

Article

A Note on Burg's Modified Entropy in Statistical Mechanics

Amritansu Ray ^{1,*} and S. K. Majumder ²

¹ Department of Mathematics, Rajyadharpur Deshbandhu Vidyapith, Serampore, Hooghly 712203, West Bengal, India

² Department of Mathematics, Indian Institute of Engineering Science and Technology (IIST), Shibpur, Howrah 711103, West Bengal, India; majumder_sk@yahoo.co.in

* Correspondence: amritansu_ray06@yahoo.co.in; Tel.: +91-943-295-3102

Academic Editor: Reza Abedi

Received: 3 December 2015; Accepted: 16 February 2016; Published: 27 February 2016

Abstract: Burg's entropy plays an important role in this age of information euphoria, particularly in understanding the emergent behavior of a complex system such as statistical mechanics. For discrete or continuous variable, maximization of Burg's Entropy subject to its only natural and mean constraint always provide us a positive density function though the Entropy is always negative. On the other hand, Burg's modified entropy is a better measure than the standard Burg's entropy measure since this is always positive and there is no computational problem for small probabilistic values. Moreover, the maximum value of Burg's modified entropy increases with the number of possible outcomes. In this paper, a premium has been put on the fact that if Burg's modified entropy is used instead of conventional Burg's entropy in a maximum entropy probability density (MEPD) function, the result yields a better approximation of the probability distribution. An important lemma in basic algebra and a suitable example with tables and graphs in statistical mechanics have been given to illustrate the whole idea appropriately.

Keywords: entropy optimization; probability distribution; Shannon's Entropy; Burg's Entropy

1. Introduction

The concept of entropy [1] figured strongly in the physical sciences during the 19th century, especially in thermodynamics and statistical mechanics [2], as a measure of equilibrium and evolution of thermodynamic systems. Two main views were developed which were the macroscopic view formulated originally by Clausius and Carnot and the microscopic approach associated with Boltzmann and Maxwell. Since then, both the approaches have made introspection in natural thermodynamic and microscopically probabilistic systems possible. Entropy is defined as the measure of a system's thermal energy per unit temperature that is unavailable for doing useful work. Because work is obtained from ordered molecular motion, the amount of entropy is also a measure of molecular disorder or randomness of a system. The concept of entropy provides deep insight into the direction of spontaneous change for many day-to-day phenomena. Now, how entropy was developed by Rudolf Clausius [3] is discussed below.

1.1. Clausius's Entropy

To provide a quantitative measure for the direction of spontaneous change, Clausius introduced the concept of entropy as a precise way of expressing the second law of thermodynamics. The Clausius form of the second law states that spontaneous change for an irreversible process [4] in an isolated

system (that is, one that does not exchange heat or work with its surroundings) always proceeds in the direction of increasing entropy. By the Clausius definition, if an amount of heat Q flows into a large heat reservoir at temperature T above absolute zero, then $\Delta s = \frac{Q}{T}$. This equation effectively gives an alternate definition of temperature that agrees with the usual definition. Assume that there are two heat reservoirs R_1 and R_2 at temperatures T_1 and T_2 . (such as the stove and the block of ice). If an amount of heat Q flows from R_1 to R_2 , then the net entropy change for the two reservoirs is $\Delta s = Q \left(\frac{1}{T_2} - \frac{1}{T_1} \right)$ which is positive, provided that $T_1 > T_2$. Thus, the observation that heat never flows spontaneously from cold to hot is equivalent to requiring the net entropy change to be positive for a spontaneous flow of heat. When the system is in thermodynamic equilibrium, then $dS = 0$, i.e., if $T_1 = T_2$, then the reservoirs are in equilibrium, no heat flows, and $\Delta s = 0$. If the gas absorbs an incremental amount of heat dQ from a heat reservoir at temperature T and expands reversibly against the maximum possible restraining pressure P , then it does the maximum work and $dW = PdV$. The internal energy of the gas might also change by an amount dU as it expands. Then, by conservation of energy, $dQ = dU + PdV$. Because the net entropy change for the system plus reservoir is zero when maximum work is done and the entropy of the reservoir decreases by an amount $dS_{\text{reservoir}} = -\frac{dQ}{T}$, this must be counterbalanced by an entropy increase of $dS_{\text{system}} = \frac{dU + PdV}{T} = \frac{dQ}{T}$ for the working gas so that $dS_{\text{system}} + dS_{\text{reservoir}} = 0$. For any real process, less than the maximum work would be done (because of friction, for example), and so the actual amount of heat dQ' absorbed from the heat reservoir would be less than the maximum amount dQ . For example, the gas could be allowed to expand freely into a vacuum and do no work at all. Therefore, it can be stated that $dS_{\text{system}} = \frac{dU + PdV}{T} \geq \frac{dQ'}{T}$ with, $dQ' = dQ$ in the case of maximum work corresponding to a reversible process. This equation defines S_{system} as a thermodynamic state variable, meaning that its value is completely determined by the current state of the system and not by how the system reached that state. Entropy is a comprehensive property in that its magnitude depends on the amount of material in the system.

In one statistical interpretation of entropy, it is found that for a very large system in thermodynamic equilibrium, entropy S is proportional to the natural logarithm of a quantity Ω corresponding to S and can be realized; that is, $S = K \ln \Omega$, in which K is related to molecular energy. On the other hand, entropy generation analysis [5–11] is used to optimize the thermal engineering devices for higher energy efficiency; it has attracted wide attention to its applications and rates in recent years. In order to access the best thermal design of systems, by minimizing the irreversibility, the second law of thermodynamics could be employed. Entropy generation is a criterion for the destruction of a systematized work. The development of the theory followed two conceptually different lines of thought. Nevertheless, they are symbiotically related, in particular through the work of Boltzmann.

1.2. Boltzmann's Entropy

In addition to thermodynamic (or heat-change) entropy, physicists also study entropy statistically [12,13]. The statistical or probabilistic study of entropy is presented in Boltzmann's law. Boltzmann's equation is somewhat different from the original Clausius (thermodynamic) formulation of entropy. Firstly, the Boltzmann formulation is structured in terms of probabilities, while the thermodynamic formulation does not consist in the calculation of probabilities. The thermodynamic formulation can be characterized as a mathematical formulation, while the Boltzmann formulation is statistical. Secondly, the Boltzmann equation yields a value of entropy S while the thermodynamic formulation yields only a value for the change in entropy (dS). Thirdly, there is a shift in content, as the Boltzmann equation was developed for research on gas molecules rather than thermodynamics. Fourthly, by incorporating probabilities, the Boltzmann equation focuses on microstates, and thus explicitly introduces the question of the relationship between macrostates and microstates. Boltzmann investigated such microstates and defined entropy in a new way such that the macroscopic maximum

entropy state corresponded to a thermodynamic configuration which could be formulated by the maximum number of different microstates. He noticed that the entropy of a system can be considered as a measure of the disorder in the system and that in a system having many degrees of freedom, the number measuring the degree of disorder also measured the uncertainty in a probabilistic sense about the particular microstates.

