Non-equiprobable states and generalized entropies in the framework of BG statistics

Fábio Mendes[*] **and Anníbal Figueiredo**

Univ. de Brasília (UnB), Brazil
E-mail: fabiomendes@unb.br, annibal@fis.unb.br

In order to deal with systems that do not obey Maxwell-Boltzmann statistics, several alternative entropy functions were proposed. We present a framework to deal with these exotic entropies in the context of an orthodox BG statistics. Anomalous distributions follow from the assumption that microscopic states are not equally probable due to the presence of correlations. This article explores the consequences of breaking the hypothesis of equal *a priori* probabilities, where it is shown that the structure of thermodynamics and Gibbs/Jaynes' prescription are fully preserved.

---

[*]Speaker.

1. Introduction

It had become clear, over the last few decades, that some systems exhibit equilibrium configurations which violate Maxwell-Boltzmann (MB) statistics. Workers tried different approaches to tackle these situations, and notably, some alternate entropy functions were employed. The important question of how these novel entropic forms could be justified and interpreted, and if they're necessary at all, is not yet well understood [1].

A wealth of works in the recent literature deal with the generalization of Boltzmann-Gibbs (BG) statistics due to Tsallis [2]. Supporters of this approach claims that multifractal phase space boundaries and an alleged non-extensivity of entropy would justify the departure from standard BG statistics. The concept of a generalized entropy were also purposed, with a different bias, to explain anomalous distributions in self-gravitating stellar systems [3].

The growing list of anomalous systems not properly described by BG statistics stands as a embarrassing issue to the conventional theory. With a variable degree of success, non-standard entropic functions seems to provide a better description of such systems, even if as a mere fitting scheme. This dismaying background could be interpreted as if a radical departure from the usual structure of statistical physics is necessary to accommodate these novel phenomena.

The authors understand that such a radical departure is unnecessary. Our approach improves on conventional methods by assuming that microscopic states are not necessarily equally probable. In a specific context, this also implies the existence of non-standard entropies. The cornerstone concept of statistical physics, which is fully kept here, is the more than 100 years old formula due to Boltzmann,

$$S = k_B \ln W, \tag{1.1}$$

which establishes the connection between statistical physics and thermodynamics. Apart from some straightforward modifications necessary to deal with non-equiprobable states, the authors assume that definition (1.1) is not subject to further revision.

We outline the article as follows. The introductory remarks on section 2 justify our commitment with Boltzmann's definition (1.1). This section also contains some arguably simpler demonstrations of orthodox results which might exhibit some interest. The second part deals with two very distinct technical matters. In the equiprobable case, Gibbs picture of statistical physics is compatible with (1.1) and we show that this still holds for non-equiprobable states. At last, a methodology where non-standard entropies may arise is presented together with proofs for compatibility with the general structure of thermodynamics. The approach is not always consonant with the way generalized entropies are used across the literature, so we believe that the present discussion may shed some light to outstanding matters.

2. Fundamental concepts

Consider a system described by a microscopic state x and a macroscopic specification $F = \mathscr{F}(x)$. The probability of a macroscopic configuration can be calculated as

$$P(F) = \sum_x P(x \cdot F) = \sum_x P(x)P(F|x), \tag{2.1}$$

where the following notation was used: $P(A|B)$ is the probability of A under the assumption that B is certain and the dot product refers to a joint event "A and B", whose probability is calculated as $P(A \cdot B) = P(A)P(B|A) = P(B)P(A|B)$.

Since microscopic states fully determine the macroscopic behaviour, the conditional probability term takes the form $P(F|x) = \delta(\mathscr{F}(x) - F)\,dF$. This solves only part of our problem since the microscopic "*a priori* probability" $P(x)$ must be known in order to compute $P(F)$. In most cases, one would set $P(x)$ equal to a constant for the lack of contrary evidence.

A more physically sounded justification exploits the fact that different observers, provided with the same information, should came to consistent inferences. This is satisfied, in general, if $P(x)$ is invariant under the physical operation that distinguishes each observer. Ordinary (Galilean) motion symmetries are usually sufficient to determine $P(x)$ for a single particle system.

