**Title**
Fisher info and thermodynamics' first law

**Permalink**
https://escholarship.org/uc/item/76n5f7zq

**Journal**
Physica A-Statistical Mechanics and its Applications, 369(2)

**ISSN**
0378-4371

**Authors**
Plastino, A
Plastino, A R
Soffer, Bernard H

**Publication Date**
2006-09-01

Peer reviewed

# Fisher information and thermodynamics' 1st. law

A. Plastino[1], A. R. Plastino[2], and B. H. Soffer[3]

[1] Exact Sciences' Faculty-Universidad Nacional de La Plata and National Research Council-CONICET
C.C. 727, 1900 La Plata, Argentina

[2] Department of Physics, University of Pretoria, Pretoria 0002, South Africa

[3] University of California at Los Angeles, Department of Electrical Engineering
Postal address: 665 Bienveneda Avenue, Pacific Palisades, California 90272

It was recently shown, starting from first principles, that thermodynamics' first law (TFL) can be microscopically obtained without need of invoking the adiabatic theorem (AT) [Physica A, **356**, 167 (2005)]. We show here that a TFL can also be found for Fisher's information measure, following a similar procedure. Further, it is proved that enforcing the Fisher-first law requirements in a process in which the probability distribution is infinitesimally varied is equivalent to minimizing Fisher's information measure subject to appropriate constraints.

**Pacs:** 05.30.-d, 05.30.Jp

## I. INTRODUCTION

The study and revision of some of the principles of statistical mechanics is still an active field (see, for instance, among many valuable works, [1–7]). In particular, it was shown in [8] that the entire Legendre-transform (LT) of thermodynamics can be microscopically reproduced using Fisher's information measure in place of Boltzmenn's entropy. This abstract LT constitues an essential ingredient that permits one to build up a "statistical mechanics". Fisher information $I$ allows for such a construction. The desired concavity property of $I$, demonstrated in [8], further illustrates on its utility as a "generator" of an underlying microscopic explanation for thermodynamics.

In another, totally disctint vein, it was also recently shown, starting from first principles and using the point of view of information theory, that thermodynamics' first law (TFL) can be microscopically obtained without

- the need of invoking the adiabatic theorem (AT) with

- its awkward dependence on hypothetical external, *infinitely slowly* varying parameters, and also,

- without needing some other usually invoked constraints and arguments [9].

We will show here that a "TFL" can also be found for Fisher's information measure rather than Shannon's, following a similar procedure. As a corollary result we demonstrate that, *given this Fisher TFL*, using a process in which a sought after unknown probability distribution $f$ is infinitesimally varied is equivalent to seeking

$f$ by minimizing Fisher's information measure subject to appropriate constraints .

## II. PRELIMINARIES

One of the most important extant information measures was advanced by R.A. Fisher in the twenties –for details and discussions we refer to (for instance) [10–12]–. Fisher information (FI) arises as a measure of the expected error in the measurement of a parameter $\theta$ in a situation governed by a family of probability densities $\rho(x,\theta)$, where $x$ is, of course, a random variable, and $\theta$ is the parameter characterizing the alluded mono-parameter family of probability densities [10]. It reads

$$I_{General} = \int dx\, \rho(x,\theta) \left[\frac{\partial \ln \rho}{\partial \theta}\right]^2 . \tag{1}$$

A particular FI-case of great importance [10] is that of translation families [10,13], that is, distribution functions of the form

$$\rho(x,\theta) = \rho(x-\theta). \tag{2}$$

In this particular case $\theta$ is a shift parameter. In other words, two probability densities with different values of $\theta$ have the same shape, but are displaced with respect to each other. It is clear that

$$\frac{\partial \rho}{\partial \theta} = -\frac{\partial \rho}{\partial x}, \tag{3}$$

and, consequently,

$$\int dx\, \rho(x,\theta) \left[\frac{\partial \rho}{\partial \theta}\right]^2 = \int dx\, \rho(x,\theta) \left[\frac{\partial \rho}{\partial x}\right]^2 . \tag{4}$$

