MDPI

Article

Calibration Invariance of the MaxEnt Distribution in the Maximum Entropy Principle

Jan Korbel [1,2,3]

1 Section for the Science of Complex Systems, Center for Medical Statistics, Informatics, and Intelligent Systems (CeMSIIS), Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria; jan.korbel@meduniwien.ac.at
2 Complexity Science Hub Vienna, Josefstädterstrasse 39, 1080 Vienna, Austria
3 Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, 11519 Prague, Czech Republic

Abstract: The maximum entropy principle consists of two steps: The first step is to find the distribution which maximizes entropy under given constraints. The second step is to calculate the corresponding thermodynamic quantities. The second part is determined by Lagrange multipliers' relation to the measurable physical quantities as temperature or Helmholtz free energy/free entropy. We show that for a given MaxEnt distribution, the whole class of entropies and constraints leads to the same distribution but generally different thermodynamics. Two simple classes of transformations that preserve the MaxEnt distributions are studied: The first case is a transform of the entropy to an arbitrary increasing function of that entropy. The second case is the transform of the energetic constraint to a combination of the normalization and energetic constraints. We derive group transformations of the Lagrange multipliers corresponding to these transformations and determine their connections to thermodynamic quantities. For each case, we provide a simple example of this transformation.

Keywords: maximum entropy principle; MaxEnt distribution; calibration invariance; Lagrange multipliers

1. Introduction

The maximum entropy principle (MEP) is one of the most fundamental concepts in equilibrium statistical mechanics. It was originally proposed by Jaynes [1,2] in order to connect information entropy introduced by Shannon and thermodynamic entropy introduced by Clausius, Boltzmann, and Gibbs. Although the MEP was originally introduced for the case of Shannon entropy, with the advent of generalized entropies [3–17] the natural effort was to apply the maximum entropy principle beyond the case of Shannon entropy. Another question that arose naturally is whether the MEP can be applied to other than ordinary linear constraints. Examples of the constraints that might be considered in connection with the MEP are *escort constraints* [18–20], *Kolmogorov–Nagumo means* [21,22], or more exotic types of constraints [23]. It brought some discussion about the applicability of the principle for the case of generalized entropies [24,25] and nonlinear constraints and its thermodynamic interpretation [26–30]. Indeed, MEP is not the only one extremal principle in statistical physics, let us mention, e.g., the *principle of maximum caliber* [31] which is useful in non-equilibrium physics. In this paper, we stick, however, to MEP, as it is the most widespread principle and the theory of generalized thermostatistics has been mainly focused on MEP. For a recent review of other principles, see also in [32]. For the discussion between entropy arising from information theory and thermodynamics, see in [33]. For the sake of simplicity, let us consider canonical ensemble, i.e., fluctuations in internal energy. For the case of the grand-canonical ensemble, one can obtain similar results to the ones presented in this paper for the case of a chemical potential μ.

In order to grasp the debate about the applicability of the MEP, let us emphasize that the MEP consists of two main parts:

(I)　Finding a distribution (*MaxEnt* distribution) that maximizes entropy under given constraints.

(II)　Plugging the distribution into the entropic functional and calculating physical quantities as thermodynamic potentials, temperature, or response coefficients (specific heat, compressibility, etc.).

The first part is rather a mathematical procedure of finding a maximum subject to constraints. This is done by the *method of Lagrange multipliers*, by defining a Lagrange function in the form

$$Lagrange\ function = entropy - (Lagrange\ multiplier) \cdot (constraint)$$

The Lagrange multipliers' role at this stage is to ensure fulfillment of constraints as they are determined from the set of equations obtained from the maximization of the Lagrange function. This procedure is known in statistics as *Softmax*, a method used to infer distribution from given data. Shore and Johnson [34,35] therefore studied MEP as a statistical inference procedure and established a set of consistency axioms. Shore and Johnson's work heated a debate about whether MEP for generalized entropies can be also understood as a statistical inference method satisfying the consistency requirements [24,36–41]. In [42], it was shown that the class of entropies satisfying the original Shore–Johnson axioms is wider than previously thought. Moreover, in [43], the connection between Shore–Johnson axioms and Shannon–Khinchin axioms was investigated and the equivalence of information theory and statistical inference axiomatics was established.

In the second part, the physical interpretation of entropy starts to arise. Similar to the case of Lagrangian mechanics, where the Lagrangian is the difference between kinetic and potential energy and the Lagrange multipliers play the role of the normal force to the constraints, here the entropy becomes a thermodynamic state variable. For Shannon entropy and linear constraints, the Lagrange multipliers become inverse temperature and free entropy, respectively.