However the existence of many particles establishes an additional degree of freedom for correlations may distort $P(x)$ as particles are accommodated in each state[1]. A better way to represent the same system is the occupation vector $\mathbf{n} = (n_1, n_2, \ldots, n_m)$, which ascribe an occupation n_i to the i-th set of quantum numbers or phase space cell. By setting $P(x) = P(\mathbf{n}(x)) \equiv \gamma(\mathbf{n})$, the probability of $\mathbf{n}$ is,

$$P(\mathbf{n}) = \frac{N!}{n_1! n_2! \ldots n_m!} \gamma(\mathbf{n}), \qquad (2.2)$$

where the factorials represents the multiplicity of this state.

Generally speaking, there is no compelling justification to the absence of correlations in $P(x)$ other than an argument for simplicity. Retaining this fact as an open issue, we rather develop the theory without the explicit assumption of equal *a priori* probabilities. Let us now define entropy in the context of non-equiprobable statistics.

A very general formulation for the laws of macroscopic motion is contained in the transition probabilities between macroscopic states. These probabilities can be calculated by integration over all microscopic states at the initial time. After the manipulation of some terms[2], they are written as

$$P(F_t|F_0) = \sum_x P(F_t \cdot x_0|F_0) = \frac{1}{P(F_0)} \sum_x P(x)P(F_0|x_0)P(F_t|x_0). \qquad (2.3)$$

Let $\hat{\mathscr{U}}_t$ denote the time evolution operator, hence the conditional probabilities are given by $P(F_0|x_0) = \delta(\mathscr{F}(x) - F_0)\,dF_0$ and $P(F_t|x_0) = \delta\left(\mathscr{F}\left(\hat{\mathscr{U}}_t x\right) - F_t\right)\,dF_t$.

Differently from microscopic dynamics, these probabilities are not necessarily symmetric under a time-reversal operation. Let the initial state be I, the final one F and their time-reversed counterparts $\bar{I}$ and $\bar{F}$[3]. The probabilities $P(F_t|I_0)$ and $P(\bar{I}_t|\bar{F}_0)$ are easily comparable as, by the time reversal symmetry of microscopic laws, we have that $P(\bar{I}_t|\bar{F}_0) = P(I_{-t}|F_0)$. Using a similar procedure which lead to (2.3), we get

$$P(F_t|I_0) = \frac{1}{P(I)} \sum_x P(x)\delta\left(\mathscr{F}\left(\hat{\mathscr{U}}_t x\right) - F\right)\delta(\mathscr{F}(x) - I)\,dF\,dI, \qquad (2.4)$$

---

[1]Consider a fermionic system: microscopic configurations where two particles share the same set of quantum numbers are assigned with null probability.

[2]It's necessary to apply Bayes theorem on $P(x_0|F_0)$ and set $P(F_t|x_0 F_0) = P(F_t|x_0)$ on the basis that F_0 is redundant.

[3]Since the direction of all momentum variables is changed upon time reversal, the time-reversed macroscopic state may not coincide with the original one.

$$P(I_{-t}|F_0) = \frac{1}{P(F)} \sum_x P(x)\delta\left(\mathscr{F}\left(\hat{\mathscr{U}}_t x\right) - F\right)\delta\left(\mathscr{F}(x) - I\right) dF\, dI. \tag{2.5}$$

This result is equivalent to Boltzmann's argument for the "probable increase of disorder" in a more mathematically sounded presentation. It can be summarized by the *time arrow equation* derived from (2.4) and (2.5):

$$P(F_t|I_0) = \frac{P(F)}{P(I)} P(\bar{I}_t|\bar{F}_0). \tag{2.6}$$

This simple equation shows that, not only macroscopic transition probabilities are usually time-asymmetric, but they are bounded by the ratio between the probability of the final and initial states. For example, if $P(I) \gg P(F)$, then the transition from I to F is *almost* forbidden. Equation (2.6) is the basis for macroscopic irreversibility.

