle's quote "The whole is greater than the sum of the parts" (Metaphysics). A concrete explicit example of such situations was recently considered by introducing the so called pairing model [11] in which particles may combine to form new paired states

which are entirely different from the single particle states. For a specific example think of hydrogen atoms. When two hydrogen atoms combine to form a hydrogen molecule,  $H_2$  bound states are formed which cannot be reached within the cartesian product of the phase space of the two individual hydrogen atoms [12]. More generally, when dealing with complex systems the independent combination of subsystems will typically be different from the whole [13]. Let us by  $AB$  denote the system obtained by bringing the  $N_A$  particles of system  $A$  together with the  $N_B$  particles of system  $B$  and allowing for the two sets of particles to establish all possible interactions or interdependencies among the particles from  $A$  and those from  $B$ . In the example of the pairing model [11],  $AB$  will also contain new “paired states” among particles in  $A$  and particles in  $B$ . Therefore  $AB \neq A \times B$  since  $A \times B$  consists only of the states that can be labelled as  $(i, j)$  where  $i = 1, \dots, W_A$  runs through all the states of system  $A$  and  $j = 1, \dots, W_B$  runs through all the states of system  $B$ . New emergent states formed by combining particles from  $A$  and  $B$  are not included in  $A \times B$ .

To illustrate this point, think of system  $A$  and system  $B$  as consisting of a single hydrogen atom each. Then  $A \times B$  is the set  $A \times B = \{(\mathbf{r}_a, \mathbf{p}_a), (\mathbf{r}_b, \mathbf{p}_b)\}$  where  $\mathbf{r}_i$  and  $\mathbf{p}_i$  with  $i = a, b$  are the position and momenta of the hydrogen atom  $A$  or  $B$ . The combined system  $AB$  in contrast contains new emergent molecular states  $H_2$  consisting of the hydrogen atom  $A$  bound together with the hydrogen atom  $B$ . We recall that the conventional description considers  $H$  and  $H_2$  as two distinct ideal gases, introduces a chemical potential for each and minimises the Helmholtz free energy for the  $H$  and  $H_2$  mixture, see e.g., Section 8.10 in [14]. In this way one does not need to handle super-exponentially fast growing phase spaces since  $H_2$  is not considered a paired state of  $H$  atoms. The profound, though by now of course very familiar, concept of the chemical potential makes it possible to escape the combinatorial explosion in this specific case.

We require that the entropy evaluated on the equal probability ensemble for the fully interacting system satisfies (asymptotically) extensivity, i.e., that

$$S(AB) = S(1/W(AB)) \simeq \lambda N_{AB} = \lambda(N_A + N_B) \text{ for } N_A \gg 1 \text{ and } N_B \gg 1. \quad (3)$$

However, we cannot in general insist that

$$S(AB) = \phi(S(A), S(B)) = S(A \times B). \quad (4)$$

In the Boltzmann-Shannon case for which  $\phi(x, y) = x + y$  the relation in Equation (4) holds when  $W(AB) = W(A)W(B)$ , i.e., we have an exponential dependence  $W(N) = k^N$ . Below we'll discuss in detail how the functional dependence  $W(N)$  of the total number of states on the number of particles will determine the properties of both the entropy and the composition law  $\phi(x, y)$  and we'll see that typically  $S(AB) \neq S(A \times B)$ . When  $W(N)$  does not have an exponential behaviour, one either gets entropies equivalent to the Tsallis entropy, for sub-exponential algebraic dependence  $W(N) = N^a$ , or new group entropies for super-exponential phase space growth rates as, for instance,  $W(N) = N^{\gamma N}$ .

For complex systems, for which entropies are typically non-additive, the group entropies discussed here immediately suggest a measure of how complex a system is. Precisely because  $A \times B \neq AB$  for complex systems and therefore the entropy of the fully interdependent system  $AB$  will be different from the cartesian combination  $A \times B$ , a measure of the essential emergent interdependence can be constructed as

$$\Delta(AB) = S(A \times B) - S(AB) = \phi(S(A), S(B)) - S(AB). \quad (5)$$

This measure can be thought of as a possible generalisation of the usual mutual information and could perhaps be useful e.g., as an alternative to Tononi's Integrated Information [15,16] as a measure that can quantify very entangled complex systems such as, say, consciousness. A thorough discussion of this complexity measure will appear in [17].