The main aim of this paper is to discuss the relation between points (I) and (II). In the first part, it is possible to find a class of entropic functionals and constraints leading to the same MaxEnt distribution. However, in the second part, different entropy and/or constraints lead to different thermodynamics and different relations between physical quantities and Lagrange multipliers. The two main messages of this paper are listed below.

(i)　For each MaxEnt distribution, there exists the whole class of entropies and constraints leading to generally different thermodynamics.

(ii)　It is possible to establish transformation relations of Lagrange parameters (and subsequently the thermodynamic quantities) for classes of entropies and constraints giving the same MaxEnt distribution.

We call the latter transformation relation *calibration invariance* of the MaxEnt distribution. A straightforward consequence is that in order to fully determine the statistical properties of a thermal system in equilibrium, it is not enough to measure the statistical distribution of energies.

The rest of the paper is organized as follows. In the next section, we briefly discuss the main aspects of MEP for the case of general entropic functional and general constraints. In the following two sections, we introduce two simple transformations of entropic functional (Section 3) and constraints (Section 4) that lead to the same MaxEnt distribution and derive transformations between the Lagrange multipliers. These transformations form a group. After the general derivation, we provide a few simple examples for each case. The last section is devoted to conclusions.

2. Maximum Entropy Principle in Statistical Physics

Maximum entropy principle is the way of obtaining the representing probability distribution from the limited amount of information. Our aim is to find the probability

distribution of the system $P = \{p_i\}_{i=1}^n$ under the set of given constraints. In the simplest case, the principle can be formulated as follows.

Maximum entropy principle: *Maximize entropy $S(P)$ under the normalization constraint $f_0(P) = 0$, and energy constraint $f_E(P) = 0$.*

The normalization condition is considered in the regular form, i.e., $f_0(P) = \sum_i p_i - 1 = \langle 1 \rangle - 1$. Moreover, we have a class of constraints, which originally described the average energy of the system. Therefore, we call them *energy constraints*. We consider only one energy constraint, for simplicity, although there can be more constraints, and they do not have to consider only internal energy but also other thermodynamic quantities. In the original formulation, the energy constraint is linear in probabilities, i.e.,

$$f_E(P) = \sum_i p_i E_i - E = \langle E \rangle - E, \tag{1}$$

but it can be generally any nonlinear function of probabilities—escort means provide an example. A large class of energy constraints can be written in a *separable form*, which means that $f_E(P) = \mathcal{E}(P) - E$, i.e., in the form expressing the "expected" internal energy (macroscopic variable) as a function of probability distribution (microscopic variable). This class of constraints plays a dominant role in the thermodynamic systems.

In order to find a solution of the Maximum entropy principle, we use a common method of Lagrange multipliers, which can be done through maximization of Lagrange function:

$$\mathcal{L}(P; \alpha, \beta) = S(P) - \alpha f_0(P) - \beta f_E(P) \tag{2}$$

The maximization procedure leads to the set of equations

$$\begin{aligned}
\frac{\partial \mathcal{L}(P; \alpha, \beta)}{\partial p_i} &= 0 \quad \forall\, i \in \{1, \ldots, n\} \\
\frac{\partial \mathcal{L}(P; \alpha, \beta)}{\partial \alpha} &= f_0(P) = 0 \\
\frac{\partial \mathcal{L}(P; \alpha, \beta)}{\partial \beta} &= f_E(P) = 0
\end{aligned} \tag{3}$$

from which we determine the resulting MaxEnt distribution. In order to obtain a unique solution, we require that the entropic functional should be a Schur-concave symmetric function [42].

As a consequence, we obtain the values of Lagrange multipliers α and β. From the strictly mathematical point of view, Lagrange multipliers are just auxiliary parameters to be solved from the set of Equation (3). However, in physics, Lagrange parameters also have a physical interpretation. In Lagrangian mechanics, Lagrange parameters play the role of normal force to the constraints. Similarly, in ordinary statistical mechanics based on Shannon entropy $H(P) = -\sum_i p_i \log p_i$ and linear constraints (1), the Lagrange multipliers have the particular physical interpretation:

$$\beta = \frac{1}{T} \quad \text{(inverse temperature)}, \tag{4}$$

$$\alpha = S - \frac{1}{T} E \quad \text{(free entropy)}. \tag{5}$$

Note that the free entropy is, similarly to Helmholtz free energy, a Legendre transform of entropy w.r.t. internal energy. For the case of ordinary thermodynamics (Shannon entropy and linear constraints), it is equal to the logarithm of the partition function.

