Maximum Entropy Principle for the Microcanonical Ensemble

Michele Campisi[1, *] and Donald H. Kobe[1, †]

^{1}Department of Physics,University of North Texas Denton, TX 76203-1427, U.S.A.

(Dated: February 13, 2013)

We derive the microcanonical ensemble from the Maximum Entropy Principle (MEP) using the phase space volume entropy of P. Hertz. Maximizing this entropy with respect to the probability distribution with the constraints of normalization and average energy, we obtain the condition of constant energy. This approach is complementary to the traditional derivation of the microcanonical ensemble from the MEP using Shannon entropy and assuming a priori that the energy is constant which results in equal probabilities.

I. INTRODUCTION

The seminal works of Jaynes [1, 2] presents the information theory approach to statistical physics using the Maximum Entropy Principle (MEP). In the original papers, Jaynes maximized the Shannon information entropy using constraints of normalization and average energy to obtain the canonical ensemble. Later on, Tsallis [3] maximized generalized information entropies, like the Rényi and Tsallis entropies, using constraints of normalization and average energy to obtain deformed exponential distributions that describe the behavior of nonextensive systems.

In this paper we show that there is also a special information entropy associated with the microcanonical ensemble. This microcanonical information entropy is the phase-space volume entropy, originally due to P. Hertz [4] (see also [5]) that satisfies the heat theorem [6, 7, 8]). Using this entropy in the MEP with constraints of normalization *and* average energy, we obtain the condition that the energy distribution is a delta function, *i.e.*, we derive the microcanonical ensemble from the MEP.

In Section 2 we review the traditional application of the MEP to the microcanonical ensemble. The quantum statistical application of the MEP with discrete probabilities using the volume entropy is treated in Section 3. The classical statistical application is given in Section 4, which employs integration and functional differentiation with continuous probability distribution functions. The conclusion is given in Section 5.

II. TRADITIONAL APPROACH TO THE MICROCANONICAL ENSEMBLE

The traditional MEP is reviewed here to contrast it with our approach and to establish the notation. The traditional approach to the quantum microcanonical ensemble starts with the assumption that the system is isolated and has a fixed energy U. Such a macrostate of energy U can be realized in a number W of possible ways each corresponding to a microstate i. Then one looks for the probability p_i that the system is in a certain state i with energy U. In quantum mechanics U is an eigenvalue of the Hamiltonian operator E_β and W is its degeneracy g_β, *i.e.*, $U = E_\beta$ and $W = g_\beta$. Since we are looking for the probability of a state i that is already assumed to belong to the eigenvalue E_β, the traditional MEP does not have to use the energy constraint and is

$$-\sum_{j\in\{j|E_j=E_\beta\}} p_j \log p_j - \lambda \left(\sum_{j\in\{j|E_j=E_\beta\}} p_j - 1 \right) = \text{ maximum}, \tag{1}$$

where the first term is Shannon entropy and the sums are over states restricted to $j \in \{j\,|E_j = E_\beta\}$.

*Electronic address: campisi@unt.edu

†Electronic address: kobe@unt.edu:

The MEP in (1) gives Laplace's Principle of Insufficient Reason

$$p_i = \frac{1}{g_\beta} = \text{constant for} \ \ i \in \{j \,|E_j = E_\beta\} \,, \tag{2}$$

that shows the states j in the given macrostate β with energy E_β are equiprobable. Thus the maximization procedure gives us a *flat* distribution. With some abuse of terminology Eq. (2) is often referred to as the "microcanonical ensemble," but it is defined only for the states j such that $E_j = E_\beta$. Strictly speaking, the microcanonical ensemble is defined on the whole phase space and constrains the system state to lie on a given surface of constant energy. The microcanonical ensemble of energy E_β is really given as [9]

$$p_i = \frac{1}{g_\beta} \delta_{Kr}(E_i, E_\beta) \tag{3}$$

where δ_{Kr} is the Kronecker delta $[\delta_{Kr}(x, y) = 1$ for $x = y$ and 0 for $x \neq y]$. The Kronecker delta does not appear in Eq. (2) because it is assumed *a priori*.