The value W was originally intended to be proportional to the Wahrscheinlichkeit (means probability) of a macrostate for some probability distribution of a possible microstate, in which the thermodynamic state of a system can be realized by assigning different ξ and ρ of different molecules. The Boltzmann formula is the most general formula for thermodynamic entropy; however, his hypothesis was for an ideal gas of N identical particles, of which $N_i (i = 1, 2, \dots, l)$ are the i th microscopic condition of position and momentum of a given distribution $D_i = (N_1, N_2, \dots, N_l)$. Here, $D_1 = (N, 0, \dots, 0)$, $D_2 = (N - 1, 1, \dots, 0)$, ... etc. For this state, the probability of each microstate system is equal, so it was equivalent to calculating the number of microstates associated with a macrostate. Then the statistical disorder is given by [14] $W = \frac{N!}{N_1! N_2! \dots N_l!}$, $[N_1 + N_2 + \dots + N_l = N]$. Therefore, the entropy given by Boltzmann is: $S_B = K \ln W$ Where $\ln W = \ln \frac{N!}{N_1! N_2! \dots N_l!} =$

$$\ln N! - \sum_{i=1}^l \ln (N_i!)$$

Let us now take an approximate value of W for a large $N!$. Using Stirling's approximation $\ln N! \cong N \ln N - N$, we have:

$$\begin{aligned} \ln W &\cong N \ln N - N - \left[\sum_{i=1}^l N_i \ln (N_i) - \sum_{i=1}^l (N_i) \right] \\ &= - \sum_{i=1}^l N_i \ln N_i = -N \sum_{i=1}^l p_i \ln p_i \end{aligned}$$

$S_B = -K \sum_{i=1}^n p_i \ln p_i$ where $p_i = \frac{N_i}{N}$ is the probability of the occurrence of i th microstates. Boltzmann was the first to emphasize the probabilistic meaning of entropy and the probabilistic nature of thermodynamics.

1.3. Information Theory and Shannon's Entropy

Unlike the first two entropy approaches (thermodynamic entropy by Clausius and Boltzmann's entropy), the third major form of entropy did not fall within the field of physics, but was developed instead in a new field known as information theory [15–17] (also known as communication theory). A fundamental step in using entropy in new contexts unrelated to thermodynamics was provided by Shannon [18], who came to conclude that entropy could be used to measure types of disorder other than that of thermodynamic microstates. Shannon was interested in information theory [19,20], particularly in the ways in which information can be conveyed via a message. This led him to examine probability distributions in a very general sense and he worked to find a way of measuring the level of uncertainty in different distributions.

For example, suppose the probability distribution for the outcome of a coin toss experiment is $P(H) = 0.999$ and $P(T) = 0.001$. One is likely to notice that there is much more "certainty" than "uncertainty" about the outcome of this experiment and, consequently, the probability distribution. If, on the other hand, the probability distribution governing that same experiment were $P(H) = 0.5$ and $P(T) = 0.5$, then there is much less "certainty" and much more "uncertainty" when compared to the previous distribution. However, how can these uncertainties can be quantified? Is there some algebraic function which measures the amount of uncertainty in any probabilistic distribution in terms of the individual probabilities? From these types of simple examples and others, Shannon was able to devise a set of criteria which any measure of uncertainty may satisfy. He then tried

to find an algebraic form which would satisfy his criteria and discovered that there was only one formula which fit. Let the probabilities of n possible outcomes E_1, E_2, \dots, E_n of an experiment be p_1, p_2, \dots, p_n , giving rise to the probability distribution $P = (p_1, p_2, \dots, p_n)$; $\sum_{i=1}^n p_i = 1$; $p_i \geq 0$. There is an uncertainty as to the outcome when the experiment is performed. Shannon suggested the measure $-\sum_{i=1}^n p_i \ln p_i$, which is identical to the previous entropy relation if the constant of probability is taken as the Boltzmann constant K . Thus, Shannon showed that entropy, which measures the amount of disorder in a thermodynamic system, also measures the amount of uncertainty in any probability distribution. Let us now give the formal definition of Shannon's entropy as follows: Consider a random experiment $P = (p_1, p_2, \dots, p_n)$ whose possible outcomes have probabilities $p_i, i = 1, 2, \dots, n$ that are known. Can we guess in advance which outcome we shall obtain? Can we measure the amount of uncertainty? We shall denote such an uncertainty measure by $H(P) = H_n(p_1, p_2, \dots, p_n)$. The most common as well as the most useful measure of uncertainty is Shannon's informational entropy (which should satisfy some basic requirements), which is defined as follows:

Definition I: Let (p_1, p_2, \dots, p_n) be the probability of the occurrence of the events E_1, E_2, \dots, E_n associated with a random experiment. The Shannon's entropy probability distribution (p_1, p_2, \dots, p_n) of the random experiment system P is defined by $H(P) = H_n(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n p_i \ln p_i$ where, $0 \ln 0 \cong 0$. The above definition is generalized straightforwardly as the definition of entropy of a random variable.

Definition II: Let $X \in R$ be a discrete random variable which takes the value $x_i (i = 1, 2, 3, \dots, n)$ with the probability $p_i, i = 1, 2, \dots, n$; then the entropy $H(X)$ of X is defined by the expression $H(X) = -\sum_{i=1}^n p_i \ln p_i$. Examination of H or $H_n(P)$ reveals why Shannon's measure is the most satisfactory measure of entropy because of the following:

- (i) $H_n(P)$ is a continuous function of p_1, p_2, \dots, p_n .
- (ii) $H_n(P)$ is a symmetric function of its arguments.
- (iii) $H_{n+1}(p_1, p_2, \dots, p_n, 0) = H_n(p_1, p_2, \dots, p_n)$, i.e., it should not change if there is an impossible outcome to the probability.
- (iv) Its minimum is 0 when there is no uncertainty about the outcome. Thus, it should vanish when one of the outcomes is certain to happen so that

$$H_n(p_1, p_2, \dots, p_n) = 0; p_i = 1, p_j = 0, i \neq j, i = 1, 2, \dots, n$$

- (v) It is the maximum when there is maximum uncertainty, which arises when the outcomes are equally likely so that $H_n(P)$ is the maximum when $p_1 = p_2 = \dots = p_n = \frac{1}{n}$.
- (vi) The maximum value of $H_n(P)$ increases with n .
- (vii) For two independent probability distributions $P(p_1, p_2, \dots, p_n)$ and $Q(q_1, q_2, \dots, q_m)$, the uncertainty of the joint scheme $P \cup Q$ should be the sum of their uncertainties: $H_{n+m}(P \cup Q) = H_n(P) + H_m(Q)$

Shannon's entropy has various applications in the field of portfolio analysis, the measurement of economic analysis, transportation, and urban and regional planning as well as in the fields of statistics, thermodynamics, queuing theory, parametric estimation, etc. It has been used in the field non-commensurable and conflicting criteria [21] and in the nonlinear complexity of random sequences [22] as well.