This interpretation is valid only in this case. In the case, when we use different entropy functional or different constraints, these relation between Lagrange multipliers

and thermodynamic quantities are no longer valid. This is even the case, when the resulting MaxEnt distribution is the same.

The main aim of this paper is to show how the invariance of MaxEnt distribution affects the Lagrange multipliers and their relations to thermodynamic quantities. Let us now solve Equation (3). The first set of equations leads to

$$\frac{\partial S(P)}{\partial p_i} - \alpha \frac{\partial f_0(P)}{\partial p_i} - \beta \frac{\partial f_E(P)}{\partial p_i} = 0. \tag{6}$$

Let us assume the normalization in the usual way which leads to $\frac{\partial f_0(P)}{\partial p_i} = 1$. Moreover, let us consider separable energy constraint, so $\frac{\partial f_E(P)}{\partial p_i} = \frac{\partial \mathcal{E}(P)}{\partial p_i}$. The resulting probability distribution can be expressed as

$$p_i^\star = \frac{\partial S}{\partial p_i}^{(-1)} \left[\alpha + \beta \frac{\partial \mathcal{E}(P)}{\partial p_i} \right]. \tag{7}$$

where $^{(-1)}$ denotes inverse function of $\partial S / \partial p_i$ (provided it exists and is unique). We can express α by multiplying the equation by p_i and summing over i, which leads to

$$\alpha = \langle \nabla_P S(P) \rangle - \beta \langle \nabla_P \mathcal{E}(P) \rangle \tag{8}$$

where $\langle X \rangle = \sum_i x_i p_i$ and $\nabla_P = (\frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_n})$. By plugging back to the previous equation, we can get β as

$$\beta = \frac{\Delta_i(\nabla S(P))}{\Delta_i(\nabla \mathcal{E}(P))} \tag{9}$$

where $\Delta_i(X) = x_i - \langle X \rangle$ is the difference from the average.

The solution of Equation (3) depends on the internal energy E. However, in thermodynamics it is natural to invert the relation $\beta = \beta(E)$ and express the relevant quantities in terms of β, so $E = E(\beta)$. With that, we can calculate dependence of entropy on β:

$$\frac{\partial S}{\partial \beta} = \sum_i \frac{\partial S}{\partial p_i} \frac{\partial p_i}{\partial \beta} = \sum_i \left(\alpha + \beta \frac{\partial \mathcal{E}(P)}{\partial p_i} \right) \frac{\partial p_i}{\partial \beta} = \beta \sum_i \frac{\partial f_E}{\partial p_i} \frac{\partial p_i}{\partial \beta} = \beta \left(-\frac{\partial f_E}{\partial E} \frac{\partial E}{\partial \beta} \right) \tag{10}$$

For separable energy constraints, $\frac{\partial f_E}{\partial E} = -1$, so we obtain the well-known relation

$$\frac{\partial S}{\partial \beta} = \beta \frac{\partial E}{\partial \beta} \Rightarrow \beta = \frac{\partial S}{\partial E}. \tag{11}$$

Let us now define the Legendre conjugate of entropy called *free entropy* (also called Jaynes parameter [44] or Massieu function [45]):

$$\psi = S - \frac{\partial S}{\partial E} E = S - \beta E \tag{12}$$

Free entropy is connected to Helmholtz free energy as $\psi = -\beta F$. The difference between α and ψ can be expressed as

$$\psi - \alpha = (S - \langle \nabla_P S \rangle) - \beta(E - \langle \nabla_P \mathcal{E} \rangle) \tag{13}$$

Therefore, we can understand the difference $\psi - \alpha$ as the Legendre transform of ψ with respect to P. From this, we see that the difference between ψ and α is a constant (not depending on thermodynamic quantities), if two independent conditions are fulfilled, i.e., $E = \langle \nabla_P \mathcal{E}(P) \rangle$ and $S = \langle \nabla_P S \rangle + a$. The former constraint leads to linear energy

constraints, while the latter one leads to the the conclusion that the entropy must be in trace form $S(P) = \sum_i g(p_i)$. Moreover, the function g has to fulfill the following equation,

$$g(x) - ax = xg'(x) \tag{14}$$

leading to $g(x) = -ax \log(x) + bx$ which is equivalent to Shannon entropy.

