Consistency of the structure of Legendre transform in thermodynamics with the Kolmogorov-Nagumo average

A.M. Scarfone

Istituto dei Sistemi Complessi (ISC-CNR) c/o Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy.

H. Matsuzoe

Department of Computer Science and Engineering, Nagoya Institute of Technology
Nagoya 466-8555, Japan

T. Wada

Department of Electrical and Electronic Engineering, Ibaraki University
Nakanarusawacho, Hitachi 316-8511, Japan

Abstract

We show the robustness of the structure of Legendre transform in thermodynamics against the replacement of the standard linear average with the Kolmogorov-Nagumo nonlinear average to evaluate the expectation values of the macroscopic physical observables. The consequence of this statement is twofold: 1) the relationships between the expectation values and the corresponding Lagrange multipliers still hold in the present formalism; 2) the universality of the Gibbs equation as well as other thermodynamic relations are unaffected by the structure of the average used in the theory.

Keywords: Kolmogorov-Nagumo average, structure of Legendre transform, Sharma-Mittal entropy, Sharma-Taneja-Mittal entropy, Rényi entropy

1. Introduction

In the last decades, the interest toward generalized entropic forms is growing after the more and more large evidence that power-law distributions are recurrently found in the description of the statistical proprieties of several complex systems running from the physical and biological realm to the social and economical sciences. In order to account for such phenomenologies several entropic forms have been proposed.
Basically, the most of them can be grouped in two different large classes. Often, they are obtained as continuous deformation of the well-known Shannon

or Boltzmann-Gibbs (SBG) entropy, by means of suitable deformation parameters. The first class corresponds to the trace-like entropic forms, defined as the linear average of an appropriate Hartley information function that represents the elementary information gain

$$S[p] = \langle I(p) \rangle_{\text{lin}} \ , \tag{1}$$

where $\langle x \rangle_{\text{lin}} = \sum_i p_i \, x_i$. In particular, for the Shannon information, the Hartley function is given by $I(p) = -\ln p$. The Sharma-Tanja-Mittal entropy [1], defined in

$$S_{\alpha\gamma}^{\text{STM}}[p] = \langle I_{\alpha\gamma}(p) \rangle_{\text{lin}} \ , \tag{2}$$

belongs to this class. In Eq. (2)

$$I_{\alpha\gamma}(p_i) = -\frac{p_i^\alpha - p_i^\gamma}{\alpha - \gamma} \ , \tag{3}$$

where α and γ are two real deformation parameters and hereinafter, $p = \{p_i \geq 0; \sum_i p_i = 1; (i = 1, \ldots, W)\}$ denotes a normalized discrete distribution function.

The second class is given by the kernel-like entropic forms defined in

$$S[p] = K \left(\langle I(p) \rangle_{\text{lin}} \right) \ , \tag{4}$$

that are functions of trace-like entropic forms. The Sharma-Mittal entropy [2], defined in

$$S_{\alpha\gamma}^{\text{SM}}[p] = K_{\alpha\gamma} \left(\langle I_\alpha(p) \rangle_{\text{lin}} \right) \ , \tag{5}$$

belongs to this class. In Eq. (5)

$$I_\alpha(p_i) = p_i^{\alpha-1} \ , \tag{6}$$

and the kernel function $K_{\alpha\gamma}(x)$, is given by

$$K_{\alpha\gamma}(x) = \frac{x^{\frac{1-\gamma}{1-\alpha}} - 1}{1 - \gamma} \ . \tag{7}$$

Both the entropy families (2) and (5) have been introduced previously in information theory and then rediscovered in statistical mechanics. Several one-parameter deformations of the SBG entropy introduced in statistical physics belong to these families. Among them, the Kaniadakis-entropy, the Abe-entropy and the Tsallis-entropy are special cases of the STM-entropy (see [3] and references therein), whilst the second family embodies the Tsallis-entropy, the Rényi entropy, the escort entropy, the Landsberg-Vedral entropy and others (see [4] and reference therein).
Often, in statistical mechanics, the approach to obtain the equilibrium distribution of the system, is based on the variational principle [5]. It consists in

maximizing the entropic form under suitable constraints given by the values of several thermodynamic variables. Consistence with definitions (1) and (4) requires that constraints are introduced by means of linear average.

A different approach to look at no-trace-form entropies, often ignored in the field of the statistical mechanics, exists. It is based on the nonlinear (o quasi-linear) average originally advanced by Kolmogorov and Nagumo [6] and then generalized by de Finetti [7]. The Kolmogorov-Nagumo (KN) average reads

$$\langle x \rangle_{\mathrm{KN}} = f^{-1} \left(\sum_i p_i \, f(x_i) \right) \ , \tag{8}$$

where $f(x)$ is an arbitrary strictly monotonic and continuous function, first time employed in statistics by Rényi in the seminal work [8]. Today, the KN average is recurrently used in the framework of statistical mathematics and information theory [9], however, as already observed, it is often ignored in statistical mechanics.

In analogy with Eq. (2), one could define a kernel-like entropy by means of a quasi-linear trace-like form with the KN average for a suitable Hartley information function, that is

$$S[p] = \langle \tilde{I}(p) \rangle_{\mathrm{KN}} \ . \tag{9}$$

The two definitions (4) and (9) are equivalent each other for $K(x) \equiv f^{-1}(x)$ and $I(x) = f(\tilde{I}(x))$.

Actually, as suggested by some Authors [10, 11], if we employ the KN average in the definition of the entropy, then the same prescription should be used, for consistence, in the construction of the constraints used in the variational problem. This may change significantly the shape of the distribution at equilibrium.