We stress that the traditional approach does not maximize on the whole set of eigenstates of the Hamiltonian but rather on the subset of eigenstates belonging to the eigenvalue E_β. This approach is quite different from Jaynes's derivation of the canonical ensemble, where i runs over *all* the energy eigenstates. In the following section we ask the question: *Is it possible to derive the microcanonical ensemble in (3) from a suitable MEP performed on the whole set of eigenstates, as Jaynes did for the canonical ensemble?*

III. DERIVATION OF THE MICROCANONICAL DISTRIBUTION: QUANTUM CASE

In order to answer to the question posed above, let us proceed by analogy with Jaynes's approach to the canonical ensemble. In order to obtain the canonical distribution,

$$p_i = Z^{-1} e^{-\beta E_i}. \tag{4}$$

where Z is the partition function, and β^{-1} is the absolute temperature, one maximizes the Shannon entropy $-\sum_i p_i \log p_i$ under the energy constraint $U = \sum_i p_i E_i$ and the normalization constraint $\sum_i p_i = 1$, where i runs over all energy eigenstates. When the Shannon entropy is evaluated with the maximal distribution (4) we obtain the correct *thermodynamic entropy*

$$\beta U + \log \sum_n e^{-\beta E_n}. \tag{5}$$

This thermodynamic entropy is correct in the sense that it satisfies the *heat theorem* whenever the averages are calculated over the canonical ensemble [10].

In the microcanonical case the correct *thermodynamic entropy* that satisfies the *heat theorem* is given by the logarithm of the volume of phase space enclosed by the hypersurface of energy $U = E_\beta$ [7, 10]. In the quantum version such entropy is

$$S(U) = \log \Phi(U) \doteq \log \sum_j \theta(U - E_j), \tag{6}$$

where $\theta(x)$ is the step function $[\theta(x) = 1$ for $x \geq 0$, and 0 for $x < 0]$.

Since we are now performing the maximization on the totality of eigenstates, we *must* use the energy constraint as we do with the canonical ensemble. Thus we are maximizing (6) under the normalization *and* average energy conditions,

$$\sum_j p_j = 1, \quad \sum_j p_j E_j = U, \tag{7}$$

Using the constraints in Eq. (7), we can rewrite the entropy in Eq. (6) as

$$S(p) = \log \sum_j \theta \left(\sum_k p_k E_k - E_j \sum_k p_k \right) \tag{8}$$

where the sums on j and k are over all states. The discrete probability distribution $p = \{p_i\}$ for the microcanonical ensemble is obtained when this entropy is an extremum. Differentiating Eq. (8) with respect to p_i and setting the result equal to zero, we obtain

$$\frac{\partial S}{\partial p_i} = \frac{1}{\Phi(U)} \sum_j \delta\left(U - E_j\right)\left(E_i - E_j\right) = 0, \tag{9}$$

for each state i, where $\theta'(x) = \delta(x)$ is the Dirac delta function. We can see by inspection that Eq. (9) is satisfied if $E_j \neq U$. When $E_j = U$ the state i must be such that $E_i = E_j$ [because $x\delta(x) = 0$]. In the latter case we have $E_i = U$. The probability distribution for states i is therefore

$$p_i = A_i \ \delta_{Kr}(E_i, U = E_\beta), \tag{10}$$

where A_i are yet to be determined. The Kronecker delta $\delta_{Kr}(E_i, E_\beta)$ imposes the restriction that the probability of states $i \notin \{i|E_i = E_\beta\}$ are zero.

Since there is nothing to distinguish different states $i \in \{i|E_i = E_\beta\}$, we can invoke Laplace's Principle of Insufficient Reason, obtained from the traditional MEP approach, to choose $A_i = A_\beta$ to be the same for all states belonging to the same eigenenergy E_β. Using the constraint of normalization in Eq. (7), we obtain

$$p_i = \frac{1}{g_\beta}\delta_{Kr}(E_i, E_\beta), \tag{11}$$

which is the microcanonical probability distribution. The only nonzero contributions are from states i with fixed energy $E_i = E_\beta$.