The remainder of the article is organised as follows. In Section 2, we present a brief and self-consistent introduction to group entropies. We explain in Section 3 how the phase space growth volume  $W(N)$  determines a specific group law that, in turn, enables us to characterise the functional form of allowed entropies and the rule for composing statistically independent systems. Precisely, we show that for a given function  $W(N)$  there exists a construction of dual entropies, a trace-form and a non-trace-form one, sharing over the uniform distribution the same composition law  $\phi(x, y)$ .

We relate the group entropies to existing entropies and discuss in Section 4 the probabilities  $p_i$  derived by maximizing the entropy under constraints.

## 2. Basic Results on Group Entropies

In this section, we shall present a brief introduction to some basic aspects of the theory of group entropies. The mathematical apparatus will be kept to a minimum. For an more complete discussion, the reader is referred to the original papers [2–4,18]. We start out with the composition requirement in Equation (2). We need to require that (i)  $\phi(x, y) = \phi(y, x)$ , since  $A$  and  $B$  are just labels that can obviously be interchanged. At the same time, we also require that the process of composition can be made in a associative way: (ii)  $\phi(x, \phi(y, z)) = \phi(\phi(x, y), z)$ . Finally, if system  $B$  is in a state of zero entropy, we wish that the entropy of the composed state  $A \times B$  coincides with the entropy of  $A$ . In other words, (iii)  $\Phi(x, 0) = x$ . We shall say that an entropy satisfies the *composability axiom* if there exists a function  $\phi(x, y)$  such that Equation (2) is satisfied, jointly with the previous properties of commutativity, associativity and composition with a zero-entropy state [1,3].

In order to ascertain the plausibility of the composability axiom, observe that, first of all, it is satisfied by Boltzmann's entropy. It is a crucial requirement for possible thermodynamical applications. Indeed, it means that the entropy of a system composed of independent constituents depends on the macroscopical configuration of the constituents only, not on their microscopic properties. Therefore we can reconstruct the entropy of the whole system, in all possible configurations, just by knowing the entropy of its macroscopic parts. At the same time, property (2) is related to Einstein's likelihood principle [19].

From a mathematical point of view, the composability axiom is equivalent to the requirement that  $\phi(x, y)$  is a group law in the sense of formal group theory [20]. This is the origin of the group theoretical structure associated with the class of generalised entropies called group entropies [1–3]. To be precise, a *group entropy* is an entropic function satisfying the first three Shannon-Khinchin axioms and the composability axiom for all possible probability distributions. In this case the entropy is said to be composable in a strong sense. If an entropy is only composable on the uniform distribution, it is said to be weakly composable.

Thus, the connection between generalised entropies and group theory crucially relies on the composability axiom. Interestingly enough, the study of the algebraic structure defined by the family of power series  $\phi(x, y)$  fulfilling the previous requirements has been developed in a completely different context, namely algebraic topology, during the second half of the past century. Here all we need to state is simply that a *one-dimensional formal group law* over a ring  $R$  [20] is a formal power series in two variables of the form

$$\phi(x, y) = x + y + \sum_{ij} a_{ij} x^i y^j \quad (6)$$

That satisfies the properties (i)–(iii). The theory of formal groups was introduced by Bchner in the seminal paper [21] and developed in algebraic topology, analysis, and other branches of pure and applied mathematics by G. Faltings, S. P. Novikov, D. Quillen, J. P. Serre and many others [20,22]. For recent applications in number theory, see also [23,24].