In order to fruitfully introduce the KN average in statistical mechanics and thermodynamics, we wonder which impact on the epistemological structure of the emerging theory it has. The structure of the Legendre transform of the theory is an important aspect to be considered. This is a fundamental question that joins the two theories since, the Legendre transform structure binds the phenomenological aspect of thermodynamics with the microscopical aspect of statistical mechanics and, for a successful application of the KN average in statistical mechanics, such mathematical structure should be reproduced in some way.

In the past, the robustness of the Legendre transform structure has been shown to hold [12] when the thermostatistic theory is based on trace-like entropy and the thermodynamic observables are defined by linear averages. This result has been generalized in [13] for the most general expression of the constraints including, in principle, the KN average. However, at the best of our knowledge, no such investigation has been done when the entropic form is of kernel-like, as in Eq. (9). Our mayor effort in this work is to elucidate this problem.

The plan of the paper is as it follows. In the next section 2, we revisit the Jaynes variational principle in the framework of the KN nonlinear average, while,

3

in section 3, we derive the structure of Legendre transform of thermodynamics based on the KN average and we obtain the corresponding Gibbs equation and other related thermostatistic relations. A concluding summary is given in the final section 4 whilst in Appendix A, as an example, we explicit our results in the case of the SM family.

2. Jaynes variational principle in the framework of Kolmogorov-Nagumo average

Let us consider a statistical system described by the entropic form (9), with a suitable Hartley function $\tilde{I}(x)$ that we suppose to be monotonic and normalizable but not necessarily additive. By introducing the function $\Lambda(x) = f_1(I(x))$, we rewrite the entropy (9) in

$$S[p] = f_1^{-1}\left(\langle\Lambda(p)\rangle_{\mathrm{lin}}\right) , \tag{10}$$

where $\Lambda(x)$ is assumed to be a strictly positive, monotonic decreasing $d\Lambda(x)/dx < 0$ and convex $d^2\Lambda(x)/dx^2 > 0$ function.

Let $\mathcal{O}^{(j)}$, with $j = 1, \ldots, M$, be a set of macroscopical quantities that, with abuse of language, we call thermodynamic variables. Consistently with definition (9), we pose $\mathcal{O}^{(j)} = \langle o^{(j)}\rangle_{\mathrm{KN}}$ so that, they correspond to the KN averages of M random variables, where $o_{i_j}^{(j)}$ ($i_j = 1, \ldots, n_j$, with $n_j \gg M$), are the respective possible outcomes. Thus

$$\mathcal{O}^{(j)} = f_2^{-1}\left(\left\langle f_2\left(o^{(j)}\right)\right\rangle_{\mathrm{lin}}\right) , \tag{11}$$

where, for sake of exposition, we have introduced in Eqs. (10) and (11) two different Kolmogorov functions, although in general $f_1(x) = f_2(x) \equiv f(x)$. By posing $\theta_i^{(j)} = f_2\left(o_{i_j}^{(j)}\right)$, the M constraints (11) can be written as functions of the linear average of $\theta_{i_j}^{(j)}$, or equivalently

$$f_2\left(\mathcal{O}^{(j)}\right) = \langle\theta^{(j)}\rangle_{\mathrm{lin}} . \tag{12}$$

Therefore, the KN average and the linear average are related each other according to

$$f_2\left(\langle x\rangle_{\mathrm{KN}}\right) = \langle f_2(x)\rangle_{\mathrm{lin}} . \tag{13}$$

In the same way, we can introduce the trace-like entropy

$$\Sigma[p] = f_1(S[p]) \equiv \langle\Lambda(p)\rangle_{\mathrm{lin}} . \tag{14}$$

We call $\Sigma[p]$ the associated entropic form to the kernel-like entropy $S[p]$.

In order to obtain the probability distribution for the system at equilibrium, we apply the Jaynes maximum entropy principle that consists in solving

the extremal problem for the maximum of the entropy $S[p]$, under the $M + 1$ constraints given by the thermodynamic variables $\mathcal{O}^{(j)}$ and the normalization

$$\sum_{\{i\}} p_{\{i\}} = 1 \ . \tag{15}$$

In the remainder, for sake of exposition, we employ a multi-index notation with $p_{\{i\}} \equiv p_{i_1, i_2, \ldots i_M}$ and summation run over all possible combinations, that is

$$\sum_{\{i\}} p_{\{i\}} \equiv \sum_{i_1, i_2, \ldots i_M} p_{i_1, i_2, \ldots i_M} \ . \tag{16}$$

In this notation $\{i\} = (i_1, \ldots, i_M)$ is an unordered M-tuple of integers, with entries $1 \leq i_j \leq n_j$ indicating the possible outcome of j-th observable $o_{i_j}^{(j)}$.
Therefore, let β_j and β_0 be the respective $M + 1$ Lagrange multipliers, we consider the following variational problem

$$\delta \left[f_1^{-1} \left(\sum_{\{i\}} p_{\{i\}} \Lambda \left(p_{\{i\}} \right) \right) - \beta_0 \sum_{\{i\}} p_{\{i\}} - \sum_j \beta_j f_2^{-1} \left(\sum_{\{i\}} p_{\{i\}} \theta_{i_j}^{(j)} \right) \right] = 0 \ , \tag{17}$$

that is solved for the unknown quantities $p_{\{i\}}$. The Lagrange multipliers, that ultimately turn out to be function only of the quantities $\mathcal{O}^{(j)}$, may be eventually obtained by using constraints (11) and (15).
By performing the variation, we get

$$F_1(S) \left(\Lambda(p_{\{i\}}) + p_{\{i\}} \frac{\partial \Lambda(p_{\{i\}})}{\partial p_{\{i\}}} \right) - \beta_0 - \sum_j \beta_j \, F_2 \left(\mathcal{O}^{(j)} \right) \theta_{i_j}^{(j)} = 0 \ , \tag{18}$$

where the functions $F_i(x)$ are defined in

$$F_i(x) = \left. \frac{\partial f_i^{-1}(y)}{\partial y} \right|_{y = f_i(x)} \ . \tag{19}$$

