Internat. J. Math. & Math. Sci.
Vol. 23, No. 4 (2000) 243–251
S0161171200000375

BOLTZMANN-GIBBS ENTROPY: AXIOMATIC CHARACTERIZATION AND APPLICATION

C. G. CHAKRABARTI and KAJAL DE

(Received 15 August 1997 and in revised form 10 April 1998)

ABSTRACT. We present axiomatic characterizations of both Boltzmann and Gibbs entropies together with an application.

Keywords and phrases. Degree of disorder, Boltzmann entropy, Gibbs entropy, axiomatic characterization, total entropy.

2000 Mathematics Subject Classification. Primary 82B03.

1. Introduction. The concept of entropy is fundamental in the foundation of statistical physics. It first appeared in thermodynamics through the second law of thermodynamics. The notion of entropy has been broadened by the advent of statistical mechanics and has been still further broadened by the later advent of information theory.

In statistical mechanics, Boltzmann was the first to give a statistical or a probabilistic definition of entropy. Boltzmann entropy is defined for a macroscopic state of a system while Gibbs entropy is defined over an ensemble, that is over the probability distribution of macrostates. Both Boltzmann and Gibbs entropies are, in fact, the pillars of the foundation of statistical mechanics and are the basis of all the entropy concepts in modern physics. A lot of work on the mathematical analysis and practical applications of both Boltzmann and Gibbs entropies was done [13], yet the subject is not closed, but is open awaiting a lot of work on their characterization, interpretation, and generalization.

In this paper, we have considered three problems. The first is the axiomatic characterization of Boltzmann entropy. Basis of this characterization is the Carnop's notion of "degree of disorder" [2]. The second is concerned with the derivation of Gibbs entropy from Boltzmann entropy, its generalization, and axiomatic characterization and the third deals with the derivation of Boltzmann entropy for classical system consistent with the above formalism.

2. Boltzmann entropy: axiomatic characterization. In statistical mechanics, we are interested in the disorder in the distribution of the system over the permissible microstates [1]. The measure of disorder first provided by Boltzmann principle (known as Boltzmann entropy) is given by

$$S = K \ln W, \tag{2.1}$$

where K is the thermodynamic unit of measurement of entropy and is known as

Boltzmann constant. $K = 1.33 \times 10^{-16}\,\text{erg}/{}^\circ\text{C}$. W, called thermodynamic probability or statistical weight, is the total number of microscopic complexions compatible with the macroscopic state of the system. We avoid the name thermodynamic probability for the term W as it leads to many confusions [10]. Following Carnop [2], we, however, call the quantity "W" the "degree of disorder". Let us now derive the expression (2.1) axiomatically which is independent of any micromodel of the system, classical or quantal. For this, we make the following two axioms in conformity with our intuitions [9].

Axiom (i). The entropy S of the system is a monotonic increasing function of "degree of disorder" W, that is, $S(W) \le S(W+1)$.

Axiom (ii). The entropy S is assumed to be an additive function for two statistically independent systems with degrees of disorder W_1 and W_2, respectively. The entropy of the composite system is given by $S(W_1 \cdot W_2) = S(W_1) + S(W_2)$.

Theorem 2.1. *If the entropy function $S(W)$ satisfies the above two axioms (i) and (ii), then $S(W)$ is given by*

$$S = K \ln W, \tag{2.2}$$

where K is a positive constant depending on the unit of measurement of entropy.

Proof. Let us assume that $W > e$. Then, for any integer n, we can define the integer $m(n)$ such that