2. Discussion

2.1. Jaynes' Maximum Entropy (MaxEnt) Principle

Let the random variable of an experiment be X , and assume the probability mass associated with the value x_i is p_i , i.e., $P_X(i) \equiv P(X) = p_i, i = 1, 2, \dots, n$. The set $(X, P) = \{(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n)\}$, $\sum_{i=1}^n p_i = 1$ is called the source ensemble as described by Karmeshu [23]. In general, we may find expected values of the functions $g_1(X), g_2(X), \dots, g_z(X)$ to get $\sum_{i=1}^n p_i g_{ri} = a_r, r = 1, 2, \dots, z, p_i \geq 0$ and with natural constraint $\sum_{i=1}^n p_i = 1$, given a number of constraints. Thus, we have $z + 1$ relations between (p_1, p_2, \dots, p_n) . There may be infinite probability distributions (p_1, p_2, \dots, p_n) satisfying the above equation. If we know only $g_1(x), g_2(x), \dots, g_z(x)$, then we get a family of max entropy distributions. If, in addition, we know the values of a_1, a_2, \dots, a_z , we get a specific member of this family and we call it the max entropy probability distribution. According to a great article by Jaynes [24–26], we choose the probability distribution out of all these which maximizes the measure of entropy as shown by Shannon's equation, $S(P) = -\sum_{i=1}^n p_i \ln p_i$.

Any distribution of the form $p_i = \exp[-\lambda_0 - \lambda_1 g_1(x_i) - \lambda_2 g_2(x_i) - \dots - \lambda_z g_z(x_i)]$, $i = 1, 2, \dots, n$ may be regarded as the maximum entropy distribution where $\lambda_0, \lambda_1, \dots, \lambda_z$ are determined as functions of a_1, a_2, \dots, a_z ; then the maximum entropy S_{\max} is given by $S_{\max} = \lambda_0 + \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_z a_z$.

Kapur [27,28] showed that there is always a concave function of a_1, a_2, \dots, a_m . We also note that all the probabilities given by p_i are always positive. We naturally want to know whether there is another measure of entropy other than Shannon's entropy which, when maximized, subject to, $\sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g_{ri} = a_r, r = 1, 2, \dots, z, p_i \geq 0$, gives positive probabilities and for which S_{\max} is possibly a concave function [29] of parameters. Kapur [30] studied that Burg's [31] measure of entropy, which has been very successfully used in spectral analysis, does always give positive probabilities. The maximum entropy principal of Jaynes has been used frequently to derive the distribution of statistical mechanics by maximizing the entropy of the system subject to some given constraints. The Maxwell-Boltzman distribution is obtained when there is only one constraint on a system which prescribes the expected energy per particle of the system by Bose-Einstein (B.E.) distribution, Fermi-Dirac (F.D.) distribution and intermediate statistics (I.S.) distributions; these are obtained by maximizing the entropy subject to two constraints by Kapur and Kesavan, and Kullback [32,33] and also by the present authors [34].

2.2. Formulation of MEPD in Statistical Mechanics Using Shannon's Measure of Entropy

Let p_1, p_2, \dots, p_n be the probabilities of a particle having energy levels $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, respectively, and let the expected value of energy be prescribed as $\bar{\varepsilon}$; then, to get MEPD, we maximize the Shannon's measure of entropy:

$$S(P) = -\sum_{i=1}^n p_i \ln p_i \quad (1)$$

Subject to

$$\sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \varepsilon_i = \bar{\varepsilon} \quad (2)$$

Let the Lagrangian be

$$L = -\sum_{i=1}^n p_i \ln p_i - (\lambda - 1) \left(\sum_{i=1}^n p_i - 1 \right) - \mu \left(\sum_{i=1}^n p_i \varepsilon_i - \bar{\varepsilon} \right) \quad (3)$$

Differentiating with respect to p_i 's, we get:

$\ln p_i + \lambda + \mu \varepsilon_i = 0 \Rightarrow p_i = \exp(-\lambda - \mu \varepsilon_i)$ where, λ, μ are to be determined by using Equation (2) so that

$$p_i = \frac{\exp(-\mu \varepsilon_i)}{\sum_{i=1}^n \exp(-\mu \varepsilon_i)}, i = 1, 2, \dots, n, \quad (4)$$

Where

$$\frac{\sum_{i=1}^n \varepsilon_i \exp(-\mu \varepsilon_i)}{\sum_{i=1}^n \exp(-\mu \varepsilon_i)} = \bar{\varepsilon} \quad (5)$$

Equation (4) is the well-known Maxwell-Boltzmann distribution from statistical mechanics which is used in many areas [35–37].

2.3. Burg's Entropy Measure and MEPD

When $S(P) = -\sum_{i=1}^n p_i \ln p_i$ was replaced by Burg's measure of entropy $B(P) = \sum_i \ln p_i$, it gave interesting results as shown by Kapur. Burg's measure of entropy is always negative, but this does not matter in entropy maximization, where it has been found that a probability distribution with maximum entropy satisfies the same constraint and it does not matter if all the entropies are negative. So, in Equation (1) when we use

$$B(P) = \sum_i \ln p_i, \text{ with the constraints } \sum_{i=1}^n p_i = 1, \sum_{i=1}^n i p_i = m \quad (6)$$

we get

$$p_i = \frac{1}{\lambda + \mu i}, i = 1, 2, 3, \dots, n \quad (7)$$

where λ & μ are obtained by solving the equations

$$\sum_i \frac{1}{\lambda + \mu i} = 1, \sum_i \frac{i}{\lambda + \mu i} = m \quad (8)$$

Multiplying first and second Equation (8) by λ & μ respectively then adding
We get

$$\sum_i \frac{\lambda + \mu i}{\lambda + \mu i} = \lambda + \mu m \text{ or, } n = \lambda + \mu m \quad (9)$$

so that from Equation (8),

$$\sum_{i=1}^n \frac{1}{n - \mu(m - i)} = 1 \quad (10)$$

Then, $\mu = 0$ is an obvious solution but that will give us $\lambda = n$ and this will satisfy the second equation of (8)

if $m = \frac{n+1}{2}$. Now, Equation (10) is the n th degree polynomial in μ , and one of its roots is zero. Its non-zero solutions will be obtained by solving an equation of $(n-1)$ th degree in μ . Lemma has been proved by Kapur as the following:

Lemma: All the roots of $\sum_{r=1}^n \frac{1}{n - \mu(m - i)} = 1$ are real; in other words, none of the roots can be complex.