They reduce to identity $F_i(x) = 1$ in the case of linear average $f_i(x) = x$.
By posing

$$h(x) = \Lambda(x) + x \, \frac{d \, \Lambda(x)}{dx} \ , \tag{20}$$

that is a strictly monotonics function for $x > 0$ and then invertible, we can solve formally Eq. (18) according to

$$p_{\{i\}} = h^{-1} \left(\tilde{\beta}_0 + \sum_j \tilde{\beta}_j \, \theta_{i_j}^{(j)} \right) \ , \tag{21}$$

where

$$\tilde{\beta}_0 = \frac{\beta_0}{F_1(S)} \ , \qquad \tilde{\beta}_j = \beta_j \, \frac{F_2 \left(\mathcal{O}^{(j)} \right)}{F_1(S)} \ . \tag{22}$$

At this step, Eq. (21) turns out to be a function of $2M + 2$ quantities: the entropy S, the Lagrange multipliers β_0 and β_j and the thermodynamic variables $\mathcal{O}^{(j)}$. Clearly, this is a consequence of the nonlinear flavor of the average used in the definition of the entropy and the constraints. In a theory based on the standard linear average, $f_i(x) = x$ and therefore $F_i(x) = 1$, so that $\tilde{\beta} \to \beta$ and the dependence from the entropy disappears.

It is worth to observe that, being $f_1(x)$ a monotonic function, the variational problem for the associate entropy $\Sigma[p]$ gives the same distribution (21), although the relationships between the Lagrange multipliers β_0 and β_j, and the corresponding constraints, differ in the two cases. In fact, from the following variational problem

$$\delta \left[\sum_{\{i\}} p_{\{i\}} \Lambda\big(p_{\{i\}}\big) - \beta_0 \sum_{\{i\}} p_{\{i\}} - \sum_j \beta_j f_2^{-1} \left(\sum_{\{i\}} p_{\{i\}} \theta_{i_j}^{(j)} \right) \right] = 0 \;, \qquad (23)$$

it is straightforward to obtain the distribution

$$p_{\{i\}} = h^{-1} \left(\beta_0 + \sum_j \bar{\beta}_j \, \theta_{i_j}^{(j)} \right) \;, \qquad (24)$$

that is formally equal to Eq. (21) while $\bar{\beta}_j = \beta_j \, F_2(\mathcal{O}^{(j)})$. Finally, if we use, in this case, the linear average for constraints coherently with the trace form of entropy $\Sigma[p]$, we obtain a simplest expression for the distribution given by $p_{\{i\}} = h^{-1}(\beta_0 + \sum_j \beta_j \, o_{i_j}^{(j)})$.

3. Legendre transform structure, Gibbs equation and the reciprocity relations

In standard thermostatistics, the Legendre structure of the theory establishes a deep connection between thermodynamics and statistical mechanics. From one hand, it is introduced to derive potential functions to study particular thermodynamics transformations. On the other hand, it is used to establish a relationships between the thermodynamic variables that, in the orthodox theory are assumed to be extensive quantities, and their conjugate variables, corresponding to the Lagrange multipliers, that in the orthodox theory are assumed to be intensive quantities.

A different interpretation of the Legendre transform structure arises in the framework of the geometric thermodynamics. As firstly suggested by Hermann [14], thermodynamic manifold can be considered like a $2M + 1$ dimensional space, derived from the thermodynamic principles [15]. The equilibrium configuration is then represented by a M dimensional sub-manifold, named Legendre manifold, where only M quantities, typically identified with the thermodynamic variables $\mathcal{O}^{(j)}$, are assumed to be independent. They play the rôle of coordinates. In the Legendre manifold, the Gibbs equation holds. Otherwise, the

remaining $M + 1$ variables, given by the Lagrange multipliers β_j and the entropy S, are assumed to be functions of $\mathcal{O}^{(j)}$. In this geometrical pictures, the Legendre transform introduces a kind of coordinate transformation in the Legendre manifold, so that one can define new thermodynamic functions on the equilibrium manifold where the rôle of the independent quantities is, partially or completely, interchanged.

This interpretation of the Legendre transform and its related structure still holds in the framework of the KN average if the equilibrium distribution is derived from a variational principle. To show this, let us firstly describe, in a formal way, how distribution (21) can be surrounded into the Legendre manifold. Apparently, this distribution depends on $2\,M+2$ variables. However, these quantities are not all independent since p_i must satisfy $M + 1$ constrained relations. Among them, the normalization condition is saturated by solving equation (15) for β_0. Therefore, $\beta_0 = \beta_0(\beta_1, \ldots, \beta_n) \equiv \beta_0[\beta]$ becomes a function of the other Lagrange multipliers. In addition, the M relations (11) can be solved for the corresponding β_j so that, $\beta_j = \beta_j \left[\mathcal{O}^{(1)}, \ldots, \mathcal{O}^{(M)}; S \right] \equiv \beta_j \left[\mathcal{O}; S \right]$. At this step, each Lagrange multiplier depends only on the thermodynamic variables and the entropy. However, by inserting distribution (21) in Eq. (10), we can derive an implicit relation for $S[p]$, that once solved, gives us the entropy as function of the thermodynamic variables only: $S[p] \equiv S\left[\mathcal{O} \right]$ and, ultimately, $\beta_j \equiv \beta_j \left[\mathcal{O} \right]$.