A property crucial for the subsequent discussion is the following: given a one-dimensional formal group law  $\phi(x, y)$  over a ring of characteristic zero, there exist a unique series  $G(t) = t + \sum_{k=2}^{\infty} \beta_k t^k$  such that

$$\phi(x, y) = G(G^{-1}(x) + G^{-1}(y)). \quad (7)$$

The relation between group entropies and formal group laws is therefore immediate. Indeed, a group entropy possesses a group law associated with it, expressed by a suitable function  $\phi(x, y)$  of the form (6) which is responsible for the composition process for any possible choice of the probability distributions on  $A$  and  $B$ . A natural question is how to classify group entropies. To this aim, we recall that, generally speaking, we can distinguish between two large classes of entropy functions, the trace-form class and the non-trace-form one. In the first case, we shall deal with entropies that can be written as  $S = \sum_i f(p_i)$  for a suitable one-variable function  $f(x)$ . The prototype of this family is Boltzmann's entropy. If an entropy cannot be written in this way, it is said to be a non-trace-form one. The most well-known example of a non-trace-form entropy is Rényi's entropy. In this paper we shall focus on the following realizations of the two classes.

For the *trace-form class*, we shall analyse the general functional [3]

$$S[p] = \sum_{i=1}^W p_i G\left(\ln \frac{1}{p_i}\right) \quad (8)$$

called the universal-group entropy (since it is related with the algebraic structure called Lazard's universal formal group). Here  $G(t)$  is an arbitrary real analytic invertible function such that  $G(0) = 0$ ,  $G'(0) = 1$ .

For the *non-trace-form class* we shall consider the functional [3]

$$S[p] = \frac{G\left(\ln\left(\sum_{i=1}^W p_i^\alpha\right)\right)}{1 - \alpha} \quad (9)$$

that has been called Z-entropy. Both families of entropies are assumed to satisfy the first three Shannon-Khinchin axioms for suitable choices of  $G(t)$ . The main difference between the trace-form and the non-trace-form class is encoded in a theorem proved in [18], stating that the most general trace-form entropy satisfying Equation (2) is Tsallis entropy, with Boltzmann's entropy as an important special case. The infinitely many other possible trace-form entropies only fulfil the composition law Equation (2) on the restricted part of probability space consisting of uniform probabilities  $p_{i,j}^{A \times B} = 1/W_{A \times B} = 1/(W_A W_B) = p_i^A p_j^B$ . Therefore, these entropies are said to be weakly composable [3]. Instead, the non-trace-form entropy (9) is composable for any combination  $A \times B$  of systems  $A$  and  $B$  with  $p_{i,j}^{A \times B} = p_i^A p_j^B$ .

### 3. From Phase Space Volume to Group Entropies

Extensivity and the dependence on the size of phase space have often played a role in the analysis of entropies. For the case of Tsallis entropy, the requirement of extensivity is used to determine the value of the parameter  $q$  [8,25]; the importance of the dependence of the entropy on the available number of microstates  $W$  was discussed in [26]. Here we describe how exploiting the relation between the number of microstates  $W$  and the number of particles  $N$  allows one to find the functional form of the group entropies, see [11,17]. For a discussion not assuming the composability requirement and hence nor the group structure, see [27,28].

We consider how the group-theoretic entropies deal with the three asymptotic dependencies of the phase space volume

(I)	Algebraic	$W(N) = N^a$	with	$W^{-1}(t) = t^{\frac{1}{a}},$
(II)	Exponential	$W(N) = k^N$	with	$W^{-1}(t) = \frac{\ln t}{\ln k},$
(III)	Super-exponential	$W(N) = N^{\gamma N}$	with	$W^{-1}(t) = \exp[L(\frac{\ln t}{\gamma})].$

Here  $L(t)$  denotes the Lambert function.

Now we shall discuss how the extensivity requirement and the functional form of  $W(N)$  determine the function  $G(t)$ , which in turn characterizes the entropy according to formulae (8) and (9). Preliminary, before entering into technical details let us clarify how the present theory relates to previous investigation.