In order to avoid an explicit computation of the probabilities $P(F_t|I_0)$, we exploit a property ubiquitous to systems with a large number of particles. Besides the states in a neighborhood of I, the probabilities of F usually differs from $P(I)$ by several orders of magnitude. This is equivalent to say that for almost all states, either $P(F) \gg P(I)$ or $P(I) \gg P(F)$ will hold.

By the time arrow equation (2.6), this implies that across any reasonably probable macroscopic trajectory, the final state can't be assigned with a (much) smaller probability than the initial condition. Assuming the existence of a statistically dominant trajectory $F(t)$, this amounts to say that $P(F(t))$ is a non-decreasing function of time. All that said, it is clear that $P(F)$ can be identified with entropy, more specifically with the function

$$S(F) \equiv k_B \ln \bar{P}(F), \tag{2.7}$$

where, for the purposes of this work, $k_B = 1$ and $\bar{P}(F)$ is an non-normalized probability. Equation (2.7) is equivalent to Boltzmann's formula (1.1) if all states are equally probable, in which case $\bar{P}(F)$ reduces to a simple accounting of the number of microscopic states constrained to $\mathscr{F}(x) = F$.

3. Non-equiprobable statistics

3.1 Gibbs picture

Let us turn our attention to the more technical problem of an explicit entropy calculation. For many thermodynamical systems, it is possible to identify the macroscopic state with energy, assuming the specific form $U = \sum_{i=1}^m \varepsilon_i n_i$. Other parameters implicit on the value of energy levels (such as volume), may also be necessary for a complete characterization of this system.

Entropy is the calculated from the logarithm of $P(U)$, even though the computation of this term can be rather tricky. Several methods were designed perform this calculation, Darwin and Fowler's was chosen for the pragmatical reason that it was the first one where the authors envisioned a way to account for non-equiprobability.

It is clear that the conditional $P(U|x)$ in (2.1) can be regarded as a constraint on the summation over microscopic states. This translates to the occupation vector picture as

$$P(U) = \sum_{[\mathbf{n}]_U} \frac{N!}{n_1! n_2! \dots n_m!} \gamma(\mathbf{n}), \tag{3.1}$$

where the notation $[\mathbf{n}]_N$ is used to remind of the condition $U(\mathbf{n}) = U$.

Summations such as (3.1) are difficult to handle due to the presence of a constraint. For the equiprobable case where $\gamma(\mathbf{n})$ is constant, an explicit calculation of (3.1) can be accomplished using the Darwin-Fowler method in the thermodynamical limit [4]. We purpose a modification that makes the problem soluble for general $\gamma(\mathbf{n})$. Lets us define

$$G_U(\eta) \equiv \sum_{[\mathbf{n}]_U} \frac{N!}{n_1! n_2! \ldots n_m!} e^{i\eta_1 n_1 + i\eta_2 n_2 + \ldots + i\eta_m n_m}, \tag{3.2}$$

and note that the operator $\hat{\gamma} \cdot f(\eta) \equiv \gamma(-i\partial_\eta) f(\eta)$ maps the function $G_U(\eta)$ into $P(A)$. The operator $\gamma(-i\partial_\eta)$ is determined from the series expansion for $\gamma(\mathbf{n})$, where it is straightforward to show that

$$P(U) = \hat{\gamma} \cdot G_U(\eta)|^{\eta=0}. \tag{3.3}$$

We define a new function of the complex parameter z as

$$G(z, \eta) \equiv \sum_{[\mathbf{n}]} \frac{N!}{n_1! n_2! \ldots n_m!} z^{U(\mathbf{n})} e^{i\eta_1 n_1 + i\eta_2 n_2 + \ldots + i\eta_m n_m},$$

$$= \left(z^{\varepsilon_1} e^{i\eta_1} + z^{\varepsilon_2} e^{i\eta_2} + \ldots + z^{\varepsilon_m} e^{i\eta_m} \right)^N \equiv \phi(z)^N, \tag{3.4}$$

without the energy constraint in the sum. In order to relate $G(z, \eta)$ with $G_U(\eta)$, it is convenient to assume an energy scale where all numbers ε_i can be approximated, without loss of generality, by a set of integers where the maximum common divisor is one. If this can be achieved, the terms $G_U(\eta)$ for each integer value of U are the coefficients of the Laurent series expansion for $G(z, \eta)$, or equivalently, the residue of $G(z, \eta) z^{-1-U}$.