In order to derive the Legendre transform structure of the theory, we firstly evaluate the derivatives of entropy with respect to β_j

$$
\begin{aligned}
\frac{\partial S[p]}{\partial \beta_j} &= \sum_{\{i\}} \frac{\partial S[p]}{\partial p_{\{i\}}} \frac{\partial p_{\{i\}}}{\partial \beta_j} \\
&= F_1(S) \sum_{\{i\}} \left(\Lambda\left(p_{\{i\}}\right) + p_{\{i\}} \frac{\partial \Lambda\left(p_{\{i\}}\right)}{\partial p_{\{i\}}} \right) \frac{\partial p_{\{i\}}}{\partial \beta_j} ,
\end{aligned}
\tag{25}
$$

and accounting for definition (20) and Eqs. (21) and (22), we obtain

$$
\begin{aligned}
\frac{\partial S[p]}{\partial \beta_j} &= F_1(S) \sum_{\{i\}} h(p_{\{i\}}) \frac{\partial p_{\{i\}}}{\partial \beta_j} \\
&= \sum_{\{i\}} \left(\beta_0 + \sum_k \beta_k F_2\left(\mathcal{O}^{(k)}\right) \theta_{i_k}^{(k)} \right) \frac{\partial p_{\{i\}}}{\partial \beta_j} \\
&= \sum_k \beta_k F_2\left(\mathcal{O}^{(k)}\right) \left(\sum_{\{i\}} \theta_{i_k}^{(k)} \frac{\partial p_{\{i\}}}{\partial \beta_j} \right) ,
\end{aligned}
\tag{26}
$$

where we used also the relation $\sum_{\{i\}} dp_{\{i\}} = 0$, which follows from condition (15). On the other hand, from Eq. (11), rewritten in

$$
\mathcal{O}^{(k)} = f_2^{-1} \left(\sum_{\{i\}} p_{\{i\}} \theta_{i_k}^{(k)} \right) ,
\tag{27}
$$

7

we have

$$\frac{\partial \mathcal{O}^{(k)}}{\partial \beta_j} = F_2\left(\mathcal{O}^{(k)}\right) \sum_{\{i\}} \theta_{i_k}^{(k)} \frac{\partial p_{\{i\}}}{\partial \beta_j} , \tag{28}$$

that, inserted in relation (26), gives us the generalized Euler equations

$$\frac{\partial S\left[\mathcal{O}\right]}{\partial \beta_j} = \sum_k \beta_k \frac{\partial \mathcal{O}^{(k)}}{\partial \beta_j} . \tag{29}$$

By multiplying both side by $\partial \beta_j / \partial \mathcal{O}^{(i)}$, and summing on the index j, we finally obtain

$$\frac{\partial S\left[\mathcal{O}\right]}{\partial \mathcal{O}^{(i)}} = \beta_i , \tag{30}$$

that establish the conjugate character between the Lagrange multipliers and the mean values. Therefore, β_j and $\mathcal{O}^{(j)}$ are related each other like the intensive versus extensive variables of thermodynamic, although they have not longer necessarily the same intensive/extensive character.

Actually, a straightforward derivation of Eqs. (30) can be given by using the relation $\sum_{\{i\}} \theta_{i_j}^{(j)} \, dp_{\{i\}} = df_2\left(\mathcal{O}^{(j)}\right)$, as it follows from Eq. (12). Therefore, we have

$$
\begin{aligned}
dS\left[\mathcal{O}\right] &= F_1(S) \sum_{\{i\}} h(p_{\{i\}}) \, dp_{\{i\}} = \sum_j \beta_j \, F_2\left(\mathcal{O}^{(j)}\right) \left(\sum_{\{i\}} \theta_{i_j}^{(j)} \, dp_{\{i\}} \right) \\
&= \sum_j \beta_j \, F_2\left(\mathcal{O}^{(j)}\right) df_2\left(\mathcal{O}^{(j)}\right) \\
&= \sum_j \beta_j \, F_2\left(\mathcal{O}^{(j)}\right) \frac{df_2\left(\mathcal{O}^{(j)}\right)}{d\mathcal{O}^{(j)}} \, d\mathcal{O}^{(j)} ,
\end{aligned}
\tag{31}
$$

and observing that

$$\frac{df_2\left(\mathcal{O}^{(j)}\right)}{d\mathcal{O}^{(j)}} = \frac{1}{F_2\left(\mathcal{O}^{(j)}\right)} , \tag{32}$$

we obtain

$$dS\left[\mathcal{O}\right] = \sum_j \beta_j \, d\mathcal{O}^{(j)} , \tag{33}$$

that corresponds to the Gibbs equation for the present formalism. Accounting for the independence of $d\mathcal{O}^{(j)}$, we recover again Eq. (30).