First, what could make the entropy non-additive? For the exponential case (II) we will find that the composition in Equation (2) corresponds to simple addition  $\phi(x, y) = x + y$ . This is the traditional Boltzmann-Shannon case. All four S-K axioms, including the 4th additivity axiom, are satisfied and in accordance with the uniqueness theorem [6] we find  $S[p] = -\sum_i p_i \ln p_i$ . So, as one could expect, an exponential-type phase space volume is related with additivity and no essential emergence of interdependence among the components of the considered system. The situation turns out to be different for the cases (I) and (III) above. In both these cases  $W(AB) \neq W(A \times B) = W(A)W(B)$ . In the sub-exponential case (I) the fully interdependent system  $AB$  has fewer states available than  $W(A)W(B)$ . This situation is akin to how the Pauli principle prevents a set of fermions from occupying all possible combinations of single particle states. Instead, in case (III) the system  $AB$  has more states available than  $W(A)W(B)$ , new collective states have emerged when  $A$  and  $B$  are combined [11].

Lieb and Yngvason has argued [29] that from standard classical thermodynamics, without any use of statistical mechanics, it follows that entropy must be additive and extensive. We recall that the fourth Shannon-Khinchin [1] axiom assumes additivity and since the four SK axioms uniquely lead to the Boltzmann-Shannon functional form we can only be consistent with traditional thermodynamics if we remain within the Shannon-Khinchin axiomatic framework. This implies that only case (II)  $W(N) = k^N$  is consistent with traditional thermodynamics. The two cases (I)  $W(N) = N^a$  and  $W(N) = N^{\gamma N}$  turns out not to be consistent with additivity, which takes one outside the framework of Boltzmann-Shannon-Khinchin and therefore in accordance with Lieb and Yngvason outside standard thermodynamics [8,30,31] i.e., we are naturally lead to the abstract conceptual framework of information theory. We wish to stress that group entropies represent measures of complexity by information geometric means [32] and can characterise limiting probability distributions by means of a maximum entropy procedure for systems where interdependence among its components makes  $W(N)$  deviate from the exponential form.

Stepping outside the SK framework can of course be done in multiple ways. One may simply decide to entirely give up on the 4th axiom and only assume the first three. This approach was considered in [26,27]. The group theoretic approach described here is of course related, however, importantly, it requires that a entropy must be defined in a way that allows the computation of the entropy of the independent combination  $A \times B$  to be related in a consistent and unique way to the entropy of the parts  $A$  and  $B$ .

### 3.1. From $W(N)$ to $G(t)$

We start from the requirement that the group entropy is extensive on the equal probability ensemble  $p_i = 1/W$ , i.e., we require asymptotically for large  $N$ , and therefore large  $W$  (here we are assuming that  $W(N)$  is a monotonically increasing function of  $N$ ) that

$$S\left(p_i = \frac{1}{W}\right) = \lambda N. \tag{10}$$

We now consider separately the trace form case (8) and the non-trace form (9) one. For the first case, we have asymptotically

$$S\left(\frac{1}{W}\right) = \sum_{i=1}^W \frac{1}{W} G(\ln(W)) \approx \lambda N. \tag{11}$$

Inverting the relation between  $S$  and  $G$ , which by Equation (11) amounts to inverting the relation between  $G$  and  $N$ , we obtain

$$G(t) \approx \lambda W^{-1}[\exp(t)]. \tag{12}$$

This is a consequence of the asymptotic extensivity. However, we also need  $G(t)$  to generate a group law, which requires  $G(0) = 0$  [1,3], so we adjust the expression for  $G(t)$  in Equation (12) accordingly and conclude

$$G(t) = \lambda \{W^{-1}[\exp(t)] - W^{-1}(1)\}. \tag{13}$$

Assuming the non-trace form in Equation (9) when inverting Equation (10), and ensuring  $G(0) = 0$  leads to

$$G(t) = \lambda(1 - \alpha) \{W^{-1}[\exp(\frac{t}{1 - \alpha})] - W^{-1}(1)\}. \tag{14}$$

Assuming that  $W(N)$  is sufficiently regular, it is easy to see that the simple choice

$$\lambda = 1/(W^{-1})'(1)$$

in both cases makes  $G$  of the form  $G(t) = t + O(t^2)$ .