Proof: Let $\mu = \alpha + i\beta, i = \sqrt{-1}$ be a pair of complex conjugate roots of Equation (10). Then,

$\sum_{r=1}^n \frac{1}{n - (\alpha + i\beta)(m - i)} = 1$ and $\sum_{r=1}^n \frac{1}{n - (\alpha - i\beta)(m - i)} = 1$; subtracting the second from the first, we get

$\sum_{r=1}^n \frac{2i\beta(m - r)}{\{n - \alpha(m - r)\}^2 + \beta^2} = 0$, which gives $\beta = 0$, or $\sum_{r=1}^n \frac{(m - r)}{\{n - \alpha(m - r)\}^2 + \beta^2} = 0$. The second possibility can easily be ruled out. To find the actual location of the n real roots, let us assume

$f(\mu) \equiv \sum_r \frac{1}{n - \mu \frac{(m - r)}{n}} - 1$ and this function is discontinuous at the following points:
 $-\frac{n}{n - m}, -\frac{n}{n - (m - 1)}, \dots, -\frac{n}{1 - l}, \frac{n}{l}, \dots, \frac{n}{m - 2}, \frac{n}{m - 1}$, where $m = k + l, k = [m]$ and l is a "+" fraction. More precisely, $f(\mu) \rightarrow +\infty$ when $\mu \rightarrow$ points from one side and $f(\mu) \rightarrow -\infty$ when $\mu \rightarrow$ points from other side. Again,

$$\begin{aligned} f(\mu) &\equiv \sum_r \frac{1}{n - \mu \frac{(m - r)}{n}} - 1 = \frac{1}{n} \left[\sum_r \left\{ 1 - \mu \left(\frac{m - r}{n} \right) \right\}^{-1} \right] - 1 \\ &= \frac{1}{n} \left[\sum_r \left\{ 1 + \mu \left(\frac{m - r}{n} \right) \right\} + O(\mu^2) \right] - 1 = \frac{1}{n} \left[n + \frac{\mu}{n} \left\{ mn - \frac{n(n + 1)}{2} \right\} \right] - 1 + O(\mu^2) \\ &= \frac{\mu}{n} \left\{ m - \frac{(n + 1)}{2} \right\} + O(\mu^2) \end{aligned}$$

2.4. d Burg's Modified Entropy (MBE) Measure and MEPD

2.4.1. Monotonic Character of MBE

We propose use of Burg's modified entropy instead of Burg's entropy. Maximizing the Burg's modified measure of entropy:

$$B_{\text{mod}}(P) = \sum_{i=1}^n \ln(1 + ap_i) - \ln(1 + a), a > 0 \quad (11)$$

$$\frac{d}{da}(B_{\text{mod}}(P)) = \sum_{i=1}^n \frac{p_i}{1 + ap_i} - \frac{1}{1 + a} = \sum_{i=1}^n p_i \left[\frac{1}{1 + ap_i} - \frac{1}{1 + a} \right] = \sum_{i=1}^n \frac{ap_i(1 - p_i)}{(1 + a)(1 + ap_i)} > 0 \quad (12)$$

Therefore, $B_{\text{mod}}(P)$ is the monotonic increasing function of a . For the probability distribution

$$(p, 1 - p)$$

$$\frac{d}{da}(B_{\text{mod}}(P)) = \sum_{i=1}^n \frac{p_i}{1 + ap_i} - \frac{1}{1 + a} = \sum_{i=1}^n p_i \left[\frac{1}{1 + ap_i} - \frac{1}{1 + a} \right] = \sum_{i=1}^n \frac{ap_i(1 - p_i)}{(1 + a)(1 + ap_i)} > 0 \quad (13)$$

it is showed (see: Table 1.) that $\text{Max}_P(B_{\text{mod}}(P)) = \ln \left(1 + \frac{a^2}{4(1 + a)} \right) = \ln \left(\frac{(a + 2)^2}{4(1 + a)} \right)$.

The measure of entropy $\text{Max}_P(B_{\text{mod}}(P))$ is the Burg's modified entropy. This is a better measure than the standard Burg's measure since it is always positive and there is no computational problem when p_i is very small. In the above case, the maximum value increases with the number of possible outcomes n .

Table 1. The a values & maximum values of Burg's Modified Entropy.

a	$Max_p (B_{\text{mod}}(P))$
0.5	0.04082
1.0	0.11778
1.5	0.20294
2.0	0.28768
5.0	0.71376
10.0	1.1856
20.0	1.7512
30.0	2.1111
40.0	2.3754
50.0	2.5843
∞	∞

2.4.2. MBE and Its Relation with Burg's Entropy

$$\begin{aligned}
 B_{\text{mod}}(P) &= \sum_{i=1}^n \ln(1 + ap_i) - \ln(1 + a) = \sum_{i=1}^n \left[\ln a + \ln \left(p_i + \frac{1}{a} \right) \right] - \ln a - \ln \left(1 + \frac{1}{a} \right) \\
 &= (n-1) \ln a + \sum_{i=1}^n \ln \left(p_i + \frac{1}{a} \right) - \ln \left(1 + \frac{1}{a} \right)
 \end{aligned}$$

So, when $a \rightarrow \infty$, maximizing $(B_{\text{mod}}(P))$ and $(B(P))$ will give the same result in both cases; again, if $(B_{\text{mod}}(P))$ is maximized under the constraints

$$\sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i g_{ri} = a_r, r = 1, 2, \dots, m, p_i \geq 0 \quad (14)$$

we get

$$\frac{1}{1 + ap_i} = [\lambda_0 + \lambda_1 g_1(x_i) + \lambda_2 g_2(x_i) + \dots + \lambda_m g_m(x_i)], i = 1, 2, \dots, n. \text{ Letting, we have}$$

$$\frac{1}{p_i} = \lambda'_0 + \lambda'_1 g_1(x_i) + \lambda'_2 g_2(x_i) + \dots + \lambda'_m g_m(x_i) \quad (15)$$

The λ'_i 's are determined by using Constraints (14) and (15) and this gives the MEPD when Burg's entropy is maximized as subject to Equation (14).

Therefore, when $a \rightarrow \infty$, the MEPD of BME \rightarrow Burg's MEPD; in fact:

$$\begin{aligned}
 (B_{\text{mod}}(P)) - (B(P)) &= (n-1) \ln a + \sum_{i=1}^n \ln \frac{\left(p_i + \frac{1}{a} \right)}{p_i} - \ln \left(1 + \frac{1}{a} \right) \rightarrow \infty \text{ as} \\
 &a \rightarrow \infty.
 \end{aligned}$$

2.4.3. MBE and Its Concavity of S_{max} under Prescribed Mean

$$\text{Maximize } B_{\text{mod}}(P) = \sum_{i=1}^n \ln(1 + ap_i) - \ln(1 + a) \quad (16)$$

$$\text{subject to } \sum_{i=1}^n p_i = 1, \sum_{i=1}^n ip_i = m.$$

We obtain this using Lagranges multiplier mechanics:

$$\left(p_i + \frac{1}{a}\right) = \frac{1}{\lambda + \mu i} \text{ Where, } \sum_i \frac{1}{\lambda + \mu i} = \sum_i \left(p_i + \frac{1}{a}\right) = 1 + \frac{n}{a} \quad (17)$$

From the above equation:

$$\sum_i \frac{i \cdot 1}{\lambda + \mu i} = \sum_i \left(p_i + \frac{1}{a}\right) = \sum_{i=1}^n \left(ip_i + \frac{i}{a}\right) \quad (18)$$

$$\sum_i \frac{i}{\lambda + \mu i} = m + \frac{n(n+1)}{2a}, \text{ i.e., } \frac{n(n+1)}{2a} + m = \sum_i \frac{i}{\lambda + \mu i} \text{ and}$$

$$\frac{\mu n(n+1)}{2a} + \mu m = \sum_i \frac{\mu i}{\lambda + \mu i}, \text{ therefore } \sum_i \frac{\lambda}{\lambda + \mu i} = \sum_i \left(p_i + \frac{1}{a}\right) = \lambda \left(1 + \frac{n}{a}\right) \quad (19)$$