Given the extensivity of U, it can be expressed by $U \simeq U + 1 = \bar{u}N$. The theorem of residues shows that

$$G_U(\eta) = \frac{1}{2\pi i} \oint_C dz\, e^{Nf(z)}; \qquad f(z) \equiv \ln(\phi(z)) - \bar{u} \ln(z), \tag{3.5}$$

where C is any closed path around $z = 0$. For the sake of simplicity, C is chosen to be the circular path that crosses the point x_0 which minimizes $G(z, \eta)$ at the positive real axis (a saddle point).

An approximate calculation of this integral can be done by the saddle point method due to Laplace. It is applicable if $|f(z)|$ has a known global maximum at the contour of integration and if N is large enough. That x_0 is a global maximum can be easily seen: it is the only point in the integration contour where the terms $e^{i\eta_i} z^{\varepsilon_i} = e^{-\beta\varepsilon_i} e^{i(\varepsilon_i \theta + \eta_i)}$ interfere constructively for $\phi(z)$ as $\eta_i \to 0$. Here we defined $x_0 = e^{-\beta}$ to account for the fact that x_0 is positive. The result of this calculation, up to the first term in a series expansion for N is

$$G_U(\eta) = \phi(x_0)^N e^{\beta U} \sqrt{\frac{1}{2\pi N |f''(x_0)|}}. \tag{3.6}$$

The probability of a state with a given energy is

$$P(U) = \frac{e^{\beta U}}{\sqrt{2\pi N}} \, \hat{\gamma} \cdot e^{N \ln[\phi(x_0)] - \frac{1}{2} \ln|f''(x_0)|} \Big|^{\eta=0},$$

$$\propto \sum_{[\mathbf{n}]} \frac{N!}{n_1! n_2! \ldots n_m!} \gamma(\mathbf{n}) e^{\beta(U - U(\mathbf{n}))} \tag{3.7}$$

where the term $\ln|f''(x_0)|$ was neglected with respect to $N\ln\phi(x_0)$.

This method also allows the computation of mean values using the operator $\hat{\chi} \equiv \hat{\gamma} \cdot A\left(-i\partial_\eta\right)$, where $A(\mathbf{n})$ is the parameter of interest. It follows, from the same procedure as (3.7), that

$$\langle A(\mathbf{n})\rangle \propto \sum_{[\mathbf{n}]} A(\mathbf{n})\frac{N!}{n_1!n_2!\ldots n_m!}\gamma(\mathbf{n})e^{\beta(U-U(\mathbf{n}))}. \tag{3.8}$$

Expressions (3.8) and (2.2) makes it possible to identify the occupation vector probability with

$$P(\mathbf{n}|\beta) = P(\mathbf{n})e^{-\beta U(\mathbf{n})-\lambda}, \tag{3.9}$$

where the proper choice of λ normalizes (3.9). Differently to $P(x)$, the conditional probability $P(\mathbf{n}|\beta)$ accounts information regarding energy via the temperature parameter β.

Entropy follows trivially from (3.7). The final result, up to an addictive factor is:

$$S(U) = \beta U + \ln Z(\beta); \qquad Z(\beta) \equiv \sum_{[\mathbf{n}]} P(\mathbf{n})e^{-\beta U(\mathbf{n})}. \tag{3.10}$$

The importance of the previous calculation stems from the fact that it connects Boltzmann's entropy (2.7) with Gibbs picture of statistical physics. A modern account of Gibbs picture is done in the context of information theory and Bayesian probabilities. It is a methodology of statistical inference which, in principle, has nothing to do with thermodynamics.