From the expressions (13) and (14) we can now list the entropies corresponding to the three classes (I), (II) and (III) of phase space growth rates. A straightforward calculation gives the following results:

*Trace-form case*

(I) Algebraic,  $W(N) = N^a$ :

$$S[p] = a \sum_{i=1}^{W(N)} p_i \left[ \left(\frac{1}{p_i}\right)^{\frac{1}{a}} - 1 \right] \tag{15}$$

$$= \frac{1}{q-1} (1 - \sum_{i=1}^{W(N)} p_i^q). \tag{16}$$

To emphasize the relation with the Tsallis  $q$ -entropy, we have introduced  $q = 1 - 1/a$ . Please note that the parameter  $q$  is determined by the exponent  $a$ , so it is controlled entirely by  $W(N)$ .

(II) Exponential,  $W(N) = k^N, k > 0$ :

$$S[p] = \sum_{i=1}^{W(N)} p_i \ln \frac{1}{p_i}. \tag{17}$$

This is of course the Boltzmann-Gibbs case.

(III) Super-exponential,  $W(N) = N^{\gamma N}, \gamma > 0$ :

$$S[p] = \sum_{i=1}^{W(N)} p_i \left\{ \exp \left[ L \left( -\frac{\ln p_i}{\gamma} \right) \right] - 1 \right\}. \tag{18}$$

*Non-trace-form case*

(I) Algebraic,  $W(N) = N^a$ :

$$S[p] = \left\{ \exp \left[ \frac{\ln(\sum_{i=1}^{W(N)} p_i^a)}{a(1 - \alpha)} \right] - 1 \right\}. \tag{19}$$

- (II) Exponential,  $W(N) = k^N$ :

$$S[p] = \frac{\ln(\sum_{i=1}^{W(N)} p_i^\alpha)}{1 - \alpha}. \tag{20}$$

This is of course the Rényi entropy.

- (III) Super-exponential,  $W(N) = N^{\gamma N}$ :

$$S[p] = \left\{ \exp \left[ L \left( \frac{\ln \sum_{i=1}^{W(N)} p_i^\alpha}{\gamma(1 - \alpha)} \right) \right] - 1 \right\}. \tag{21}$$

This entropy was recently studied in relation to a simple model in which the components can form emergent paired states in addition to the combination of single particle states [11].

### 3.2. The Composition Law $\phi(x, y)$

We now derive the composition law introduced in Equation (2) above. The composition is given in terms of the function  $G(t)$  as in [3,4] according to the relations

Trace-form case [33]

$$\phi(x, y) = G[G^{-1}(x) + G^{-1}(y)]. \tag{22}$$

Non-trace-form case

$$\phi(x, y) = \frac{1}{1 - \alpha} G[G^{-1}((1 - \alpha)x) + G^{-1}((1 - \alpha)y)]. \tag{23}$$

When we express  $\phi(x, y)$  directly in terms of the phase space volume  $W(N)$  by use of Equations (13) and (14) we arrive at the following expression valid for both trace and non-trace forms

$$\phi(x, y) = \lambda \left\{ W^{-1} \left[ W \left( \frac{x}{\lambda} + W^{-1}(1) \right) W \left( \frac{y}{\lambda} + W^{-1}(1) \right) \right] - W^{-1}(1) \right\} \tag{24}$$

where  $\lambda = 1/(W^{-1})'(1)$ .

To obtain from Equation (24) specific realisations of  $\phi(x, y)$  for the three phase space growth rates, we substitute the appropriate expressions for  $W(N)$  and  $W^{-1}(t)$  and obtain the following results.

- (I) Algebraic,  $W(N) = N^a$ :

$$\phi(x, y) = x + y + \frac{1}{\lambda} xy = x + y + (1 - q)xy. \tag{25}$$

The case of Tsallis and and Sharma-Mittal entropies (see also [32] for new examples).