$$e^{m(n)} \le W^n \le e^{m(n)+1} \tag{2.3}$$

or

$$\frac{m(n)}{n} \le \ln W \le \frac{m(n)+1}{n}. \tag{2.4}$$

Consequently,

$$\lim_{n\to\infty} \frac{m(n)}{n} = \ln W. \tag{2.5}$$

Let S denote a function that satisfies axioms (i) and (ii), then, by axiom (i),

$$S(W_1) \le S(W_2) \tag{2.6}$$

for $W_1 < W_2$. Combining (2.3) with (2.6), we get

$$S(e^{m(n)}) \le S(W^n) \le S(e^{m(n)+1}). \tag{2.7}$$

By axiom (ii), we get

$$S(W^n) = nS(W). \tag{2.8}$$

Here,

$$S(e^{m(n)}) = m(n)S(e), \tag{2.9}$$

$$S(e^{m(n)+1}) = (m(n)+1)S(e). \tag{2.10}$$

Applying (2.8) and (2.9) to the inequality (2.7), we get

$$m(n)S(e) \le nS(W) \le (m(n)+1)S(e) \tag{2.11}$$

and consequently,

$$\lim_{n\to\infty} \frac{m(n)}{n} = \frac{S(W)}{S(e)}. \tag{2.12}$$

Comparing (2.12) with (2.5), we conclude that

$$S(W) = K \ln W, \tag{2.13}$$

where $K = S(e)$ is a positive constant depending on the unit of measurement of entropy. Thus, if the number of microscopic complexions or statistical weight W is taken as the measure of the "degree of disorder", then Boltzmann entropy provides the interrelation between the concepts of disorder and entropy. □

The Boltzmann entropy (2.1) and its axiomatic derivation need some further explanation. Boltzmann entropy plays a crucial role in the connection of the nonmechanical science of thermodynamics with mechanics through statistics. In spite of its great importance, little attention is paid to the derivation of this entropy rigorously. In thermodynamics, two fundamental properties of Boltzmann entropy are (see [11]):

(i) its nondecrease: if no heat enters or leaves a system, its entropy cannot decrease;

(ii) its additivity: the entropy of two systems, taken together, is the sum of their separate entropies.

However, in statistical mechanics of finite system, it is impossible to satisfy both the properties exactly. It is only for an infinitely large system that the properties can be reconciled and an appropriate expression of entropy can be derived [11]. In most of the books on statistical physics [11], the expression of Boltzmann entropy is derived from the additivity property of entropy. But this is not correct. The axioms we have assumed are variant forms of the above thermodynamic properties (i) and (ii), and the method of derivation of the Boltzmann entropy is in conformity with the statistical mechanical requirement of an infinitely large system and Hartley measure of information [9].

3. Boltzmann-Gibbs entropy: axiomatic characterization. Boltzmann entropy S is defined for a macroscopic state whereas Gibbs entropy is defined over a statistical ensemble. It is, in fact, the entropy of a coarse-grained distribution and can be expressed in terms of the probability distribution of the observational states of the system. More formally, let us introduce a more general definition of Gibbs entropy and call it "Generalized Boltzmann-Gibbs Entropy" or simply, BG entropy [8].

Definition. Let $\Omega = \{A_1, A_2, \dots, A_n\}$ be a denumerable set of macroscopic or observable states of a system, $\wp$ a σ-algebra of the elements of Ω, and $W(A)$, $(A \in \wp)$ is a measure on $\wp$. The measure W is uniquely defined by its values $W(A_n)$, $(n = 0, 1, 2, \dots)$. We call the measure $W(A_n)$ the statistical weight or the structure function of the observational state A_n which is nothing but the total number of microscopic complexions compatible with the observational state A_n in the definition of Boltzmann entropy. We consider the class P_n of all probability measures $P(A)$ on $\wp$ absolutely continuous with respect to W

$$P_n = \Big\{P(A),\ A \in \wp \mid P(\Omega) = 1,\ P \ll w\Big\}. \tag{3.1}$$

By Radon-Nikodym theorem, we may write

$$P(A) = \sum_{A_n \varepsilon A} \rho(A_n) W(A_n) \quad \text{for all } A \in \wp \text{ and } P \in P_n, \tag{3.2}$$

where

$$\rho_n = \frac{P(A_n)}{W(A_n)} \quad (n = 0, 1, 2, \dots). \tag{3.3}$$

This is analogous to the coarse-grained density of microstates defined over the phase-space of the usual definition of Gibbs entropy. Then, we define BG entropy as

$$\overline{S} = -\sum_{A_n} P(A_n) \ln \frac{P(A_n)}{W(A_n)}. \tag{3.4}$$

The BG entropy $\overline{S}$, defined by (3.4), reduces to the form of Boltzmann entropy

$$S = K \ln W(A_n), \tag{3.5}$$

if the probability distribution $P(A_n)$ is so sharp that $P(A_n) = 1$ for a particular event or observational state A_n. Thus, the BG entropy becomes identical to Boltzmann entropy when we know in which observational or microscopic state A_n the system belongs to. The Boltzmann entropy is thus a boundary value of BG entropy.