By means of a partial Legendre transform we can introduce up to $2^M - 2$ different functions called Massieu potentials. They follow from any possible

partition $I + J$ of the set of indices $\{1, \ldots, M\}$ into two disjoint subsets. Thus, function

$$\Phi[\mathcal{O}^{(I)}, \beta_J] = S[\mathcal{O}] - \sum_J \beta_J \, \mathcal{O}^{(J)} \, , \tag{34}$$

depend on the thermodynamic variables $\mathcal{O}^{(i)}$, with $i \in I$, and the Lagrange multipliers β_j, with $j \in J$. In fact

$$d\Phi[\mathcal{O}^{(I)}, \beta_J] \;=\; dS[\mathcal{O}] - \sum_J d\beta_J \, \mathcal{O}^{(J)} - \sum_J \beta_J \, d\mathcal{O}^{(J)} \, , \tag{35}$$

and recalling Eq. (33), we obtain

$$d\Phi[\mathcal{O}^{(I)}, \beta_J] - \sum_I \beta_I \, d\mathcal{O}^{(I)} + \sum_J \mathcal{O}^{(J)} \, d\beta_J = 0 \, , \tag{36}$$

or equivalently

$$\frac{\partial}{\partial \mathcal{O}^{(i)}} \Phi[\mathcal{O}^{(I)}, \beta_J] = \beta_i \, , \qquad \frac{\partial}{\partial \beta_j} \Phi[\mathcal{O}^{(I)}, \beta_J] = -\mathcal{O}^{(j)} \, , \tag{37}$$

that state again the duality between the Lagrange multipliers and the thermodynamic variables.

In contrast, a total Legendre transform, with $J = \{1, \ldots, M\}$, defines a thermodynamic potential $\psi[\beta]$, that is a function of all Lagrange multipliers β_i

$$\psi[\beta] = S[\mathcal{O}] - \sum_j \beta_j \, \mathcal{O}^{(j)} \, , \tag{38}$$

with

$$d\psi[\beta] = - \sum_j \mathcal{O}^{(j)} \, d\beta_j \, . \tag{39}$$

The role of dependent and independent variables is exchanged so that, $\mathcal{O}^{(j)} \equiv \mathcal{O}^{(j)}[\beta]$ are now the dependent variables, functions of the Lagrange multipliers β_i that turn out to be the independent variables. Therefore, Eq. (39) implies

$$\frac{\partial \psi[\beta]}{\partial \beta_i} = -\mathcal{O}^{(i)} \, . \tag{40}$$

Many thermodynamic relations hold also in the present formalism. For instance, from Eqs. (30) and (40), it is straightforward to obtain

$$\frac{\partial \mathcal{O}^{(j)}}{\partial \beta_i} = \frac{\partial \mathcal{O}^{(i)}}{\partial \beta_j} \, , \tag{41}$$

that, in the standard thermodynamic, correspond to the Maxwell relations. These equations imply the following relationships between the potentials $S[\mathcal{O}]$ and $\psi[\beta]$:

$$\frac{\partial^2 S[\mathcal{O}]}{\partial \mathcal{O}^{(i)} \, \partial \mathcal{O}^{(j)}} = \left(\frac{\partial^2 \psi[\beta]}{\partial \beta_i \, \partial \beta_j} \right)^{-1} \, . \tag{42}$$

In the framework of information geometry the two Hessian matrices appearing in Eq.(42) define, respectively, the metric tensor

$$g^{ij} = \frac{\partial^2 S[\mathcal{O}]}{\partial \mathcal{O}^{(i)} \, \partial \mathcal{O}^{(j)}} \ , \tag{43}$$

and its inverse

$$g_{ij} = \frac{\partial^2 \psi[\beta]}{\partial \beta_i \, \beta_j} \ , \tag{44}$$

of the corresponding statistical manifold [16] that, according to Eq. (42), is equivalents to

$$g^{ik} \, g_{kj} = \delta^i_j \ . \tag{45}$$

Proceeding further, we consider again the entropy (10) in the probability space and looking at the following relation

$$\frac{\partial S[p]}{\partial p_{\{i\}}} = \ell_{\{i\}}[p] \ , \tag{46}$$

where

$$\ell_{\{i\}}[p] = F_1(S) \, h\big(p_{\{i\}}\big) \ . \tag{47}$$

According to Eq. (46), $\big(p_{\{i\}}, \ell_{\{i\}}\big)$ forms a set of canonically conjugate variables in the statistical manifold, like as like, the set $(\mathcal{O}^{(j)}, \beta_j)$ does in the Legendre manifold. This suggest us to introduce a Legendre transform of the entropy in the statistical manifold [16], to obtain a new functional $\mathcal{S}[\ell]$

$$\mathcal{S}[\ell] = S[p] - \sum_{\{i\}} p_{\{i\}} \, \ell_{\{i\}}[p] \ . \tag{48}$$

However, taking into account that $\ell_{\{i\}}$ actually depend on $p_{\{i\}}$, we can pull-back the functional $\mathcal{S}[\ell]$ onto the probability space and define the new quantity $\mathcal{I}^*[p]$, where $\mathcal{I}^*[p] \equiv \mathcal{I}\big[\beta[p]\big]$, as

$$\begin{aligned} \mathcal{I}^*[p] &\equiv \mathcal{S}[(\ell[p])] \\ &= S[p] - F_1(S[p]) \, \langle h(p_{\{i\}}) \rangle_{\mathrm{lin}} \\ &= S[p] - \beta_0 - \sum_j \beta_j \, F_2\left(\mathcal{O}^{(j)}\right) \langle \theta^{(j)} \rangle_{\mathrm{lin}} \ . \end{aligned} \tag{49}$$

By using Eq. (38), we introduce the pull-back of the ψ-potential $\psi^*[p] \equiv \psi\big[\beta[p]\big]$, according to

$$\psi^*[p] = \mathcal{I}^*[p] + \beta_0 + \sum_j \beta_j \, F_2\left(\mathcal{O}^{(j)}\right) \mathcal{D}\left(\theta^{(j)}\right) \ , \tag{50}$$