- (II) Exponential,  $W(N) = k^N$ :

$$\phi(x, y) = x + y. \tag{26}$$

The Boltzmann and Rényi case.

- (III) Super-exponential,  $W(N) = N^{\gamma N}$ :

$$\phi(x, y) = \lambda \left\{ \exp \left[ L \left( \left( 1 + \frac{x}{\lambda} \right) \ln \left( 1 + \frac{x}{\lambda} \right) + \left( 1 + \frac{y}{\lambda} \right) \ln \left( 1 + \frac{y}{\lambda} \right) \right) \right] - 1 \right\} \tag{27}$$

For examples of models relevant to this growth rate and composition law see [11,28].



#### 4. Maximum Entropy Ensembles

Let us now consider the probability distributions derived from the group entropies by maximizing them under very simple constraints. As usual, we shall introduce the constraints by means of Lagrange multipliers and analyse the functional

$$J[p] = S[p] - \sum_{n=1}^M \lambda_n g_n[p], \tag{28}$$

with  $M$  constraints expressed by the functionals  $g_n[p]$ . Traditionally, one uses the first constraint to control the normalization

$$g_1[p] = \sum_{i=1}^W p_i - 1 \tag{29}$$

and the second one to determine the average of some observable  $E$ . In physics, this observable is typically the average of the systems energy  $\bar{E} = \langle E \rangle - E_0$  measured from the ground state level  $E_0$

$$g_2[p] = \sum_{i=1}^W (p_i(E_i - E_0) - \bar{E}). \tag{30}$$

With these two constraints, from the extremal condition  $\delta J / \delta p_i = 0$  we obtain

$$\frac{\delta S[p]}{\delta p_i} = \lambda_1 + \lambda_2(\Delta E_i - \bar{E}). \tag{31}$$

Here  $\Delta E_i = E_i - E_0$ .

##### 4.1. Trace-Form Entropies

The derivatives of  $S[p]$  for the three trace-form ones in Equations (16)–(18) are

(I) Algebraic –  $W(N) = N^a$ :

$$\frac{\delta S}{\delta p_i} = a \left(1 - \frac{1}{a}\right) \left(\frac{1}{p_i}\right)^{\frac{1}{a}}. \tag{32}$$

(II) Exponential –  $W(N) = k^N$ :

$$\frac{\delta S}{\delta p_i} = \left(\ln \frac{1}{p_i} - 1\right). \tag{33}$$

(III) Super-exponential –  $W(N) = N^{\gamma N}$ :

$$\frac{\delta S}{\delta p_i} = \exp [L(X_i)] - \frac{1}{\gamma} \frac{1}{1 + L(X_i)} - 1, \text{ with } X_i = \frac{1}{\gamma} \ln \frac{1}{p_i}. \tag{34}$$

The functional form of  $\delta S / \delta p_i$  in Equations (32) and (33) allows us to solve Equation (31) to express  $p_i$  as

(I) Algebraic –  $W(N) = N^a$ :

$$p_i = \frac{[1 + \beta(\Delta E_i - \bar{E})]^{-a}}{Z}, \tag{35}$$

where formally  $\beta = \lambda_2 / \lambda_1$  and  $Z = \sum_i [1 + \beta(\Delta E_i - \bar{E})]^{-a}$ . These probabilistic weights correspond to Tsallis’s q-exponentials.

(II) Exponential –  $W(N) = k^N$ :

$$p_i = \frac{\exp[-\beta(\Delta E_i - \bar{E})]}{Z} \tag{36}$$

where formally  $\beta = \lambda_2/\lambda_1$  and  $Z = \sum_i \exp[-\beta(\Delta E_i - \bar{E})]$ . As expected, we have re-derived the Boltzmann weights starting from the trace form of the group entropies and exponential phase space growth rates.