IV. DERIVATION OF THE MICROCANONICAL DISTRIBUTION: CLASSICAL CASE

The derivation of the classical microcanonical distribution proceeds in a way analogous to the quantum derivation. Because we need to use a continuous probability distribution, we must use integration and functional differentiation in the MEP. However, the treatment is sufficiently different to merit some discussion.

Equation (9) for the classical volume entropy of P. Hertz [4] is

$$S(U) = \log \Phi(U), \tag{12}$$

where U is again the energy. In the classical case the function $\Phi(U)$ is now the volume of phase space enclosed by the hypersurface of energy U [7]

$$\Phi(U) \doteq \int_{\mathbf{z}\in\{\mathbf{z}|H(\mathbf{z})\leq U\}} d\mathbf{z} = \int d\mathbf{z}\ \theta(U - H(\mathbf{z})), \tag{13}$$

where the Hamiltonian is $H(\mathbf{z})$ and the step function $\theta(U - H(\mathbf{z}))$ provides the limits for the integral. The phase space coordinate $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ consists of the set of canonical coordinates $\mathbf{q} = \{q_i\}_{i=1}^n$ in n-dimensional space and the set of their conjugate canonical momenta $\mathbf{p} = \{p_i\}_{i=1}^n$. The element of volume in $2n$-dimensional phase space is $d\mathbf{z} = d^n q\, d^n p$ and integration is over all phase space if no limits are shown.

For the classical case, the constraints on normalization and average energy corresponding to Eq. (7) are

$$\int d\mathbf{z}\rho(\mathbf{z}) = 1, \quad \int d\mathbf{z}\rho(\mathbf{z})H(\mathbf{z}) = U, \tag{14}$$

respectively, where $\rho(\mathbf{z})$ is the probability density in phase space. The MEP for the classical microcanonical ensemble is analogous to the quantum case. Using Eq. (13) and the constraints of normalization and average energy in Eq. (14), we can rewrite the entropy in Eq. (12) as a functional

$$S[\rho] = \log \int d\mathbf{z}'\ \theta\left(\int d\mathbf{z}''\rho(\mathbf{z}'')H(\mathbf{z}'') - H(\mathbf{z}')\int d\mathbf{z}''\rho(\mathbf{z}'')\right), \tag{15}$$

where the integration is over all phase space. The continuous probability distribution $\rho = \rho(\mathbf{z})$ for the microcanonical ensemble is obtained when this entropy is an extremum. Functionally differentiating Eq. (15) with respect to $\rho(\mathbf{z})$ and setting the result equal to zero, we obtain

$$\frac{\delta S[\rho]}{\delta \rho(\mathbf{z})} = \frac{1}{\Phi(U)} \int d\mathbf{z}'\delta\left(U - H(\mathbf{z}')\right)\left(H(\mathbf{z}) - H(\mathbf{z}')\right) = 0. \tag{16}$$

By inspection we see that this equation is satisfied if $\mathbf{z}'$ is such that $H(\mathbf{z}') \neq U$. When $H(\mathbf{z}') = U$ for some values $\mathbf{z}'$ we must also have $H(\mathbf{z}') = H(\mathbf{z})$ for some values of $\mathbf{z}$ [because $x\delta(x) = 0$]. In the latter case we therefore have $H(\mathbf{z}) = U$. The distribution function $\rho(\mathbf{z})$ therefore has a delta function that restricts the Hamiltonian to the hypersurface of energy U,

$$\rho(\mathbf{z}) = A(\mathbf{z}) \, \delta(U - H(\mathbf{z})), \tag{17}$$

where $A(\mathbf{z})$ is an arbitrary function of $\mathbf{z}$, Since there is nothing to distinguish different points in phase space $\mathbf{z} \in \{\mathbf{z}|H(\mathbf{z}) = U\}$ that are all on the energy hypersurface, we can invoke Laplace's Principle of Insufficient Reason to choose $A(\mathbf{z}) = A_U$, which is constant for fixed U for all these phase space points. The normalization condition in Eq. (14) then becomes

$$\int d\mathbf{z}\rho(\mathbf{z}) = \int d\mathbf{z} \, A(\mathbf{z}) \, \delta(U - H(\mathbf{z})) = A_U \int d\mathbf{z} \, \delta(U - H(\mathbf{z})) = 1. \tag{18}$$