10

where function $\mathcal{D}(x)$ is defined in

$$\mathcal{D}(x) = \langle x \rangle_{\text{lin}} - \frac{\langle x \rangle_{\text{KN}}}{F_2\left(\langle x \rangle_{\text{KN}}\right)} \ , \tag{51}$$

and measures the discrepancy between the KN average and the standard linear average. In fact, it is easy to see as $\mathcal{D}(x) = 0$ when $f_2(x) = x$, that is the natural setup for trace-form entropies. In this case, the expression of the ψ-potential simplifies in

$$\psi^*[p] = \mathcal{I}^*[p] + \beta_0 \ . \tag{52}$$

The same relation holds also for the associate entropy $\Sigma[p]$ introduced in Eq. (14), where the $\mathcal{I}$-function is given by

$$\mathcal{I}^*[p] = \Sigma[p] - \sum_{\{i\}} p_{\{i\}} \, h\left(p_{\{i\}}\right) \ , \tag{53}$$

since, in this case, $\ell\left(p_{\{i\}}\right) \equiv h\left(p_{\{i\}}\right)$.

It could be of a certain interest to observe as the function $\mathcal{D}(x)$ can be introduced in

$$\mathcal{D}(x) = f_2(x) - x \, \frac{df_2(x)}{dx} \ , \tag{54}$$

that is reminiscent of a Legendre transform of the Kolmogorov function $f_2(x)$.

4. Final summary

In this letter, we proved in the formulation of a thermostatistic theory the robustness of the structure of Legendre transform against the replacement of the standard linear average with the Kolmogorov-Nagumo nonlinear average. This result can be summarized up as it follows

$$S[\mathcal{O}] - \psi[\beta] + \sum_j \beta_j \, \mathcal{O}^{(j)} = 0$$

$$dS[\mathcal{O}] - \sum_j \beta_j \, d\mathcal{O}^{(j)} = 0 \qquad \Rightarrow \qquad \frac{\partial S[\mathcal{O}]}{\partial \mathcal{O}^{(j)}} = \beta_j$$

$$d\psi[\beta] - \sum_j \mathcal{O}^{(j)} \, d\beta_j = 0 \qquad \Rightarrow \qquad \frac{\partial \psi[\beta]}{\partial \beta_j} = -\mathcal{O}^{(j)}$$

$$\left. \begin{array}{l} \frac{\partial^2 S[\mathcal{O}]}{\partial \mathcal{O}^{(i)} \, \partial \mathcal{O}^{(j)}} = \frac{\partial \beta_i}{\partial \mathcal{O}^{(j)}} = \frac{\partial \beta_j}{\partial \mathcal{O}^{(i)}} \\[2ex] \frac{\partial^2 \psi[\beta]}{\partial \beta_i \, \partial \beta_j} = \frac{\partial \mathcal{O}^{(j)}}{\partial \beta_i} = \frac{\partial \mathcal{O}^{(i)}}{\partial \beta_j} \end{array} \right\} \quad \Rightarrow \quad \frac{\partial^2 S[\mathcal{O}]}{\partial \mathcal{O}^{(i)} \, \partial \mathcal{O}^{(j)}} = \left(\frac{\partial^2 \psi[\beta]}{\partial \beta_i \, \partial \beta_j} \right)^{-1}$$

It is shown, one more time, that the structure of Legendre transform of thermodynamics is strictly related to the maximal entropy principle used to derive the equilibrium distribution, confirming and extending in this way the discussion advanced in [12, 13]. An important consequence of our result concerns the Gibbs equation of thermodynamics that, as showed, still holds also in the present formalism.

11

Appendix A.

In this appendix we specify the results obtained in the paper by considering, as an explicit example, the kernel-like entropic form belonging to the SM family (5). It can be derived from the KN average according to

$$S_{\alpha\gamma}^{\mathrm{SM}}[p] = \varphi_{\alpha\gamma}^{-1}\left(\sum_i p_i\,\varphi_{\alpha\gamma}\left(I_\gamma(p_i)\right)\right) \equiv \frac{\left(\sum_{\{i\}} p_{\{i\}}^\alpha\right)^{\frac{1-\gamma}{1-\alpha}} - 1}{1-\gamma}\ . \tag{A.1}$$

where the Kolmogorov function $f_1(x) = f_2(x) = \varphi_{\alpha\gamma}(x)$ and the elementary information gain $\tilde{I}(p) \equiv I_\gamma(p)$ are given, respectively, by

$$\varphi_{\alpha\gamma}(x) = \ln_\alpha\left(\exp_\gamma(x)\right)\ , \tag{A.2}$$

and

$$I_\gamma(p_i) = \ln_\gamma(1/p_i)\ , \tag{A.3}$$

where $\ln_\alpha(x)$ and $\exp_\alpha(x)$, with $\ln_\alpha(\exp_\alpha(x)) = \exp_\alpha(\ln_\alpha(x)) = x$, are the deformed logarithm and exponential [17], defined in

$$\ln_\alpha(x) = \frac{x^{1-\alpha} - 1}{1 - \alpha}\ , \tag{A.4}$$

$$\exp_\alpha(x) = \left(1 + (1-\alpha)\,x\right)^{\frac{1}{1-\alpha}}\ . \tag{A.5}$$

The above two definitions reduce to standard logarithm and exponential in the $\alpha \to 1$ limit, i.e $\ln_1(x) = \ln(x)$ and $\exp_1(x) = \exp(x)$.

For simplicity, we consider a canonical system with $M = 1$, where the only constraint, in addition to the normalization of p_i, is given by the energy average $U = \langle\epsilon\rangle_{\mathrm{KN}}$.