The transcendental nature of the expression for  $\delta S/\delta p_i$  in Equation (34) seems to prevent one from deriving a closed formula for  $p_i$  in the case of super-exponential phase space growth rate  $W(N) = N^{\gamma N}$ , having assumed the trace-form (weakly composable) expression (18) for the entropy. We shall see below that the situation is different when starting from non-trace-form entropies.

#### 4.2. Non-Trace-Form Entropies

The form of the entropies for the non-trace case given in Equations (19)–(21) all lead to the same functional expression as when starting from the trace-form algebraic case in Equation (35), namely

$$p_i = \frac{[1 + \beta(\Delta E_i - \bar{E})]^{-\frac{1}{1-\alpha}}}{Z} \quad (37)$$

where formally  $\beta = \lambda_2/\lambda_1$  and  $Z = \sum_i [1 + \beta(\Delta E_i - \bar{E})]^{-\frac{1}{1-\alpha}}$ . This expression for  $p_i$  is reminiscent of the Tsallis q-exponential.

### 5. Discussion

We have seen that the group theoretic entropies offer a systematic classification of entropies according to how the phase space grows with number of particles. The formalism allows for a systematic generalisation of a statistical mechanics description to non-exponential phase spaces and reduces to the standard Boltzmann-Gibbs picture when  $W(N)$  is an exponential. A new measure of complexity, see Equations (5) and (24), determined by the phase space volume's dependence on system size follows right away.

We wish to point out that group entropies represent an interesting new tool in information geometry, since they can be used to construct Riemannian structures in statistical manifolds via suitable divergences (or relative entropies) [34] associated with them. This has been proved in [32] for Z-entropies and in [35] for the universal-group entropy. Also, a quantum version of these entropies can be used to define a family of entanglement measures for spin chains [1]. Work is in progress along all these lines.

**Author Contributions:** H.J.J. and P.T. developed the proposed research, wrote the paper and reviewed it in a happy and constructive collaboration.

**Funding:** This research of P.T. has been partly supported by the research project FIS2015-63966, MINECO, Spain, and by the ICMAT Severo Ochoa project SEV-2015-0554 (MINECO).

**Conflicts of Interest:** The authors declare no conflict of interest.

### References and Notes

1. Tempesta, P. Formal groups and Z-entropies. *Proc. R. Soc. A* **2016**, *472*, 20160143. [[CrossRef](#)] [[PubMed](#)]
2. Tempesta, P. Group entropies, correlation laws and zeta functions. *Phys. Rev. E* **2011**, *84*, 021121. [[CrossRef](#)] [[PubMed](#)]
3. Tempesta, P. Beyond the Shannon-Khinchin formulation: The composability axiom and the universal-group entropy. *Ann. Phys.* **2016**, *365*, 180–197. [[CrossRef](#)]
4. Tempesta, P. A theorem on the existence of trace-form entropies. *Proc. R. Soc. A* **2015**, *471*, 20150165. [[CrossRef](#)]
5. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [[CrossRef](#)]
6. Khinchin, A.I. *Mathematical Foundation of Information Theory*; Dover: Mineola, NY, USA, 1957.