So

$$\sum_i \frac{\lambda}{\lambda + \mu i} + \sum_i \frac{\mu i}{\lambda + \mu i} = \lambda \left(1 + \frac{n}{a}\right) + \frac{\mu n(n+1)}{2a} + \mu m = \sum_i \frac{\lambda + \mu i}{\lambda + \mu i} = \sum_i 1 = n \quad (20)$$

$$\lambda \left(1 + \frac{n}{a}\right) + \frac{\mu n(n+1)}{2a} + \mu m = n \Rightarrow \lambda = \frac{n - \mu \left(m + \frac{n(n+1)}{2a}\right)}{1 + \frac{n}{a}} \quad (21)$$

Therefore,

$$\lambda + \mu i = \frac{n + \mu \left[\left(1 + \frac{n}{a}\right)i - m - \frac{n(n+1)}{2a}\right]}{1 + \frac{n}{a}} \quad (22)$$

and

$$p_i = \frac{a + n}{an + \mu \left[(n + a)i - am - \frac{n(n+1)}{2}\right]} - \frac{1}{a}, i = 1, 2, \dots, n, \quad (23)$$

where μ is determined as a function of m and that is

$$1 + \frac{n}{a} = \sum_i^n \frac{a + n}{an + \mu \left[(n + a)i - am - \frac{n(n+1)}{2}\right]} \&$$

$$\frac{1}{a} = \sum_i^n \frac{1}{an + \mu \left[(n + a)i - am - \frac{n(n+1)}{2}\right]},$$

Therefore,

$$\begin{aligned} S_{\max} &= \sum_i \ln(1 + ap_i) - \ln(1 + a) \\ &= \sum_i \ln \frac{a}{\lambda + \mu i} - \ln(1 + a) = \ln \frac{a^n}{1 + a} - \sum_i \ln(\lambda + \mu i) \\ &= \ln \frac{a^n}{1 + a} - \sum_i \ln \left[n + \mu \left(\left(1 + \frac{n}{a}\right)i - m - \frac{n(n+1)}{2} \right) \right] + n \ln \left(1 + \frac{n}{a}\right) \end{aligned}$$

$$\begin{aligned}
\frac{dS_{\max}}{dm} &= -\sum_{i=1}^n \frac{\left[\frac{d\mu}{dm} \left\{ \left(1 + \frac{n}{a}\right) i - m - \frac{n(n+1)}{2a} \right\} - \mu \right]}{n + \mu \left[\left(1 + \frac{n}{a}\right) i - m - \frac{n(n+1)}{2a} \right]} \\
&= -\sum \frac{\left(p_i + \frac{1}{a} \right)}{a + n} a \left[\frac{d\mu}{dm} \left[\left(1 + \frac{n}{a}\right) i - m - \frac{n(n+1)}{2a} \right] - \mu \right] \\
&= \frac{d\mu}{dm} \left[m + \frac{n(n+1)}{2} - m - \frac{n(n+1)}{2} \right] + \frac{\mu a}{a + n} \left(1 + \frac{n}{a} \right)
\end{aligned}$$

So that

$$\frac{dS_{\max}}{dm} = \frac{d\mu}{dm} \& \frac{d^2 S_{\max}}{dm^2} = \frac{d\mu}{dm} = \frac{\mu \sum_i \left(p_i + \frac{1}{a} \right)^2}{\sum_i \left(p_i + \frac{1}{a} \right)^2 \left[\left(1 + \frac{n}{a}\right) i - m - \frac{n(n+1)}{2a} \right]} \quad (24)$$

S_{\max} will be a concave function of m if $\frac{d\mu}{dm} < 0$, that is, if either $\mu > 0, < 0$ when the denominator $< 0, > 0$, respectively.

In the above case, when $a = 1$ we get S_{\max} and the derivative of S_{\max} as follows:

$$\begin{aligned}
\frac{dS_{\max}}{dm} &= -\sum_i \frac{1}{\lambda + \mu i} \left[\frac{d}{dm} (\lambda + \mu i) \right] = \sum_i \left[\frac{n}{2} \frac{d\mu}{dm} + \frac{1}{n+1} \left(m \frac{d\mu}{dm} + \mu \right) - \frac{d\mu}{dm} i \right] \quad (25) \\
&= \frac{d\mu}{dm} \sum_i (1 + p_i) \left(\frac{n}{2} + \frac{m}{n+1} - i \right) + \frac{\mu}{n+1} \sum_i (1 + p_i) \\
&= \frac{d\mu}{dm} \sum_i \left(\frac{n}{2} + \frac{m}{n+1} - i \right) + \frac{d\mu}{dm} \left(\frac{n}{2} + \frac{m}{n+1} \right) - \frac{d\mu}{dm} \sum_i i p_i + \frac{\mu n}{n+1} + \frac{\mu}{n+1} \\
&= \frac{d\mu}{dm} \sum_i \left(\frac{n}{2} \cdot n + \frac{m}{n+1} \cdot n - \frac{n(n+1)}{2} \right) + \frac{d\mu}{dm} \left(\frac{n}{2} + \frac{m}{n+1} \right) - \frac{d\mu}{dm} m + \frac{\mu n}{n+1} + \frac{\mu}{n+1} \\
&= \frac{d\mu}{dm} \left(\frac{mn}{n+1} - \frac{m}{n+1} - m \right) + \frac{\mu n}{n+1} + \frac{\mu}{n+1} = \mu \left(\frac{n}{n+1} + \frac{1}{n+1} \right) = \mu \quad (26)
\end{aligned}$$

So that

$$\frac{d^2 S_{\max}}{dm^2} = \frac{d\mu}{dm} \quad (27)$$

Additionally, we have

$$0 = -\sum_i (1 + p_i)^2 \left[\frac{n}{2} \frac{d\mu}{dm} + \frac{1}{n+1} \left(m \frac{d\mu}{dm} + \mu \right) - \frac{d\mu}{dm} i \right] \quad (28)$$

So that

$$\frac{d\mu}{dm} \sum_i (1 + p_i)^2 \left(\frac{n}{2} + \frac{m}{n+1} - i \right) + \frac{\mu}{n+1} \sum_i (i + p_i)^2 = 0 \quad (29)$$

Therefore,

$$\frac{d\mu}{dm} = \frac{\frac{\mu}{n+1} \sum_i (1 + p_i)^2}{\sum_i (1 + p_i)^2 \left(i - \frac{n}{2} - \frac{m}{n+1} \right)} \quad (30)$$

So,

$$\frac{d^2 S_{\max}}{dm^2} < 0 \Rightarrow \frac{d\mu}{dm} < 0 \quad (31)$$

Therefore, from Equations (27) and (30),
 S_{\max} will be a concave function of m if either

$$(i) \mu > 0, \sum_i (1 + p_i)^2 \left(i - \frac{n}{2} - \frac{m}{n+1} \right) < 0 \quad (32)$$