From the variational problem

$$\delta\left[\frac{1}{1-\gamma}\left(1 + \sum_i p_i^\alpha - \sum_i p_i\right)^{\frac{1-\gamma}{1-\alpha}}\right.$$

$$\left. - \ \beta_0 \sum_i p_i - \frac{\beta_1}{1-\gamma}\left(1 + (1-\alpha)\sum_i p_i\,\theta_i\right)^{\frac{1-\gamma}{1-\alpha}}\right] = 0\ , \tag{A.6}$$

with $\theta_i = \ln_\alpha\left(\exp_\gamma(\epsilon_i)\right)$, we obtain

$$\frac{1}{1-\alpha}\left(\sum_i p_i^\alpha\right)^{\frac{\alpha-\gamma}{1-\alpha}}\left(\alpha\,p_i^{\alpha-1} - 1\right) - \beta_0 - \beta_1\left(\exp_\alpha\left(\langle\theta\rangle_{\mathrm{lin}}\right)\right)^{\alpha-\gamma}\theta_i = 0\ , \tag{A.7}$$

that corresponds to Eq. (18), with $h(x) = (\alpha\,x^{\alpha-1} - 1)/(1-\alpha)$ and $F_1(x) = F_2(x) \equiv \left(\exp_\gamma(x)\right)^{\alpha-\gamma}$. By solving Eq. (A.7) for p_i, we get the distribution in the form

$$p_i = c_\alpha\left[1 - (\alpha - 1)\left(\tilde{\beta}_0 + \tilde{\beta}_1\,\theta_i\right)\right]^{\frac{1}{\alpha-1}}\ , \tag{A.8}$$

where $c_\alpha = \alpha^{\frac{1}{1-\alpha}}$ is a constant.

Unless $\gamma = 1$, distribution (A.8) has an asymptotic power-law behavior for $\theta_i \gg 1$ (i.e. $\epsilon_i \gg 1$). Otherwise, for $\gamma = 1$, corresponding to the Rényi entropy

$$S_\alpha^{\mathrm{R}}[p] \equiv S_{\alpha\gamma=1}^{\mathrm{SM}}[p] = \frac{1}{1-\alpha} \ln \sum_i p_i^\alpha \ , \tag{A.9}$$

the equilibrium distribution has an exponential tail, like the Gibbs one ($\alpha = \gamma = 1$). This fact, already observed in [11], is a consequence of the KN average used in the definition of the energy average. Differently, if we use standard linear average, $f_2(x) = x$, the equilibrium distribution (A.8) has always an asymptotic power-law tail whenever $\alpha \neq 1$ and in particular, for the Rényi entropy, it becomes

$$p_i = c_\alpha \left[1 - (\alpha - 1) \left(\exp\left(S_\alpha\right) \right)^{1-\alpha} (\beta_0 + \beta_1\,\epsilon_i) \right]^{\frac{1}{\alpha-1}} \ , \tag{A.10}$$

with a power-law tail [18].

The associated entropy to $S_{\alpha\gamma}^{\mathrm{SM}}$:

$$\Sigma_\alpha[p] = \varphi_{\alpha\gamma}\left(S_{\alpha\gamma}^{\mathrm{SM}}[p]\right) \ , \tag{A.11}$$

coincides with the Tsallis entropy

$$\Sigma_\alpha[p] = \frac{\sum_i p_i^\alpha - 1}{1 - \alpha} \ . \tag{A.12}$$

This transmutation relationship is already known in the Rényi entropy formalism and actually holds for the whole SM family. As a consequence, both Tsallis and SM entropies share the same equilibrium distribution given in Eq. (A.8) but with an appropriate redefinition of its arguments $(\tilde\beta_0 + \tilde\beta_1\,\theta_i) \to (\beta_0 + \beta_1\,\epsilon_i)$.

Starting from Eq. (A.1) we can derive the expression of the $\mathcal{I}$-function

$$\mathcal{I}_{\alpha\gamma}^{*\,\mathrm{SM}}[p] = \frac{1}{1-\alpha} \left[\left(\frac{1-\alpha}{1-\gamma} - \alpha \right) z_\alpha[p]^{1-\gamma} + z_\alpha[p]^{\alpha-\gamma} \right] - \frac{1}{1-\gamma} \ , \tag{A.13}$$

with $\mathcal{I}_{\alpha\gamma}^{*\,\mathrm{SM}} \to \mathcal{I}_{00}^{*\,\mathrm{SM}} \equiv \mathcal{I}^{*\,\mathrm{BG}} = 1$, where $z_\alpha[p]$ is the partition function of the system [4], given by $z_\alpha[p]^{1-\alpha} = \sum_{\{i\}} p_{\{i\}}^\alpha$ and the ψ-potential follows directly from Eq. (50).

In addition, from Eq. (38), we can obtain the fundamental thermodynamic equation in the entropy representation [19], that links the entropy to the thermodynamic variables

$$\left(\exp_\gamma\left(S_{\alpha\gamma}^{\mathrm{SM}}\right) \right)^{\alpha-\gamma} \left(\alpha\,\varphi_{\alpha\gamma}\left(S_{\alpha\gamma}^{\mathrm{SM}}\right) - 1 \right) = \beta_0 + \sum_j \beta_j\,\overline{\mathcal{O}^{(j)}} \ , \tag{A.14}$$

where, for sake of notation, we defined $\overline{\mathcal{O}^{(j)}} = F_2(\mathcal{O}^{(j)})\,f_2(\mathcal{O}^{(j)})$.

In the $(\alpha, \gamma) \to (1, 1)$ limit, Eq. (A.14) reduces to the fundamental equation of

the Boltzmann-Gibbs theory: $S^{\mathrm{BG}} = 1 + \beta_0 + \sum_j \beta_j \, \mathcal{O}^{(j)}$.