Inference is any process that updates a "prior" probability $P(x)$ into the "posterior" $P(x|\xi)$ which takes into account an additional datum via a parameter ξ. Of great interest is the case where the mean value $F = \langle f(x)\rangle$ over $P(x|\xi)$ is known. We clearly need a method to determine this probability from the known functions $P(x)$, $f(x)$ and F.

One possible approach is to assume the existence of a functional $H[x|\xi]$ which order different assignments $P(x|\xi)$ in a scale of increasing plausibility. It is obvious that H is meant to be maximized, but what is H? The answer depends on one's own prejudices regarding what makes a good inference. Owing a lot to Shannon's celebrated work on communication theory, Khinchin [6] and Jaynes [5] were probably the responsibles to formalize this problem and translate it into the usual language of probability theory and statistical mechanics. By some plausible desiderata found in [7], the MaxEnt inference functional is uniquely determined as

$$H[x|\xi] = -\sum_x P(x|\xi)\ln\frac{P(x|\xi)}{P(x)}. \tag{3.11}$$

The maximum of $H[x|\xi]$ constrained to $F = \langle f(x)\rangle$ gives the best inference on $P(x|\xi)$ as

$$P(x|\xi) = \frac{P(x)}{Z(\xi)}e^{-\xi f(x)}; \qquad Z(\xi) = \sum_x P(x)e^{-\xi f(x)}, \tag{3.12}$$

where $H[x|\xi]$ assumes the value

$$H_{\max} = \xi F + \ln Z(\xi). \tag{3.13}$$

Strictly speaking, from the point of view of what is pertinent to statistical physics, Gibbs method can only claim a partial success. It succeeds in what probabilities $P(x|\xi)$ were obtained, but entropy was not. However the remarkable similarities between (3.9) and (3.10) with (3.13)

and (3.12) means that $H_{\max}$ can also be used as an approximation to the physical entropy given by $S(F) = k_B \ln \bar{P}(F)$. This work extends this equivalence to the non-equiprobable case. It must be stressed that a demonstration of this result for more general class of energy functions with interaction terms is still lacking [8].

If a different functional than (3.11) were employed by alternative desiderata, it doesn't seems likely that its maximum value would lead to the same expression found in (3.10). This would make it useless for the purposes of thermodynamics and, in this sense, Gibbs picture is not subject to revision.

3.2 Occupation vector entropies

Some authors recognized that the functional $-\sum_n p_n \ln p_n$ is used in different contexts and with a different operational definitions to Gibbs and Boltzmann [9]. In Gibbs picture, it arguably has an information theoretical interpretation and is employed in phase space ($\sim 6 \times 10^{23}$ coordinates) and for Boltzmann, it appears in the context of a normalized occupation vector (~ 6 coordinates). Non-standard entropic forms may arise only in the second context.

If a coarse graining procedure is assumed, the occupation vector can be regarded as a macroscopically accessible variable. Thus entropy of this *macroscopic* configuration is computed with (3.1) and (2.7), leading to

$$S(\mathbf{n}) = -\sum_i n_i \ln \frac{n_i}{N} + \ln \gamma(\mathbf{n}), \tag{3.14}$$

where Stirling approximation for $\ln n!$ was employed. This distinguishes from the previous approach where energy, rather than $\mathbf{n}$, was identified with the macroscopic state.

The occupation vector may change with time since it is not an integral of motion. Without knowing the time evolution equation for $\mathbf{n}$, it is necessary to resort to some kind of plausible reasoning. In the absence of further information, the best we can say is that, with the greatest probability, the actual occupation vector is the one with maximum entropy. Additional information such as the value of known invariants of motion are contemplated as constraints on entropy maximization. The optimal inference for $\mathbf{n}$ is given by the prescription:

$$\max \left(S(\mathbf{n}) - \beta U(\mathbf{n}) - \ldots - \lambda N \right) \Rightarrow \mathbf{n}(\beta, \lambda, \ldots), \tag{3.15}$$

where Lagrange multipliers β, λ, etc are necessary to adjust the resulting occupation vector $\mathbf{n}(\beta, \lambda, \ldots)$ up to the correct value of macroscopic constraints.