Now let us turn to the axiomatic derivation of the BG entropy $\overline{S}$ given by (3.4). Let us assume that the entropy of the system is given by

$$\overline{S} = K \sum_i P(A_i) f_i[P(A_i)], \tag{3.6}$$

where the uncertainty function f_i is a continuous function of its argument and is determined on the basis of some axioms or postulates. We make the following two axioms.

Axiom (i). For the simultaneous events of the two systems having observational states A_i and A_j with weights $W(A_i)$ and $W(A_j)$, respectively, we make the additivity postulate

$$f_{ij}(P(A_i) \cdot P(A_j)) = f_{i0}(P(A_i)) + f_{0j}(P(A_j)), \tag{3.7a}$$

where

$$0 \le P(A_i), \qquad P(A_j) \le 1. \tag{3.7b}$$

Axiom (ii). The uncertainty function f_{i0} and f_{0j} associated with the observational states A_i and A_j having weights $W(A_i)$ and $W(A_j)$, respectively, satisfy the boundary conditions

$$f_{i0}(1) = \ln W(A_i), \qquad f_{0j}(1) = \ln W(A_j). \tag{3.8}$$

Axiom (ii) takes account of the post-observational uncertainty, i.e., the uncertainty or the disorder of the microstates, measured by Boltzmann entropy, that remains even after the system has been found in a microscopic state. Then, we have the following theorem.

Theorem 3.1. *If the entropy function $\overline{S}$, given by (3.6), satisfy the above two axioms, then $\overline{S}$ is given by*

$$\overline{S} = -\sum_{A_n} P(A_n) \ln \frac{P(A_n)}{W(A_n)}. \tag{3.9}$$

Proof. First let us take $P(A_j) = 1$. Then, (3.7a) and (3.7b) give

$$f_{ij}(P(A_i)) = f_{i0}(P(A_i)) + f_{0j}(1) = f_{i0}(P(A_i)) + \ln W(A_j). \tag{3.10}$$

Similarly, taking $P(A_i) = 1$, we get

$$f_{ij}(P(A_j)) = f_{0j}(P(A_j)) + \ln W(A_j). \tag{3.11}$$

So, we have

$$\begin{aligned} f_{i0}(P(A_i)) &= f_{ij}(P(A_i)) - \ln W(A_j), \\ f_{0j}(P(A_j)) &= f_{ij}(P(A_j)) - \ln W(A_i). \end{aligned} \tag{3.12}$$

Then we take $P(A_i) = P(A_j) = 1$. Then, (3.7a) and (3.7b) give

$$f_{ij}(1) = \ln W(A_i) + \ln W(A_j). \tag{3.13}$$

Thus, from (3.7a), (3.7b), and (3.13), we can write

$$f_{ij}(P(A_i) \cdot P(A_j)) = [f_{ij} P(A_i) - \ln W(A_j)] + [f_{ij} P(A_j) - \ln W(A_i)], \tag{3.14a}$$

where $0 \leq P(A_i), P(A_j) \leq 1$,

$$f_{ij}(1) = \ln W(A_i) + \ln W(A_j). \tag{3.14b}$$

Writing

$$\begin{aligned} f_{ij}(P(A_j)) &= F(P(A_i)) + \ln W(A_i) + \ln W(A_j), \\ f_{ij}(P(A_j)) &= F(P(A_j)) + \ln W(A_i) + \ln W(A_j), \\ f_{ij}[P(A_i) \cdot P(A_j)] &= F[P(A_i) \cdot P(A_j)] + \ln W(A_i) + \ln W(A_j), \end{aligned} \tag{3.15}$$

the system of equations reduces to Cauchy's functional equation

$$F(x \cdot y) = F(x) + F(y) \tag{3.16a}$$

with boundary conditions

$$0 < x, \qquad y \leq 1, \qquad F(1) = 0, \tag{3.16b}$$

where

$$x = P(A_i), \qquad y = P(A_j). \tag{3.17}$$

The solution of the functional equation (3.16a), subject to the boundary condition (3.16b), is given by $F(x) = -\ln x$. Hence,

$$f_{ij}(P(A_i)) = -\ln P(A_i) + \ln W(A_i) + \ln W(A_j). \tag{3.18}$$

Using (3.13), we get

$$f_{i0}(P(A_i)) = \ln \frac{P(A_i)}{W(A_i)}. \tag{3.19}$$

So, the entropy $\overline{S}$, given by (3.6), reduces to the form of (3.4)

$$\overline{S} = -K \sum_{A_n} P(A_n) \ln \frac{P(A_n)}{W(A_n)}. \tag{3.20}$$

We can express BG entropy $\overline{S}$ as

$$\overline{S} = -K \sum_{A_n} P(A_n) \ln P(A_n) + K \sum_{A_n} P(A_n) \ln W(A_n). \tag{3.21}$$