In the next sections, we will explore how the transformation of the entropy and the energy constraint that leaves the MaxEnt distribution invariant affects the Lagrange multipliers and their relation to thermodynamic quantities.

3. Calibration Invariance of MaxEnt Distribution with Entropy Transformation

The simplest transformation of Lagrange functional that leaves the MaxEnt distribution invariant is to consider an arbitrary increasing function of entropy, i.e., we replace $S(P)$ by $c(S(P))$, where $c'(x) > 0$. Let us note that this transform preserves the uniqueness of the MEP because it is easy to show that if $S(P)$ is Schur-concave, $c(S(P))$ is also Schur-concave [42] which is a sufficient condition for uniqueness of the MaxEnt distribution.

In this case, the Lagrange equations are adjusted as follows,

$$c'(S(P))\frac{\partial S(P)}{\partial p_i} - \alpha_c \frac{\partial f_0(P)}{\partial p_i} - \beta_c \frac{\partial \mathcal{E}(P)}{\partial p_i} = 0 \tag{15}$$

leading to

$$\alpha_c = c'(S(P))\langle \nabla_P S(P) \rangle - \beta_c \langle \nabla_P \mathcal{E}(P) \rangle \tag{16}$$

and

$$\beta_c = c'(S(P))\frac{\Delta_i(\nabla_P S(P))}{\Delta_i(\nabla_P \mathcal{E}(P))} \tag{17}$$

so we get that the function c causes *rescaling* of α and β, so

$$\alpha_c = c'(S(P))\,\alpha \tag{18}$$
$$\beta_c = c'(S(P))\,\beta \tag{19}$$

while its ratio remains unchanged, i.e., $\alpha_c/\beta_c = \alpha/\beta$. Actually, the set of increasing functions conform a group of Lagrange multipliers, because it is easy to show that the Lagrange parameters related to the entropy $c_1(c_2(S(P)))$

$$\beta_{c_1 \circ c_2} = c_1'(c_2(S(P)) \cdot c_2'(S(P))\,\beta = c_1'(c_2(S(P))\beta_{c_2} \tag{20}$$

which can be described as the group operation $(c_1 \circ c_2) \mapsto c_1'(c_2) \cdot c_2'$.

An important property of this transformation is that it changes the extensive–intensive duality of the conjugated pair of thermodynamic variables and the respective forces while it maintains the distribution. Notably, by changing the entropic functional from extensive (i.e., $S(n) \sim U(n)$) to non-extensive, it changes β from intensive (i.e., size-independent, at least in the thermodynamic limit) to non-intensive, i.e., explicitly size-dependent. This point has been discussed in connection with q-non-extensive statistical physics of [29,30] and the relation to the zeroth law of thermodynamics was shown in [46]. As one can see from the example below, although Rényi entropy and Tsallis entropy have the same maximizer, the corresponding thermodynamics is different. While Rényi entropy is additive (and therefore extensive for systems where $U(n) \sim n$) and the temperature is intensive, Tsallis entropy is non-extensive, and the corresponding temperature explicitly depends on the size of the system.

Let us finally mention that the difference between free entropy and Lagrange parameter α transforms as

$$\psi_c - \alpha_c = (c(S) - c'(S)\langle \nabla_P S(P) \rangle - c'(S)\beta(E - \langle \nabla_P \mathcal{E}(P) \rangle) = c'(S)(\psi - \alpha) + (c(S) - c'(S) \cdot S). \tag{21}$$

While free entropy and other thermodynamic potentials are transformed, the heat change remains invariant under this transformation:

$$\text{đ}Q_c = T_c \text{d}\, c(S) = \frac{T}{c'(S)} c'(S) \text{d}S = T \text{d}S = \text{đ}Q. \tag{22}$$

Example 1. *We exemplify the calibration invariance on two popular examples of closely related entropies.*

- **Rényi entropy and Tsallis entropy**: *Two most famous examples of generalized entropies are Rényi entropy $R_q(P) = \frac{1}{1-q} \ln\left(\sum_i p_i^q\right)$ and Tsallis entropy $S_q(P) = \frac{1}{1-q}\left(\sum_i p_i^q - 1\right)$. Their relation can be expressed as*

$$R_q(P) = c_q(S_q(P)) = \frac{1}{1-q} \ln\left[(1-q)S_q(P) + 1\right] \tag{23}$$

and therefore we obtain that

$$c'_q(S_q(P)) = \frac{1}{1 + (1-q)S_q} = \frac{1}{\sum_i p_i^q}. \tag{24}$$

The difference between free entropy and α can be obtained as

$$\psi_R - \alpha_R = \frac{1}{\sum_i p_i^q}(\psi_S - \alpha_S) + \left(R_q(P) - \frac{S_q(P)}{\sum_i p_i^q}\right). \tag{25}$$