The last integral in Eq. (18) can be performed by making a change of variables to $e = H(\mathbf{z})$, which gives

$$\int d\mathbf{z} \, \delta(U - H(\mathbf{z})) = \int de \frac{d\mathbf{z}}{de}\delta(U - e) = \left(\frac{d\mathbf{z}}{de}\right)_{e=U} \equiv \Omega(U), \tag{19}$$

where the function $\Omega(U)$ is the density of states for energy U, $i.e.$, the number of states per unit energy. Substituting Eq. (19) into Eq. (18), we obtain $A_U = \Omega(U)^{-1}$. Therefore, the probability distribution function in phase space in Eq. (17) becomes

$$\rho(\mathbf{z}) = \frac{1}{\Omega(U)}\delta(U - H(\mathbf{z})), \tag{20}$$

which is in fact the well-known classical microcanonical distribution. If the phase space point $\mathbf{z}$ is not on the energy hypersurface $U = H(\mathbf{z})$ the probability density is zero. This probability density is analogous to the probability distribution in Eq. (11) for the quantum case, where the degeneracy g_β corresponds to the density of states $\Omega(U)$.

V. CONCLUSION

In this work we have reviewed the traditional information-theoretic approach to the microcanonical ensemble. In contrast to the derivation of the canonical ensemble, the maximization for the traditional approach to the microcanonical ensemble is performed on a sub-manifold of the Hilbert space (phase space in the classical case) rather than on the whole Hilbert (phase) space. Thus the microcanonical ensemble is assumed in the traditional approach rather than derived. In our approach we have used the Hertz volume entropy with constraints of normalization and average energy to show that it leads to the correct microcanonical distribution.

Some of the significant differences between the traditional approach to the MEP in Eq. (1) and our approach to the MEP in Eqs. (8) and (15) for the microcanonical ensemble are the following.

1. The traditional MEP $assumes$ that the microstate i belongs to energy eigenvalue E_β, whereas our MEP $derives$ such condition.

2. The traditional approach employs the Shannon entropy $without$ an energy constraint, whereas we employ the Hertz volume entropy $with$ the energy constraint.

3. The traditional MEP derives Laplace's Principle of Insufficient Reason for the states belonging to the eigenenergy E_β, whereas our MEP invokes Laplace's principle after deriving the condition that the state i must belong to E_β.

In the latter case we see that our approach, rather than being in contrast with the traditional one, $completes$ it. First one maximizes the Hertz entropy to select the microcanonical energy level. At this point can use the traditional method to find that all states with that energy have the same probability.

---

[1] E. T. Jaynes, Phys. Rev. 106 (1957) 620–630.

[2] E. T. Jaynes, Phys. Rev. 108 (1957) 171–190.

[3] C. Tsallis, J. Stat. Phys. 52 (1988) 479.

[4] P. Hertz, *Uber die mechanischen Grundlagen der Thermodynamik*. Annalen der Physik (Leipzig) 33 (1910) 225-274 and 537-552.

[5] J. W. Gibbs, Elementary Principles in Statistical Mechanics, New Haven, Yale University Press, 1902, p.170. Reprinted by Dover Publications, New York, 1960.

[6] L. Boltzmann, Crelle's Journal 98 (1884) 68–94. Reprinted in Hasenöhrl (ed.), Wissenschaftlic Abhandlungen, vol. 3. New York, Chelsea, pp. 122-152.

[7] M. Campisi, Studies in History and Philosphy of Modern Physics 36 (2005) 275–290.

[8] G. Gallavotti, Statistical mechanics. A short treatise, Springer Verlag, Berlin, 1995.

[9] D. Ruelle, Statistical mechanics: rigorous results, New York, W. A. Benjamin, 1969.

[10] M.Campisi, Physica A 385 (2007) 501-517,

[11] M. Campisi and G. B. Bagci, Phys. Lett. A 362 (2007) 11-15.

[12] M. Campisi, Phys. Lett. A, 366 (4-5) (2007) 335-338.

[13] A. I. Khinchin, Mathematical foundations of statistical mechanics, Dover Publications, New York, 1949.

[14] V. Gurarie, Am. J. Phys. 75 (2007) 747-751.