7. (1)  $S[p]$  is a continuous function of the  $p_i$ , (2)  $S[p]$  is maximized by the uniform distribution  $p_i = 1/W$ , (3) entropy is left unchanged if a state of zero probability is added, namely  $S[p, 0] = S[p]$  and (4) entropy is additive, see e.g., [1,26]
8. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics*; Springer: Berlin, Germany, 2009.
9. Ellis, R. *Entropy, Large Deviation, and Statistical Mechanics*; Springer: Berlin, Germany, 1985.
10. Touchette, H. The large deviation approach to statistical mechanics. *Phys. Rep.* **2009**, *478*, 1–69. [[CrossRef](#)]
11. Jensen, H.J.; Pazuki, R.H.; Pruessner, G.; Tempesta, P. Statistical mechanics of exploding phase spaces: Ontic open systems. *J. Phys. A Math. Theor.* **2018**, *51*, 375002. [[CrossRef](#)]
12. Hendry, R.F. Ontological reduction and molecular structure. *Stud. Hist. Philos. Mod. Phys.* **2010**, *41*, 183–191. [[CrossRef](#)]
13. Morin, E. *On Complexity*; Hampton Press: New York, NY, USA, 2008.
14. Reif, F. *Fundamentals of Statistical and Thermal Physics*; McGraw-Hill: New York, NY, USA, 1965.
15. Tononi, G. Information measures for conscious experience. *Arch. Ital. Biol.* **2001**, *139*, 367–371. [[PubMed](#)]
16. Balduzzi, D.; Tononi, G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput. Biol.* **2008**, *4*, e1000091. [[CrossRef](#)] [[PubMed](#)]
17. Tempesta, P.; Jensen, H.J. Group Entropies and Measures of Complexity. In Preparation.
18. Enciso, A.; Tempesta, P. Uniqueness and characterization theorems for generalized entropies. *J. Stat. Mech. Theory Exp.* **2017**, *2017*, 123101. [[CrossRef](#)]
19. Sicuro, G.; Tempesta, P. Groups, information theory, and Einstein's likelihood principle. *Phys. Rev. E* **2016**, *93*, 040101(R).
20. Hazewinkel, M. *Formal Groups and Applications*; Academic Press: Cambridge, MA, USA, 1978.
21. Bochner, S. Formal Lie groups. *Ann. Math.* **1946**, *47*, 192–201. [[CrossRef](#)]
22. Bukhshtaber, V.M.; Mishchenko, A.S.; Novikov, S.P. Formal groups and their role in the apparatus of algebraic topology. *Russ. Math. Surv.* **1971**, *26*, 63–90. [[CrossRef](#)]
23. Tempesta, P. L-series and Hurwitz zeta functions associated with the universal formal group. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2010**, *9*, 133–144.
24. Tempesta, P. The Lazard formal group, universal congruences and special values of zeta functions. *Trans. Am. Math. Soc.* **2015**, *367*, 7015–7028. [[CrossRef](#)]
25. Tsallis, C.; Gell-Mann, M.; Sato, Y. Asymptotically scale-invariant occupancy of phase space makes the entropy  $S_q$  extensive. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15377–15382. [[CrossRef](#)] [[PubMed](#)]
26. Hanel, R.; Thurner, S. A comprehensive classification of complex statistical systems and an axiomatic derivation of their entropy and distribution functions. *EPL (Europhys. Lett.)* **2011**, *93*, 20006. [[CrossRef](#)]
27. Hanel, R.; Thurner, S. When do generalized entropies apply? How phase space volume determines entropy. *EPL* **2011**, *96*, 50003. [[CrossRef](#)]
28. Korbel, J.; Hanel, R.; Thurner, S. Classification of complex systems by their sample-space scaling exponents. *New J. Phys.* **2018**, *20*, 093007. [[CrossRef](#)]
29. Lieb, E.H.; Yngvason, J. The mathematical structure of the second law of thermodynamics. *Curr. Dev. Math.* **2001**, *2001*, 89–129. [[CrossRef](#)]
30. Mathematical structures for generalised thermodynamics have been considered by several authors, see e.g., Tsallis [8] and Naudts [31].
31. Naudts, J. *Generalised Thermostatistics*; Springer: Berlin, Germany, 2011.
32. Rodríguez, M.A.; Romaniega, A.; Tempesta, P. A new class of entropic information measures, formal group theory and information geometry. *arXiv* **2018**, arxiv:1807.01581.
33. We remark that for trace-form entropies, formula (22) is valid in general only when the systems A and B we wish to combine are both in a state described by the uniform distribution (weak composability), with the remarkable exception of Tsallis entropy, which is composable on all possible probability distributions.
34. Amari, S.I. *Information Geometry and Its Applications, Applied Mathematical Sciences*; Springer: Berlin, Germany, 2016.
35. Gomez, I.S.; Borges, E.P.; Portesi, M. Fisher metric from relative entropy group. *arXiv* **2018**, arxiv:1805.11157.