Or

$$(ii) \mu < 0, \sum_i (1 + p_i)^2 \left(i - \frac{n}{2} - \frac{m}{n+1} \right) > 0 \quad (33)$$

Additionally, when $\mu = 0$, $\frac{d^2 S_{\max}}{dm^2} = 0$, all the probabilities are equal, and from Equation (24) we have:

We have,

$$m + \frac{n(n+1)}{2} = \frac{1}{\lambda} \frac{n(n+1)}{2} \text{ or, } m = \frac{n(n+1)}{2} \left[\frac{1}{\lambda} - 1 \right] \Rightarrow m = \frac{n+1}{2}. \quad (34)$$

Now, if we proceed algebraically as done Section 2.3, we get:

$$1 + p_i = \frac{1}{\lambda + \mu i} = \frac{1}{\frac{1}{n+1} (n - \mu m) - \mu \left(\frac{n}{2} - i \right)}, i = 1, 2, \dots, n,$$

$\sum_{i=1}^n \frac{1}{\frac{1}{n+1} (n - \mu m) - \mu \left(\frac{n}{2} - i \right)} = \sum_{i=1}^n (1 + p_i) = n + 1$. The obvious solution of the above problem is $\mu = 0$ which will give $\lambda = \frac{1}{p_i} = n$, i.e., uniform distribution, and thus we get, $m = \frac{n+1}{2}$.

3. An Illustrative Example in Statistical Mechanics

3.1. Example

Let p_1, p_2, \dots, p_{10} be the probabilities of a particle having energy levels $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{10}$; respectively, and let the expected value of energy be prescribed as m ; then, we get the maximum entropy probability distribution (MEPD) with MBE as follows:

$$\text{Maximize } \sum_{i=1}^{10} \ln(1 + p_i) - \ln 2 \text{ such that, } \sum_{i=1}^{10} p_i = 1 \text{ and } \sum_{i=1}^{10} i p_i = m \text{ for } m = 1 (0.25) 10$$

Solution: Maximizing the measure of entropy subject to the given constraints, we have
 $(1 + p_i) = \frac{1}{\lambda + \mu i}$.

When $\sum_i \frac{1}{\lambda + \mu i} = 10 + 1 = 11$, $\frac{10(10+1)}{2} + m = \sum_i \frac{i}{\lambda + \mu i}$ i.e., $\sum_i \frac{i}{\lambda + \mu i} = 55 + m$, we get the probability distribution in the form of a table and also get the values of S_{\max} as described by Kapur and Kesavan.

There may be two cases,

Case (i) when $m < \frac{n+1}{2} = 5.5$.

In this case, when m lies between 1 and 5.5, $\mu > 0$ implies $\frac{dS_{\max}}{dm} > 0$ and then S_{\max} is increasing.

Case (ii) when $m > \frac{n+1}{2} = 5.5$.

In this case, when m lies between 5.5 and 10, $\mu < 0$ implies $\frac{dS_{\max}}{dm} < 0$ and then S_{\max} is decreasing.

When $m = \frac{n+1}{2} = 5.5$, $\mu = 0$, $\frac{dS_{\max}}{dm} = 0$.

it can be shown that S_{\max} will be concave if we prescribe $E(g(x))$ instead of $E(x)$, where $g(x)$ is a monotonic increasing function of x ; then we will apply the necessary changes. Again, since the concavity of S_{\max} has already been proven, this will enable us to handle the inequality constraint of the type $m_1 \leq \sum_{i=1}^n ip_i \leq m_2$.

3.2. Simulated Results

Following Table (Table 2) Found Using LINGO Software 2011 where Different Max-Entropy Values Are Given for Different m Values:

Table 2. Comparative Maximum Entropy Values.

m	Maximum Entropy Value of $B_{\text{mod}}(P)$	Maximum Entropy Value of $B(P)$	Maximum Entropy Value of $S(P)$
1.00	0	−158.225	0
1.25	0.0896122	−45.13505	0.6255029
1.50	0.1345336	−38.98478	0.9547543
1.75	0.1633092	−35.43113	1.194875
2.00	0.1836655	−32.94620	1.385892
2.25	0.1989097	−31.05219	1.542705
2.50	0.2107459	−29.53735	1.675885
2.75	0.2201939	−28.28982	1.790029
3.00	0.2279397	−27.24394	1.888477
3.25	0.2344146	−26.35851	1.973503
3.50	0.2399041	−25.60649	2.046725
3.75	0.24461	−24.96958	2.109324
4.00	0.248682	−24.43512	2.162186
4.25	0.252127	−23.99419	2.205980
4.50	0.2549453	−23.64047	2.241209
4.75	0.257137	−23.36942	2.268253
5.00	0.2587024	−23.17789	2.287386
5.25	0.2596416	−23.06376	2.298794
5.50	0.2599546	−23.02585	2.302585
5.75	0.2596416	−23.06376	2.298794
6.00	0.2587024	−23.17789	2.287386
6.25	0.257137	−23.36942	2.268253
6.50	0.2549453	−23.64047	2.241209
6.75	0.252127	−23.99419	2.205980
7.00	0.248682	−24.43512	2.162186
7.25	0.24461	−24.96958	2.109324
7.50	0.2399041	−25.60649	2.046725
7.75	0.2344146	−26.35851	1.973503
8.00	0.2279397	−27.24394	1.888477
8.25	0.2201939	−28.28982	1.790029
8.50	0.2107459	−29.53735	1.675885
8.75	0.1989097	−31.05219	1.542705
9.00	0.1836655	−32.94620	1.385892
9.25	0.1633092	−35.43113	1.194875
9.50	0.1345336	−38.98478	0.9547543
9.75	0.0896122	−45.13505	0.6255029
10.0	0	−141.25285	0

Graphs obtained from the above table are given on the next page (Figures 1–3):

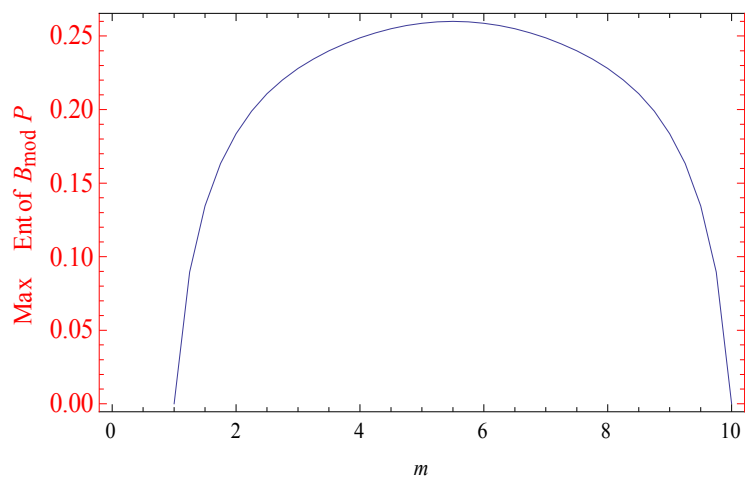


Figure 1. Maximum Burg's modified entropy.