The first term represents the uncertainty about the macroscopic or the observational states of the system. The second term represents the average of Boltzmann entropy (measuring microscopic uncertainty) which remains even when we know the macrostate the system belongs to. Thus, the BG entropy $\overline{S}$, defined by (3.4), measures the total uncertainty or disorder associated with the microscopic and macroscopic states of the system [12]. It is, therefore, an example of total entropy introduced by Jumarie [8] in Information theory. □

4. Boltzmann-Gibbs entropy: classical system. Let us consider a system consisting of N elements (molecules, organisms, etc.) classified into n classes (energy-states, species, etc.) $Q(i = 1,2,\ldots,n)$. Let N_i be the occupation number of the ith class. The macroscopic state of the system is given by the set of occupation number $A_n = (N_1, N_2, \ldots, N_n)$. The statistical weight or degree of disorder of the macrostate A_n is given by

$$W(A_n) = \frac{N!}{\Pi_{i=1}^{n} N_i!}, \tag{4.1}$$

representing the total number of microscopic states or complexions compatible with the constraints the system is subjected to. For large N_i, Boltzmann entropy (2.1), with $W(A_n)$ given by (4.1), reduces to the form of Shannon entropy [9]

$$S = -KN \sum_{i=1}^{n} P_i \ln P_i, \tag{4.2}$$

where $P_i = N_i/N$ is the relative frequency of the ith class or energy state. For large N, it is the probability that a molecule lies in the ith energy-state. Shannon entropy (4.2) is the basis of the subjective formalism of statistical mechanics initiated by Jaynes [7], Ingarden and Kossakowski [6].

Returning to the occupation numbers N_i, the entropy (4.2) can be written as

$$\overline{S} = -K\sum_{i=1}^{n} N_i \ln\left(\frac{N_i}{N}\right) = -K\sum_{i=1}^{n} N_i \ln N_i + KN\ln N. \tag{4.3}$$

The expression (4.3) has, however, a drawback. It does not satisfy the additive or the extensive property of Boltzmann entropy postulated in Section 2. This fallacy is, in fact, due to Gibbs paradox [3, 4]. To remove this fallacy, we have to subtract $K\ln N! \approx KN(\ln n - 1)$ from the right-hand side of (4.3). This leads to the correct expression of Boltzmann entropy of a classical system

$$\overline{S} = -K\sum_{i=1}^{n} N_i \ln N_i + KN, \tag{4.4}$$

which evidently satisfy the additivity property. The subtraction of the term $\ln N!$ from (4.3) is equivalent to omitting the term $N!$ in the expression (4.1) of statistical weight or degrees of disorder. So, the correct expression of the statistical weight of classical system should be

$$W(A_n) = \prod_{i=1}^{n}\left(\frac{1}{N_i!}\right). \tag{4.5}$$

Now, we consider open system model of the classical system. Then, the whole system can be considered as an aggregate of n subsystem in contact with an environment, each subsystem consisting of N_i particles, each having energy E_i $(i = 1,2,\dots,n)$. A subsystem is then characterized by the occupation number N_i, which is now a random variable having probability distribution $P(N_i)$. The entropy of the whole system is then given by BG entropy,

$$\overline{S} = -K\sum_{A_n} P(A_n)\ln\frac{P(A_n)}{W(A_n)}, \tag{4.6}$$

where $W(A_n)$ is the statistical weight of the macrostate $A_n = (N_1, N_2, \dots, N_n)$. For the gaseous system, the subsystems can be considered independent. Hence, the entropy of the whole system can be decomposed into the sum of the entropies of the different constituent subsystems

$$\overline{S} = \sum_{i=1}^{n} \overline{S}_i, \tag{4.7}$$

where

$$\overline{S}_i = -\sum_{N_i} P(N_i)\ln\frac{P(N_i)}{W(N_i)} \tag{4.8}$$

is the entropy of the ith subsystem. The statistical weight $W(N_i)$ is different for different systems. For a classical system, $W(N_i) = 1/N_i!$ $(i = 1,2,\dots,n)$ as we have arrived at earlier. The probability distribution $P(N_i)$ can be estimated by the principle of maximum entropy [7], which consists of the maximization of the entropy (4.8) subject to the condition of fixed average number of particles

$$\sum_{i=1}^{n} P(N_i)N_i = \langle N_i \rangle. \tag{4.9}$$

The maximization of (4.8) subject to the constraint (4.9) and the normalization condition:

$$\sum_{N_i} P(N_i) = 1 \tag{4.10}$$

leads to the Poisson distribution

$$P(N_i) = \frac{e^{-\langle N_i \rangle} \langle N_i \rangle^{N_i}}{N_i!}. \tag{4.11}$$