In particular, for the Rényi entropy ($\gamma = 1$), the fundamental equation (A.14) becomes

$$S_\alpha = \frac{1}{\alpha - 1} \, \ln \left(\alpha + (\alpha - 1) \left(\beta_0 + \sum_j \beta_j \overline{\mathcal{O}^{(j)}} \right) \right) , \qquad (A.15)$$

while the $\mathcal{I}$-function is given by

$$\mathcal{I}_\alpha = \log z_\alpha - \frac{z_\alpha^{\alpha - 1} - \alpha}{\alpha - 1} \; . \qquad (A.16)$$

In contrast, for the associated entropic form, the $\mathcal{I}$-function is related to the entropy by an affine relation $\mathcal{I}_\alpha^* = 1 + (1 - \alpha) \, \Sigma_\alpha$, where $\Sigma_\alpha = \varphi_{\alpha\gamma} \left(S_{\alpha\gamma}^{\mathrm{SM}} \right)$ and the fundamental equation becomes

$$\alpha \, \Sigma_\alpha = 1 + \beta_0 + \sum_j \beta_j \, \mathcal{O}^{(j)} \; . \qquad (A.17)$$

A simple expression for $\mathcal{I}^*[p]$ is obtained also when $\Sigma[p]$ belongs to STM family given in Eq. (2). In this case, posing $\Sigma \equiv S_{\alpha\gamma}^{\mathrm{STM}}$, the fundamental equation becomes [20]

$$\lambda \, S_{\alpha\gamma}^{\mathrm{STM}} \left[p^{\mathrm{scaled}} \right] = \beta_0 + \sum_j \beta_j \, \mathcal{O}^{(j)} \; , \qquad (A.18)$$

where the scaled distribution is defined in $p^{\mathrm{scaled}} = \{ p_{\{i\}}/\delta \}$ and the constants λ and δ are given by $\lambda^{\beta - \alpha} = (1 + \alpha)^{1 + \beta}/(1 + \beta)^{1 + \alpha}$ and $\delta^{\beta - \alpha} = (1 + \alpha)/(1 + \beta)$, respectively. The scaled entropy fulfills the relation

$$\lambda \, S_{\alpha\gamma}^{\mathrm{STM}} \left[p^{\mathrm{scaled}} \right] = S_{\alpha\gamma}^{\mathrm{STM}}[p] - \mathcal{I}_{\alpha\gamma}^{\mathrm{STM}}[p] \; , \qquad (A.19)$$

where

$$\mathcal{I}_{\alpha\gamma}^{* \, \mathrm{STM}}[p] = \sum_{\{i\}} \frac{p_{\{i\}}^{1+\alpha} + p_{\{i\}}^{1+\gamma}}{2} \quad \overset{(\alpha,\gamma) \to (0,0)}{\to} \quad 1 \; , \qquad (A.20)$$

so that, Eq. (A.18) can be written in [21]

$$S_{\alpha\gamma}^{\mathrm{STM}} = \mathcal{I}_{\alpha\gamma}^{* \, \mathrm{STM}} + \beta_0 + \sum_j \beta_j \, \mathcal{O}^{(j)} \; . \qquad (A.21)$$

Acknowledgments

References

References

[1] B.D. Sharma and I.J. Taneja, Metrica **22**, 205 (1975); D.P. Mittal, Metrika **22**, 35 (1975).

[2] B.D. Sharma and D.P. Mittal, J. Math. Sci. **10**, 28 (1975).

[3] G. Kaniadakis, M. Lissia and A.M. Scarfone, Phys. Rev. E **71**, 046128 (2005); A.M. Scarfone and T. Wada, Phys. Rev. E **72**, 026123 (2005).

[4] A.M. Scarfone, Phys. Lett. A **355**, 404 (2006); E.LK. Lenzi and A.M. Scarfone, Physica A **391**, 2543 (2012).

[5] E.T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).

[6] A.N. Kolmogorov, *Sur la notion de la moyenne*, Atti della R. Accademia Nazionale dei Lincei, **12**, 388 (1930); M. Nagumo, *Über eine klasse von mittelwerte*, Jpn. J. Math. **7**, 71 (1930).

[7] B. de Finetti, Giornale dell'Istituto Italiano degli Attuari **2**, 369 (1931).

[8] A. Rényi, *Probability Theory*, (North Holland, Amsterdam 1970).

[9] S. Eguchi and O. Komori, *Geometry Science of Information*, Lecture Notes in Computer Science **9389**, 615 (2015).

[10] M. Czachor and J. Naudts, Phys. Lett. A **298**, 369 (2002).

[11] G.B. Bagci and U. Tirnakli, Phys. Lett. A **373**, 3230 (2009).

[12] A. Plastino and A.R. Plastino, Phys. Lett. A **226**, 257 (1997).

[13] T. Yamano, Eur. Phys. J. B **18**, 103 (2000).

[14] R. Hermann, *Geometry, physics and systems* (Marcel Dekker, New York 1973).

[15] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, (Springer-Verlag, 1989).

[16] T. Wada, H. Matsuzoe and A.M. Scarfone, Entropy **17**, 7213 (2015).

[17] C. Tsallis *Introduction to Nonextensive Statistical Mechanics* (Springer, Berlin 2009).

[18] A.G. Bashkirov, Phys. Rev. Lett. **93**, 130601 (2004).

[19] H.B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (John Wiley & Sons, New York 1985).

[20] A.M. Scarfone and T. Wada, J. Phys. A: Math. Theor. **47**, 275002 (2014).

[21] A.M. Scarfone and T. Wada, Prog. Theor. Phys. Suppl. **162**, 45 (2006).