One can therefore see that even though Rényi and Tsallis entropy lead to the same MaxEnt distribution, their thermodynamic quantities, such as temperature or free entropy, are different. Whether the system follows Rényi or Tsallis entropy depends on additional facts, as e.g., (non)-extensivity and (non)-intensivity of thermodynamic quantities.

- **Shannon entropy and Entropy power**: *A similar example is provided with Shannon entropy $H(P) = \sum_i p_i \ln 1/p_i$ and entropy power $\mathcal{P}(P) = \prod_i (1/p_i)^{p_i}$. The relation between them is simply*

$$H(P) = c(\mathcal{P}(P)) = \log(\mathcal{P}(P)), \tag{26}$$

so we obtain that

$$c'(\mathcal{P}(P)) = 1/(\mathcal{P}(P)) = \exp(-H(P)). \tag{27}$$

For the difference between free entropy and α, we obtain that

$$0 = \psi_H - \alpha_H = \frac{1}{\mathcal{P}(P)}(\psi_{\mathcal{P}} - \alpha_{\mathcal{P}}) + (H(P) - 1) \tag{28}$$

from which we get that

$$\psi_{\mathcal{P}} - \alpha_{\mathcal{P}} = \mathcal{P}(P)(1 - \log \mathcal{P}(P)). \tag{29}$$

Therefore, we see that even that the MaxEnt distribution remains unchanged, the relation between α and free energy is different.

4. Calibration Invariance of MaxEnt Distribution with Constraints Transformation

Similarly, one can uncover the invariance of the MaxEnt distribution when the constraints are transformed in a certain way. Generally, if two sets of constraints define the same domain, the resulting Maximum entropy principle should lead to equivalent results. We will not be so general, but we focus on a specific situation, which might be quite interesting for thermodynamic applications. Let us remind two conditions, which we assume: normalization $f_0(P) = 0$ and energy constraint $f_E(P) = 0$. Let us investigate the latter. Similarly to the previous case, it is possible to take any function g of $f_E(P)$, for which $g(y) = 0$

if $y = 0$. More generally, we can also take into account the normalization constraint and replace the original energy condition by

$$g(f_0(P), f_E(P)) = 0 \tag{30}$$

for any $g(x, y)$, for which $g(x, y) = 0 \Rightarrow y = 0$. Let us investigate the Maximum entropy principle for this case. We can express the Lagrange function as

$$\mathcal{L}(P) = S(P) - \alpha_g f_0(P) - \beta_g g(f_0(P), f_E(P)) \tag{31}$$

which leads to a set of equations

$$\frac{\partial S(P)}{\partial p_i} - \alpha_g \frac{\partial f_0(P)}{\partial p_i} - \beta_g \left[G^{(1,0)} \frac{\partial f_0(P)}{\partial p_i} + G^{(0,1)} \frac{\partial \mathcal{E}(P)}{\partial p_i} \right] = 0 \tag{32}$$

where $G^{(1,0)} = \frac{\partial g(x,y)}{\partial x}\big|_{(0,0)}$ and $G^{(0,1)} = \frac{\partial g(y,x)}{\partial x}\big|_{(0,0)}$. We take again into account that $\frac{\partial f_0(P)}{\partial p_i} = 1$, multiply the equations by p_i and some over i. This gives us

$$\alpha_g = \langle \nabla_P S(P) \rangle - \beta_g \left[G^{(1,0)} + G^{(0,1)} \langle \nabla_P \mathcal{E}(P) \rangle \right]. \tag{33}$$

By plugging α_g back, we end with relation for β_g:

$$\beta_g = \frac{1}{G^{(0,1)}} \frac{\Delta_i(\nabla_P S(P))}{\Delta_i(\nabla_P \mathcal{E}(P))}. \tag{34}$$

For α_g we end with

$$\alpha_g = \langle \nabla_P S(P) \rangle - \frac{\Delta_i(\nabla_P S(P))}{\Delta_i(\nabla_P \mathcal{E}(P))} \langle \nabla f_E(P) \rangle \left[1 + \frac{G^{(1,0)}}{G^{(0,1)}} \frac{1}{\langle \nabla_P \mathcal{E}(P) \rangle} \right]. \tag{35}$$