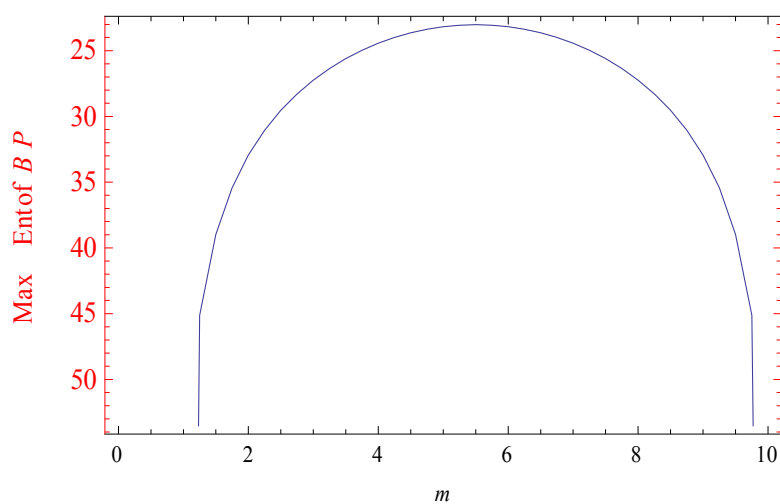


Figure 2. Maximum Burg's entropy.

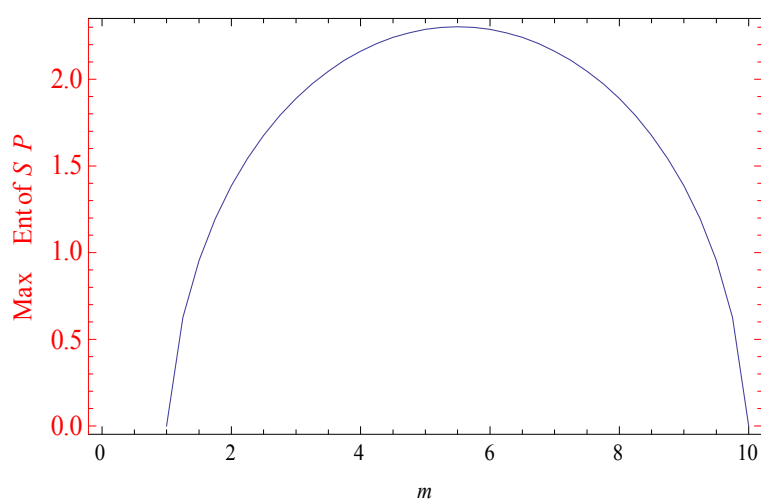


Figure 3. Maximum Shannon's entropy.

If p_i is plotted against i , we get rectangular hyperbolic types of curves (Table 3 and Figure 4).

Table 3. p_i and i graph.

i	p_i
1	1.000000
2	0.2899575
3	0.1747893
4	0.1253783
5	0.1027619
6	0.1027620
7	0.1253782
8	0.1747891
9	0.2899574
10	1.000000

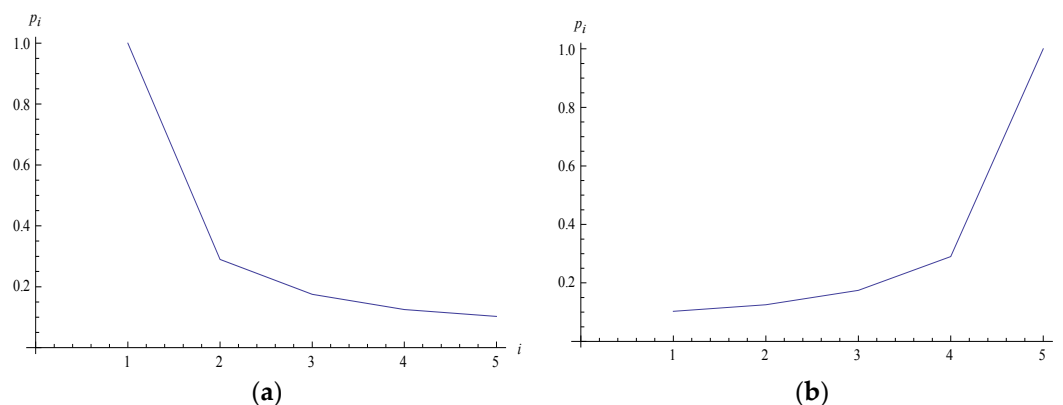


Figure 4. Rectangular hyperbolic types of graphs. (a) when $m < \frac{n+1}{2}, \mu > 0$; (b) when $m > \frac{n+1}{2}, \mu < 0$.

4. Conclusions

In the present paper we have presented different MEPDs and respective entropy measures with their properties. It has been found that MBE is a better measure than Burg's entropy when the maximized subject to the mean is prescribed, and it also has been shown that unlike Burg's entropy, the maximum value of MBE increases with n . The main problem here will consist of solving $z + 1$ simultaneous transcendental equations for the Lagranges multipliers. An application in statistical mechanics with simulated data has been studied with the help of Lingo11 software and corresponding graphs are provided. Now, one question arises: Will this result continue to hold for other moment constraints also? When we take generalized moment expectation of $g(x)$ instead of expectation of x , then $g(x)$ must be a monotonic increasing function of x , and if p_i becomes negative for some values of the moments, then we have to set those probabilities to zero and reformulate the problem for the remaining probabilities over the remaining range and solve it.

Acknowledgments: Authors would like to express their sincere gratitude and appreciation to the reviewers for their valuable comments and good suggestions to make this paper enriched and impactful, and a special thank you to the Assistant Editor.

Author Contributions: A. Ray and S. K. Majumder conceived and designed the experiments; A. Ray performed the experiments, analyzed the data and both A. Ray and S. K. Majumder contributed in analysis tools; A. Ray wrote the paper.

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

Nomenclature

S = Entropy

$p_i = \frac{N_i}{N}$, probability of i th event

$H(P) = H_n(p_1, p_2, \dots, p_n)$ = Information entropy

$S(P)$ or $H_n(P) = -\sum_{i=1}^n p_i \ln p_i$ = Shannon's entropy

$B(P) = \sum_i \ln p_i$ = Burg's entropy

$B_{\text{mod}}(P)$ = Burg's modified entropy

n = Number of energy levels/number of possible outcome

$\mu = \frac{1}{KT}$

K = Boltzmann constant

T = Absolute temperature

N = Identical particle of ideal gas

Δs = Increase of entropy

dS = Change in entropy

$[m]$ = Greatest integer value of m

Where m is the mean value:

$E(x)$ = Expectation

dV = Change in volume

$P \cup Q$ = Union of two sets

$m = 1(0.25)10$, m ranges from 1 to 10 with step length of 0.25

S_{max} = Maximum entropy

$\text{Max}_P(B_{\text{mod}}(P))$ = Maximum value under the given probability distribution