We now apply our approach to the simplest of the quantum systems: a gas of non-interacting particles with definite energy levels $\{\varepsilon_i\}$. Given the absence of interactions, energy is written as $U(\mathbf{n}) = \sum_i n_i \varepsilon_i$. We also assume that entropy have the form

$$S(\mathbf{n}) = \sum_i f(n_i), \tag{3.16}$$

which is equivalent to $\gamma(\mathbf{n})$ be factorable into $\gamma(\mathbf{n}) = \gamma(n_1)\gamma(n_2)\ldots\gamma(n_m)$. The maximum condition is

$$\frac{\partial f(n_i)}{\partial n_i} \equiv g(n_i) = \beta \varepsilon_i + \lambda, \tag{3.17}$$

where $g(n)$ must be invertible and thus $n_i = g^{-1}(\beta \varepsilon_i + \lambda)$. If the set $\{\varepsilon_i\}$ contains an infinite number of energy levels, then the occupation $n_i(\beta \varepsilon_i + \lambda)$ must vanish as $\varepsilon_i \to \infty$. This property is equivalent to the constraint $g(n \to 0) = \infty$.

The monotonic character of $g(n)$ implies that $f(n)$ has a definite sign for it is second derivative. The only case of interest is $f''(n) < 0$, a concave function, otherwise the extremum (3.17) would correspond to a minimum. This very important requirement is also necessary to guarantee the thermodynamical stability of $S(\mathbf{n})$.

A more general class of entropies which respect these constraints may be obtained from (3.16), is expressed as:

$$S(\mathbf{n}) = F\left(\sum_i \omega_i f\left(\frac{n_i}{\omega_i}\right)\right). \tag{3.18}$$

The function $F(x)$ is an arbitrary monotonically increasing and concave function and $\omega_i \neq 0$ is a probability distribution. The most probable occupation vector is written, for this class of entropies, as

$$n_i(\beta, \lambda) = \omega_i g^{-1}(\beta' \varepsilon_i + \lambda'), \tag{3.19}$$

where β' is given by $\beta' = \beta \left[F'\left(\sum_i \omega_i f\left(\frac{n_i}{\omega_i}\right)\right)\right]^{-1}$. Each function on (3.18) has a distinctive role: $f(n)$ encodes the occupation vector dependency on temperature, energy and chemical potential. The factors ω_i distort this occupation due to a prior probability assignment (they may be related to the degeneracy of the levels) and $F(x)$ controls the thermal response of the system.

We stress that most of the so called generalized entropies (Fermi-Dirac, Bose-Einstein, Tsallis, Rényi, and others) are special cases of (3.18). However the proper justification for each specific choice of $\gamma(\mathbf{n})$ is an entirely different question. Besides the equiprobable BG entropy, quantum entropies for system with fermions and bosons are the only ones that have a widely accepted interpretation [10].

Thermodynamics may provide some information regarding the functional dependency of entropy on $\mathbf{n}$. One may also speculate that compatibility with the laws of thermodynamics may set further constraints on $S(\mathbf{n})$ besides (3.18). This is not the case, however.

Let us open this discussion with a few collected remarks. First, temperature defined as $(\partial S/\partial U)^{-1}$ can be identified, quite generally, with the parameter β^{-1}[4]. Secondly, this suggests the importance of a Legendre transform structure where the Massieu potential,

$$\Phi(\beta) = S(U) - \beta U, \tag{3.20}$$

is of great interest.