From now on, and in order to simplify the notation, the $\theta-$dependence will not be explicitated. Thus, the classical Fisher information associated with translations of a, for simplicity's sake, one-dimensional observable $x$ with probability distribution $f(x)$ becomes [14]

$$I = \int dx\, f(x) \left[\frac{\partial \ln f(x)}{\partial x}\right]^2 , \tag{5}$$

and the associated Cramer–Rao inequality [10,14] satisfies the Cramer-Rao relation

$$\Delta x \geq I^{-1} \tag{6}$$

where $\Delta x$ is the variance of the stochastic variable $x$ [14]

$$\Delta x^2 = \int dx\, \rho(x)\, x^2 - \left( \int dx\, \rho(x)\, x \right)^2, \tag{7}$$

and represents the mean-square error associated to the above referred to measurement. Notice that, unlike the case for Shannon's measure, $I$ is NOT a function of $f(x)$ alone, but also of its derivatives.

We will derive a relationship involving the internal energy, Fisher information and temperature, a generalized pressure and the known a priori input constraints of measurement data. Having constructed this relationship, it will be most simply interpreted as the desired TFL.

### III. A RELATION FOR THE DIFFERENTIAL OF FISHER'S MEASURE

#### A. Notation

We start presenting now our new results. We proceed first of all to derive a relationship for $dI$. For this we need to deal with the internal energy

$$U = \langle \hat{O}_1 \rangle \equiv \langle H(x) \rangle = Tr[fH] \equiv \int dx\, f(x) H(x), \tag{8}$$

where we have used the abbreviation

$$Tr f = \int dx\, f(x), \tag{9}$$

and consider variations $\delta f$ of the distribution function, i.e., a process of the type

$$f(x) \rightarrow f(x) + \delta f(x), \tag{10}$$

where normalization-preservation entails that, during the process,

$$Tr[\delta f] = \delta Tr[f] = 0. \tag{11}$$

We also have, of course, for other physical quantities $O_i(x)$

$$Tr[\delta f(x)\, O_i(x)] = d\langle O_i(x) \rangle. \tag{12}$$

#### B. The Fisher TFL

We are going now to rely heavily on the results of Ref. [8]. It was shown there that, if we are discussing a thermodynamical setting involving $M$ extensive variables $O_i$, we can recover the formal, Legendre-structure of thermodynamics if we minimize the Fisher information measure

$I$ subject to appropriate constraints. These are given by the supposedly a-priori known values

*Constraints in the minimization process are*:

$$\langle O_i(x) \rangle = a_i, \ (i = 1, \dots, M)) \ \text{plus}$$

$$Tr[f] = 1, \ \text{normalization of } f. \tag{13}$$

It turns then out that the extremized measure $I$ is a function only of the extensive variables

$$I = I(\langle O_1(x) \rangle, \langle O_2(x) \rangle, \dots, \langle O_M(x) \rangle), \tag{14}$$

and that, further [8],

$$\frac{\partial I}{\partial \langle O_k(x) \rangle} = \lambda_k, \tag{15}$$

where $\lambda_k$, $(k = 1, \dots, M)$ are $M$ Lagrange multipliers that arise in the course of the minimization procedure. From (14) and (15) it is now clear that

$$dI = \sum_{i=1}^{M} \frac{\partial I}{\partial \langle O_i(x) \rangle}\, d\langle O_i(x) \rangle, \tag{16}$$

i.e.,

$$dI = \sum_{i=1}^{M} \lambda_i\, d\langle O_i(x) \rangle. \tag{17}$$

We explicitly separate now the energy term, and *define* $\beta = \lambda_{i=1} = 1/T$. Here $T$ will be a "Fisher-temperature" (FT), whose existence was conjectured in [10,11] and proved in [8]. Its properties have been extensively discussed in Ref. [15]. The FT behaves as $1/T_{orthodox}$, so that an ordinary thermometer, conveniently recalibrated, can measure it. Accordingly, a term of the form $TdI$ is the Fisher-equivalent of a "heat" increment $dQ_F$ [15] and we are in a position to write

$$dI = \beta dU + \sum_{i=2}^{M} \lambda_i\, d\langle O_i(x) \rangle. \tag{18}$$

Introducing, finally, the "generalized pressures"

$$p_j = T\lambda_j, \ (j = 2, \dots, M), \tag{19}$$

(18) leads us to

$$dU = TdI - \sum_{i=2}^{M} p_i\, d\langle O_i(x) \rangle, \tag{20}$$

which is a "Fisher's thermodynamics" first-law: note that the variations of the extensive quantities are not independent one of another.