The entropy of the whole system then reduces to the form

$$\overline{S} = -K \sum_{i=1}^{n} \langle N_i \rangle \ln \langle N_i \rangle + k \sum_{i=1}^{n} \langle N_i \rangle, \tag{4.12}$$

which is the open system analogue of the Boltzmann entropy (4.4) for closed system. The choice of the statistical weight (4.5) needs a further explanation. Ingarden and Kossakowski have considered this type of weight in the quantum statistics of photons for the case when the photons are registered by a counter [6]. They, however, remarked that the choice of the weight (4.5) did not imply that the statistics would be of quantum character. In fact, the statistical weight (4.5), which we have justified from Boltzmann-entropy, plays an important role in the analysis of classical system [3, 4, 5].

5. Conclusion. Boltzmann entropy plays a crucial role not only in the foundation of statistical mechanics, but also in the other branches of science. In view of its great importance, we have tried first to provide an axiomatic characterisation of Boltzmann entropy consistent with the basic properties of thermodynamics. As a generalisation of Boltzmann entropy, we have introduced generalised (BG) entropy and characterised it on the basis of some axioms. The axiomatic approach to BG entropy is completely different from the total entropy introduced by Jumarie [8], which is a generalisation of Shannon entropy to take into account the uncertainty due to the spans of the lattice where a random variable is defined. In the present paper, the statistical weight appears through the boundary condition of the functional equation, the solution of which leads to the form of BG entropy.

References

[1] B. H. Callen, *Thermodynamics and Introduction to Thermostatistics*, Wiley & Sons, New York, 1974.

[2] R. Carnap, *Two Essays on Entropy*, University of California Press, Berkeley, 1977. MR 57#4929. Zbl 399.03002.

[3] C. G. Chakrabarti, *Boltzmann-Gibbs Entropy: Thermodynamics and Statistics of Ecosystems*, Communicated.

[4] ______, *Statistical inference and quantum statistics*, Rev. Bull. Calcutta Math. Soc. **2** (1994), no. 1-2, 39–46, Proceedings of S. N. Bose Birth Centenary International Symposium on Statistical Physics Calcutta, 1993. CMP 1 354 900. Zbl 844.60100.

[5] C. G. Chakrabarti and De Kajal, *Boltzmann entropy: generalization and application*, J. Biol. Phys., in press, 1997.

[6] R. S. Ingarden and A. Kossakowski, *The Poisson probability distribution and information thermodynamics*, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. **19** (1971), 83–85. MR 46 3203. Zbl 217.28905.

[7] E. T. Jaynes, *Information theory and statistical mechanics*, Phys. Rev. (2) **106** (1957), 620–630. MR 19,335b. Zbl 084.43701.

[8] G. Jumarie, *Relative Information: Theories and Applications*, Springer Series in Synergetics, vol. 47, Springer-Verlag, Berlin, 1990, Theories and applications. MR 91h:94012. Zbl 719.94005.

[9] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1988. MR 89e:94028. Zbl 675.94025.

[10] B. H. Lavenda, *Statistical Physics: A Probabilistic Approach*, John Wiley & Sons, Inc., New York, 1991. MR 94a:82002.

[11] O. Penrose, *Foundations of Statistical Mechanics. A Deductive Treatment*, International Series of Monographs in Natural Philosophy, vol. 22, Pergamon Press, Oxford, New York, Toronto, 1970. MR 41#4982. Zbl 193.27801.

[12] S. Watanabe, *Knowing and Guessing*, John Wiley & Sons, Inc., New York, 1969, A quantitative study of inference and information. MR 52 12977. Zbl 206.20901.

[13] A. Wehrl, *General properties of entropy*, Rev. Modern Phys. **50** (1978), no. 2, 221–260. MR 58 14862.

Chakrabarti and De: Department of Applied Mathematics, University of Calcutta, 92 A.P.C. Road, Calcutta-700 009, India