Thus, we end again with rescaling of α_g and β_g, which reads

$$\alpha_g(\alpha, \beta) = \alpha - \frac{G^{(1,0)}}{G^{(0,1)}} \beta, \tag{36}$$

$$\beta_g(\beta) = \frac{\beta}{G^{(0,1)}}. \tag{37}$$

The ratio of Lagrange multipliers is also transformed, so we get

$$\frac{\alpha_g}{\beta_g} = G^{(0,1)} \frac{\alpha}{\beta} - G^{(1,0)}. \tag{38}$$

Again, the set of all functions fulfilling the aforementioned condition conform a group. The group operation can be described by the relation between coefficients $G^{(1,0)}$ and $G^{(0,1)}$ for the composite function $g(x, y) = g_1(x, g_2(x, y))$. We obtain that

$$G^{(1,0)} = G_1^{(1,0)} + G_1^{(0,1)} G_2^{(1,0)} \tag{39}$$

$$G^{(0,1)} = G_1^{(0,1)} G_2^{(0,1)} \tag{40}$$

which leads to group relations

$$\alpha_g(\alpha, \beta) = \alpha_{g_1}(\alpha_{g_2}(\alpha, \beta), \beta_{g_2}(\beta)) - \frac{G_1^{(1,0)}}{G_1^{(0,1)}} \beta_{g_2}(\beta) \tag{41}$$

$$\beta_g(\beta) = \frac{\beta_{g_2}(\beta)}{G_1^{(0,1)}}. \tag{42}$$

Example 2. *Here we mention two simple examples of the aforementioned transformation.*

- **Energy shift:** *Under this scheme, we can assume the constant shift in the energy spectrum. Let us rewrite the constraint $f(P)$ in the following form,*

$$f_E(P) = \sum p_i E_i - E = \sum p_i (E_i - E') - (E - E') \tag{43}$$

which allows us to identify the function $g(x,y)$ as

$$g(x,y) = y - E'x + E' \tag{44}$$

We obtain $G^{(1,0)} = -E'$ and $G^{(0,1)} = 1$, which means that $\alpha' = \alpha - \beta E'$.

- **Latent escort means:** *Apart from linear means, it is possible to use some generalized approaches. One of these examples is provided by so-called escort mean:*

$$E_q = \langle E \rangle_q = \frac{\sum_i p_i^q E_i}{\sum_i p_i^q} \tag{45}$$

which for $q = 1$ becomes an ordinary linear mean, when $P = \{p_i\}_{i=1}^n$ are normalized to one. When we use this class of means in the Maximum entropy principle, the normalization is enforced by the normalization condition $f_0(P) = 0$, therefore for $q = 1$ we obtain the same results. Nevertheless, by taking $q = 1$ for the results with escort distribution, the energy constraint is actually expressed as

$$\frac{\sum p_i E_i}{\sum p_i} - E \tag{46}$$

can be understood in the same way as considered before in this section, i.e., as a combination of a normalization constraint and energy constraint. In this case the function g has the following form,

$$g(x,y) = \frac{y + E}{x + 1} - E. \tag{47}$$

Therefore, we obtain that $G^{(1,0)} = -E$ and $G^{(0,1)} = 1$, which correspond to the previous example for $E' = E$. Therefore, the latent energy mean can be understood in terms of MaxEnt procedure as the shift of the energy spectrum by its average energy.

5. Conclusions

In this paper, we have discussed the calibration invariance of MEP, which means that for a given MaxEnt distribution, there exists a whole class of entropies and constraints that lead to different thermodynamics (Thermodynamic quantities and response coefficients generally have different behavior. For example, from intensive temperature we can obtain temperature that explicitly depends on the size of the system). We have stressed that the MEP procedure consists of two parts, where the first part, consisting of determining the MaxEnt distribution, is rather a mathematical tool, while the second part, making connection between Lagrange multipliers and thermodynamic quantities, is a specific for application of MEP in statistical physics. Indeed, the paper does not cover all possible transformations leading to the same MaxEnt distribution (let us mention, at least, the additive duality of Tsallis entropy, where maximizing S_{2-q} with linear constraint leads to the same result as maximizing S_q with escort constraints [47]). The main lesson of this paper is that in order to fully determine a thermal system in equilibrium, we need to measure not only probability distribution, but also all relevant thermodynamic quantities (as entropy). Moreover, the transformation between Lagrange parameters and its connection to thermodynamic potentials can be useful in situations when one is not certain about the exact form of entropy.