If we call the differential of work $(dW)$, effected at temperature $T$, in the form of

$$dW = \sum_{i=2}^{M} p_i \, d\langle O_i(x)\rangle, \qquad (21)$$

and, as stated above, recognize that $T dI$ plays the role of a heat-change $dQ_F$, as expected (but not explicitly discussed there) from the philosophy espoused in [8]. Thus, we have here shown that work is represented by changes in expectation values. These constituted part of our prior knowledge. If a posteriori we encounter changes, this entails that work has been performed, on or by the system.

## IV. TFL AS A GENERATOR OF THE PROBABILITY DISTRIBUTION F

As a corollary to the above results we infer what would be the case if we turn things around and take the TFL (16) as the given, and want to find the unknown probability distribution $f$. By considering a change from $f$ to $f + \delta f$, with $\delta f$ arbitrary, but constrained so that the TFL (16) is obeyed in effecting the probabilities change, we obtain, as a result of that constrained change, the very same $f$ that is obtained by the method of MaxEnt that maximizes Fisher information subject to constraints [8].

We now demonstrate this result formaly, in order to make all the conditions and steps explicit, as a THEOREM:

HYPOTHESIS: If we constrain the process (10)-(11) forcing it to comply with (20),

THESIS: this results in a probability distribution that minimizes $I$ subject to the constraints

$$\langle O_i(x)\rangle = a_i, \quad (i = 1, \ldots, M), \qquad (22)$$

where the $a_i$ are real numbers supposedly known a-priori.

PROOF

The process (10)-(11) generates infinitesimal changes in respectively, $I$, $U$, and the $\langle O_i(x)\rangle$, $(i = 2, \ldots, M)$, and these infinitesimal changes are constrained in the fashion (20). Thus, for the change $\delta f$ prescribed by (10)-(11) we have (as shown in [8]), starting from the $I-$definition (5),

$$dI = \int dx \, K[f] \, \delta f,$$

with a functional $-$ derivative $K = \delta I / \delta f$ of the form

$$K = [\tfrac{d \ln f}{dx}]^2 + 2[\tfrac{d^2 \ln f}{dx^2}], \qquad (23)$$

while the changes in the $\langle O_i(x)\rangle$, $(i = 2, \ldots, M)$ are governed by (12), this is,

$$\delta_f[\langle O_m(x)\rangle] = \delta_f Tr[O_m(x) f(x)] =$$
$$= Tr[O_m(x) \delta f], \quad (\forall \ m). \qquad (24)$$

Of course, $f$ is not yet known. We wish to "extract" $f$ from the 1st Law-relation (20). We introduce now, for future reference, the function $W(x)$ defined as

$$W(x) = \beta H(x) + \sum_{i=2}^{M} \lambda_i O_i(x), \qquad (25)$$

where $\beta$ and the $\lambda$'s are the quantities appearing in (20). We are thus demanding simultaneous fulfillment of (10)-(11), which leads to the two requirements

**(1)** $\int dx \, [K(x) - \beta H(x) - \sum_{i=2}^{M} \lambda_i \langle O_i(x)\rangle] \delta f =$

$$= \int dx \, [K(x) - W(x)] \delta f = 0$$

$$\textbf{(2)} \quad \int dx \, \delta f = 0. \qquad (26)$$

The second one is fullfilled by *any arbitrary* variation $\delta f(x)$ such that any possible area under the pertinent curve above the $x-$axis is compensated by another similar one under that axis. Now, a *special type of variation* among these allow us to obtain an important result from the first relation in (26): consider for the associated integral a $\delta f(x)$ that consists of two equal rectangles of vanishing widths and height $C(x)$, centered at points $x_1$ and $x_2$, respectively, such that one of them points upwards and the other downwards (and thus the contributions of their areas is zero). Since the two points $x_1$, $x_2$ may be located anywhere along the $x-$axis, we are forced to conclude that the rectangle's height, given by

$$C(x) = [K(x) - W(x)] = constant = -\alpha, \qquad (27)$$

which is identical to Eq. (13) of [8]. It is therein shown that, (i) setting the probability distribution $f$ equal to the square of some amplitude $\psi$

$$f(x) = \psi(x)^2, \qquad (28)$$

(ii) defining the new variable

$$v(x) = \frac{\partial \ln \psi(x)}{\partial x}, \qquad (29)$$

(iii) one obtains a Ricatti equation from which one immediately gets for the amplitude a Scroedinger-like equation [8]

$$-\frac{\psi''(x)}{2} - \frac{1}{8} \sum_{i}^{M} \lambda_i O_i(x) \psi(x) = \alpha \psi(x). \qquad (30)$$