MBE=Modified Burg's Entropy

Greek Symbols

Ω = The maximum number of microscopic ways in the macroscopic state

ξ = Position of the molecule

ρ = Momentum of the molecule

μ, λ =Lagrangian constant

$\varepsilon_1, \varepsilon_2, \dots$ =Different energy levels

$\bar{\varepsilon}$ =Mean energy

Subscripts

B =Boltzmann

max =Maximum

mod =modified

Superscript

λ' =New constant different from λ

References

1. Andreas, G.; Keller, G.; Warnecke, G. *Entropy*; Princeton University Press: Princeton, NJ, USA, 2003.
2. Phil, A. *Thermodynamics and Statistical Mechanics: Equilibrium by Entropy Maximisation*; Academic Press: Cambridge, MA, USA, 2002.
3. Rudolf, C. Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie. *Annalen der Physik* **1865**. (In German). [[CrossRef](#)]
4. Wu, J. Three Factors Causing the Thermal Efficiency of a Heat Engine to Be Less than Unity and Their Relevance to Daily Life. *Eur. J. Phys.* **2015**, *36*, 015008. [[CrossRef](#)]
5. Rashidi, M.M.; Shamekhi, L. Entropy Generation Analysis of the Revised Cheng-Minkowycz Problem for Natural Convective Boundary Layer Flow of Nanofluid in a Porous Medium. *J. Thermal Sci.* **2015**, *19*. [[CrossRef](#)]
6. Rashidi, M.M.; Mohammadi, F. Entropy Generation Analysis for Stagnation Point Flow in a Porous Medium over a Permeable Stretching Surface. *J. Appl. Fluid Mech.* **2014**, *8*, 753–763.
7. Rashidi, M.M.; Mahmud, S. Analysis of Entropy Generation in an MHD Flow over a Rotating Porous Disk with Variable Physical Properties. *Int. J. Energy* **2014**. [[CrossRef](#)]
8. Abolbashari, M.H.; Freidoonimehr, N. Analytical Modeling of Entropy Generation for Casson Nano-Fluid Flow Induced by a Stretching Surface. *Adv. Powder Technol.* **2015**, 231. [[CrossRef](#)]
9. Baag, S.S.R.; Dash, M.G.C.; Acharya, M.R. Entropy Generation Analysis for Viscoelastic MHD Flow over a Stretching Sheet Embedded in a Porous Medium. *Ain Shams Eng. J.* **2016**, *23*. [[CrossRef](#)]
10. Shi, Z.; Tao, D. Entropy Generation and Optimization of Laminar Convective Heat Transfer and Fluid Flow in a Microchannel with Staggered Arrays of Pin Fin Structure with Tip Clearance. *Energy Convers. Manag.* **2015**, *94*, 493–504. [[CrossRef](#)]
11. Hossein, A.M.; Ahmadi, M.A.; Mellit, A.; Pourfayaz, F.; Feidt, M. Thermodynamic Analysis and Multi Objective Optimization of Performance of Solar Dish Stirling Engine by the Centrality of Entropy and Entropy Generation. *Int. J. Electr. Power Energy Syst.* **2016**, *78*, 88–95. [[CrossRef](#)]
12. Giovanni, G. *Statistical Mechanics: A Short Treatise*; Springer Science & Business Media: New York, NY, USA, 2013.
13. Reif, F. *Fundamentals of Statistical and Thermal Physics*; Waveland Press: Long Grove, IL, USA, 2009.
14. Rudolf, C.; Shimony, A. *Two Essays on Entropy*; University of California Press: Berkeley, CA, USA, 1977.
15. Shu-Cherng, F.; Rajasekera, J.R.; Tsao, H.S.J. *Entropy Optimization and Mathematical Programming*; Springer Science & Business Media: New York, NY, USA, 2012.
16. Robert, M.G. *Entropy and Information Theory*; Springer Science & Business Media: New York, NY, USA, 2013.
17. Silviu, G. *Information Theory with New Applications*; MacGraw-Hill Books Company: New York, NY, USA, 1977.
18. Shannon, C.E. A Mathematical Theory of Communication. *SigmobileMob. Comput. Commun. Rev.* **2001**, *5*, 3–55. [[CrossRef](#)]
19. Thomas, M.C.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
20. Christodoulos, A.F.; Pardalos, P.M. *Encyclopedia of Optimization*; Springer Science & Business Media: New York, NY, USA, 2008.
21. Arash, M.; Azarnivand, A. Application of Integrated Shannon's Entropy and VIKOR Techniques in Prioritization of Flood Risk in the Shemshak Watershed, Iran. *Water Resour. Manag.* **2015**, *30*, 409–425. [[CrossRef](#)]
22. Liu, L.; Miao, S.; Liu, B. On Nonlinear Complexity and Shannon's Entropy of Finite Length Random Sequences. *Entropy* **2015**, *17*, 1936–1945. [[CrossRef](#)]
23. Karmeshu. *Entropy Measures, Maximum Entropy Principle and Emerging Applications*; Springer: Berlin, Germany; Heidelberg, Germany, 2012.
24. Jaynes, E.T. Information Theory and Statistical Mechanics. II. *Phys. Rev.* **1957**, *108*, 171–190. [[CrossRef](#)]
25. Jaynes, E.T. On the Rationale of Maximum-Entropy Methods. *Proc. IEEE* **1982**, *70*, 939–952. [[CrossRef](#)]
26. Jaynes, E.T. Prior Probabilities. *IEEE Trans. Syst. Sci. Cybernet.* **1968**, *4*, 227–241. [[CrossRef](#)]
27. Kapur, J.N. *Maximum-Entropy Models in Science and Engineering*; John Wiley & Sons: New York, NY, USA, 1989.
28. Kapur, J.N. Maximum-Entropy Probability Distribution for a Continuous Random Variate over a Finite Interval. *J. Math. Phys. Sci.* **1982**, *16*, 693–714.

29. Ray, A.; Majumder, S.K. Concavity of maximum entropy through modified Burg's entropy subject to its prescribed mean. *Int. J. Math. Oper. Res.* **2016**, *8*, to appear.
30. Kapur, J.N. *Measures of Information and Their Applications*; Wiley: New York, NY, USA, 1994.
31. Burg, J. The Relationship between Maximum Entropy Spectra and Maximum Likelihood Spectra. *Geophysics* **1972**, *37*, 375–376. [[CrossRef](#)]
32. Narain, K.J.; Kesavan, H.K. *Entropy Optimization Principles with Applications*; Academic Press: Cambridge, MA, USA, 1992.
33. Solomon, K. *Information Theory and Statistics*; Courier Corporation: North Chelmsford, MA, USA, 2012.
34. Amritansu, R.; Majumder, S.K. Derivation of some new distributions in statistical mechanics using maximum entropy approach. *Yugoslav J. Oper. Res.* **2013**, *24*, 145–155.
35. Ulrych, T.J.; Bishop, T.N. Maximum Entropy Spectral Analysis and Autoregressive Decomposition. *Rev. Geophys.* **1975**, *13*, 183–200. [[CrossRef](#)]
36. Michele, P.; Ferrante, A. On the Geometry of Maximum Entropy Problems. **2011**, arXiv:1112.5529.
37. Ke, J.-C.; Lin, C.-H. Maximum Entropy Approach to Machine Repair Problem. *Int. J. Serv. Oper. Inform.* **2010**, *5*, 197–208. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).