References

1. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620. [CrossRef]
2. Jaynes, E.T. Information Theory and Statistical Mechanics. II. *Phys. Rev.* **1957**, *108*, 171. [CrossRef]
3. Burg, J.P. The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophysics* **1972**, *37*, 375–376. [CrossRef]
4. Rényi, A. *Selected Papers of Alfréd Rényi*; Akademia Kiado: Budapest, Hungary, 1976; Volume 2.
5. Havrda, J.H.; Charvát, F. Quantification Method of Classification Processes. Concept of Structural *α*-Entropy. *Kybernetika* **1967**, *3*, 30–35.
6. Sharma, B.D.; Mitter, J.; Mohan, M. On Measures of "Useful" Information. *Inf. Control* **1978**, *39*, 323–336. [CrossRef]
7. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
8. Frank, T.; Daffertshofer, A. Exact time-dependent solutions of the Renyi Fokker-Planck equation and the Fokker-Planck equations related to the entropies proposed by Sharma and Mittal. *Physica A* **2000**, *285*, 351–366. [CrossRef]
9. Kaniadakis, G. Statistical mechanics in the context of special relativity. *Phys. Rev. E* **2002**, *66*, 056125. [CrossRef]
10. Jizba, P.; Arimitsu, T. The world according to Rényi: Thermodynamics of multifractal systems. *Ann. Phys.* **2004**, *312*, 17–59. [CrossRef]
11. Hanel, R.; Thurner, S. A comprehensive classification of complex statistical systems and an ab-initio derivation of their entropy and distribution functions. *Europhys. Lett.* **2011**, *93*, 20006. [CrossRef]
12. Thurner, S.; Hanel, R.; Klimek, P. *Introduction to the Theory of Complex Systems*; Oxford University Press: Oxford, UK, 2018.
13. Korbel, J.; Hanel, R.; Thurner, S. Classification of complex systems by their sample-space scaling exponents. *New J. Phys.* **2018**, *20*, 093007. [CrossRef]
14. Tempesta, P.; Jensen, H.J. Universality classes and information-theoretic Measures of complexity via Group entropies. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef]
15. Ilić, V.M.; Stankovixcx, M.S. Generalized Shannon-Khinchin axioms and uniqueness theorem for pseudo-additive entropies. *Physica A* **2014**, *411*, 138–145. [CrossRef]
16. Ilić, V.M.; Scarfone, A.M.; Wada, T. Equivalence between four versions of thermostatistics based on strongly pseudoadditive entropies. *Phys. Rev. E* **2019**, *100*, 062135. [CrossRef]
17. Czachor, M. Unifying Aspects of Generalized Calculus. *Entropy* **2020**, *22*, 1180. [CrossRef]
18. Beck, C.; Schlögl, F. *Thermodynamics of Chaotic Systems: An Introduction*; Cambridge University Press: Cambridge, UK, 1993.
19. Abe, S. Geometry of escort distributions. *Phys. Rev. E* **2003**, *68*, 031101. [CrossRef] [PubMed]
20. Bercher, J.-F. On escort distributions, q-gaussians and Fisher information. *AIP Conf. Proc.* **2011**, *1305*, 208.
21. Czachor, M.; Naudts, J. Thermostatistics based on Kolmogorov-Nagumo averages: Unifying framework for extensive and nonextensive generalizations. *Phys. Lett. A* **2002**, *298*, 369–374. [CrossRef]
22. Scarfone, A.M.; Matsuzoe, H.; Wada, T. Consistency of the structure of Legendre transform in thermodynamics with the Kolmogorov-Nagumo average. *Phys. Lett. A* **2016**, *380*, 3022–3028. [CrossRef]
23. Bercher, J.-F. Tsallis distribution as a standard maximum entropy solution with 'tail' constraint. *Phys. Lett. A* **2008**, *372*, 5657–5659. [CrossRef]
24. Pressé, S.; Ghosh, K.; Lee, J.; Dill, K.A. Nonadditive Entropies Yield Probability Distributions with Biases not Warranted by the Data. *Phys. Rev. Lett.* **2013**, *111*, 180604. [CrossRef] [PubMed]
25. Oikonomou, T.; Bagci, B. Misusing the entropy maximization in the jungle of generalized entropies. *Phys. Lett. A* **2017**, *381*, 207–211. [CrossRef]
26. Tsallis, C.; Mendes, R.S.; Plastino, A.R. The role of constraints within generalized nonextensive statistics. *Phys. A* **1998**, *286*, 534–554. [CrossRef]
27. Martínez, S.; Nicolás, F.; Peninni, F.; Plastino, A. Tsallis' entropy maximization procedure revisited. *Phys. A* **2000**, *286*, 489–502. [CrossRef]
28. Plastino, A.; Plastino, A.R. On the universality of thermodynamics' Legendre transform structure. *Phys. Lett. A* **1997**, *226*, 257–263. [CrossRef]
29. Rama, S.K. Tsallis Statistics: Averages and a Physical Interpretation of the Lagrange Multiplier *β*. *Phys. Lett. A* **2000**, *276*, 103–108. [CrossRef]
30. Campisi, M.; Bagci, G.B. Tsallis Ensemble as an Exact Orthode. *Phys. Lett. A* **2007**, *362*, 11–15. [CrossRef]
31. Dixit, P.D.; Wagoner, J.; Weistuch, C.; Pressé, S.; Ghosh, K.; Dill, K.A. Perspective: Maximum caliber is a general variational principle for dynamical systems. *J. Chem. Phys.* **2018**, *148*, 010901. [CrossRef]
32. Lucia, U. Stationary Open Systems: A Brief Review on Contemporary Theories on Irreversibility. *Physica A* **2013**, *392*, 1051–1062. [CrossRef]
33. Palazzo, P. Hierarchical Structure of Generalized Thermodynamic and Informational Entropy. *Entropy* **2018**, *20*, 553. [CrossRef]
34. Shore, J.E.; Johnson, R.W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theor.* **1980**, *26*, 26–37. [CrossRef]