Now, (30) yields simultaneously:

- the probability distribution that minimizes $I$ subject to the constraints posed by (i) normalization, with Lagrange multiplier $\alpha$ and (ii) (22), with Lagrange multipliers $\lambda_i$, as shown in [8]

- as demonstrated here, the probability distribution that guarantees that the process (10)-(11) is effected in such a manner that the first law requirements (20) are satisfied, with $\alpha$ equal to minus the rectangle's height in (27) and the $\lambda_i$'s equal to the ratio of the generalized pressures to $\lambda_1 \equiv \beta$, i.e., Eq. (19).

The theorem is proved.

## V. DISCUSSION AND CONCLUSIONS

Two new results have been advanced here:

- We have shown that, within Fisher information theory context, one can derive thermodynamics first law for the Fisher Thermodynamics without appeal to the adiabatic theorem, or to any explicit dependence on hypothetical external parameters. Thus, we avoid the need to add, to the theoretical description, putative infinitely slowly varying external parameters so as to obtain the first law. We find work is represented by changes in expectation values. These constituted part of our prior knowledge. If a posteriori we encounter changes, this entails that work has been performed, on or by the system.

- As a corollary we demonstrated that, given the Fisher TFL, using a process in which a sought after unknown probability distribution $f$ is infinitesimmally varied is equivelant to seeking $f$ by minimizing Fisher's information measure subject to appropriate constraints. Plastino and Curado have shown that similar results would obtain for a large class of additional entropies, namely, those that depend *only* on $f$ [16], unlike the Fisher measure treated here, that in addition depends on derivatives of $f$ as well. All these entropies with the necessary concavity property, and Legendre-transform structure of thermodynamics, allow the construction of a "statistical mechanics" [16].

[1] B. H. Lavenda, *Statistical Physics* (Wiley, New York, 1991); B. H. Lavenda, *Thermodynamics of Extremes* (Albion, West Sussex, 1995).

[2] C. Beck, Phys. Rev. Lett., **87**, 180601 (2001).

[3] C. Beck and E.G.D. Cohen, Physica A, **344**, 393 (2004).

[4] H. Touchette and C. Beck, Phys. Rev. E, **71**, 016131 (2005).

[5] P. T. Landsberg, *Seeking Ultimates*, (Institute of Physics Publishing, Bristol and Philadelphia, 2001).

[6] H. C. von Baeyer, *Information: The New Language of Science*, (Weidenfeld and Nicolson, London, 2003).

[7] H.S. Leff and A.F. Rex, eds., *Maxwell's Demons 2: Entropy, Classical and Quantum Information, Computing*, (Institute of Physics Publishing, Bristol and Philadelphia, 2003).

[8] B. R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, Phys. Rev. E **60**, 48 (1999).

[9] S. Martinez, A. Plastino, and B. H. Soffer, Physica A, **356**, 167 (2005).

[10] B.R. Frieden, *Physics from Fisher information* (Cambridge University Press, Cambridge, England, 1998); *Science from Fisher information* (Cambridge University Press, Cambridge, England, 2004)

[11] B.R. Frieden, B.H. Soffer, Phys. Rev. E **52**, 2274 (1995).

[12] F. Pennini and A. Plastino, Phys. Rev. E **69**, 057101 (2004).

[13] F. Pennini, A.R. Plastino, and A. Plastino, Physica A **258**, 446 (1998).

[14] Michael J.W Hall, Phys. Rev. A **62**, 012107 (2000).

[15] F. Pennini and A. Plastino, Phys. Rev. E **71**, 047102 (2005).

[16] A. Plastino and E. Curado, Phys. Rev. E (2005, in Press).