35. Shore, J.E.; Johnson, R.W. Properties of cross-entropy minimization. *IEEE Trans. Inf. Theor.* **1981**, *27*, 472–482. [CrossRef]
36. Uffink, J. Can the maximum entropy principle be explained as a consistency requirement? *Stud. Hist. Philos. Mod. Phys.* **1995**, *26*, 223–261. [CrossRef]
37. Tsallis, C. Conceptual Inadequacy of the Shore and Johnson Axioms for Wide Classes of Complex Systems. *Entropy* **2015**, *17*, 2853–2861. [CrossRef]
38. Pressé, S.; Ghosh, K.; Lee, J.; Dill, K.A. Reply to C. Tsallis' "Conceptual Inadequacy of the Shore and Johnson Axioms for Wide Classes of Complex Systems". *Entropy* **2015**, *17*, 5043–5046. [CrossRef]
39. Oikonomou, T.; Bagci, G.B. Rényi entropy yields artificial biases not in the data and incorrect updating due to the finite-size data. *Phys. Rev. E* **2019**, *99*, 032134. [CrossRef]
40. Jizba, P.; Korbel, J. Comment on "Rényi entropy yields artificial biases not in the data and incorrect updating due to the finite-size data". *Phys. Rev. E* **2019**, *100*, 026101 [CrossRef]
41. Oikonomou, T.; Bagci, G.B. Reply to "Comment on Rényi entropy yields artificial biases not in the data and incorrect updating due to the finite-size data". *Phys. Rev. E* **2019**, *100*, 026102. [CrossRef]
42. Jizba, P.; Korbel, J. Maximum Entropy Principle in Statistical Inference: Case for Non-Shannonian Entropies. *Phys. Rev. Lett.* **2019**, *122*, 120601. [CrossRef]
43. Jizba, P.; Korbel, J. When Shannon and Khinchin meet Shore and Johnson: Equivalence of information theory and statistical inference axiomatics. *Phys. Rev. E* **2020**, *101*, 042126. [CrossRef]
44. Plastino, A.; Plastino, A.R. Tsallis Entropy and Jaynes' Information Theory Formalism. *Braz. J. Phys.* **1999**, *29*, 50–60. [CrossRef]
45. Naudts, J. *Generalized Thermostatistics*; Springer: London, UK, 2011.
46. Biró, T.S.; Ván, P. Zeroth law compatibility of nonadditive thermodynamics. *Phys. Rev. E* **2011**, *83*, 061147. [CrossRef] [PubMed]
47. Wada, T.; Scarfone, A.M. Connections between Tsallis' formalisms employing the standard linear average energy and ones employing the normalized q-average energy. *Phys. Lett. A* **2005**, *335*, 351–362. [CrossRef]