Since the Zeroth and Second Laws are already implicit on the maximization procedure, we shall only concern with the remaining First and Third. Heat can be defined as the energy exchange due to irreversible entropy production. This is written as

$$\delta Q = \frac{\partial U}{\partial S} \delta S = T \delta S. \tag{3.21}$$

---

[4]The proof is very simple. Just take the expression at an entropy maximum,

$$\frac{\partial S}{\partial U} = \sum_{i=1}^{m} \frac{\partial S}{\partial n_i} \frac{\partial n_i}{\partial U} = \beta \sum_{i=1}^{m} \frac{\partial U}{\partial n_i} \frac{\partial n_i}{\partial U} + \lambda \sum_{i=1}^{m} \frac{\partial n_i}{\partial U} = \sum_{i=1}^{m} = \beta \frac{\partial U}{\partial U} + \lambda \frac{\partial N}{\partial U} = \beta.$$

Work, on the other hand, is any energy transfer associated to changes on energy levels due to variation of macroscopically controllable parameters (volume, electric field, etc). If both the occupation vector and the energy levels of a system are modified in a physical process, energy variations are given by

$$\delta U(\mathbf{n}, \varepsilon_i, \varepsilon_{ij}, \ldots) = \delta U(\mathbf{n}) + \langle \delta \varepsilon_i \rangle + \langle \delta \varepsilon_{ij} \rangle + \ldots = \delta U(\mathbf{n}) + W. \quad (3.22)$$

Recalling of the Legendre transform structure of (3.20), then

$$\delta S = \beta \, \delta U(\mathbf{n}, \varepsilon_i, \varepsilon_{ij}, \ldots) + U \delta \beta + \delta \Phi(\beta) = \beta \, \delta U(\mathbf{n}) + \beta W, \quad (3.23)$$

where the fact that $\delta \Phi(\beta) = \frac{\partial U}{\partial \beta} \delta \beta = -U \delta \beta$ were used. Putting it all together, the first law of thermodynamics is obtained

$$\delta U(\mathbf{n}) = \delta Q - \delta W. \quad (3.24)$$

The Third Law is also valid for arbitrary entropies of the type (3.18). Of particular interest is the property that, as temperature goes to zero, the system (most probably) evolves to the configuration with minimum energy.

We provide a proof for non-interacting systems of the type considered in section 3.2. After a simple manipulation of equation (3.18) at maximum entropy, then we obtain, at $T = 0$,

$$\frac{\varepsilon_i - \mu}{g(n_i/\omega_i)} = 0. \quad (3.25)$$

If the function $g(n)$ has no divergences in the interval $0 < n \leq N$, then equation (3.25) is only satisfied if $g(n_i/\omega_i) \to \infty$, or equivalently $n_i = 0$, and when $\varepsilon_i = \mu$. Taking normalization into account, this implies that the state $\varepsilon_i = \mu$ is singled out in the process. From the monotonic character of $g^{-1}(\beta \varepsilon_i - \lambda)$ this must be the state with minimum energy.

A different scenario appears if $g(n)$ has a negative divergence at a finite value $\bar{n}$ (this divergence doesn't violate the monotonic aspect of $g(n)$). It is clear that $\bar{n}$ represents the maximum occupation value for the system, a restriction familiar to fermionic statistics. A new solution for equation (3.25) is now possible as $n = \bar{n}$. Once again, the chemical potential plays the role of selecting the occupied states: states with $\varepsilon_i > \mu$ are empty while the others are occupied with $\bar{n}$ particles.

To sum up everything, for both fermionic-like and bosonic-like entropies, the most probable state at $T \to 0$ is the one with minimum energy. This is the essence of the Third Law of thermodynamics. The crucial point made here is that the general structure of thermodynamics is kept in a non-equiprobable statistical physics. Given the similarities between (3.11) and (3.14), the extensions of these demonstrations to Gibbs picture are, of course, trivial.

4. Final remarks

A fundamental aspect of our approach is that, equilibrium or otherwise, entropy of a macroscopic state F is given by $\ln \bar{P}(F)$. In case of a non-interacting system, Gibbs' methods follow as a valuable calculation aids, which may also be valid for a more general class of energy functions. This point of view, of course, is not new.

In line with a strict adherence to Boltzmann's entropy definition, a lot that goes under the name of generalized entropies is here interpreted as a violation of the assumption of equal *a priori* probabilities. Our main original contribution is the generalization of Darwin-Fowler method for non-equiprobable statistics, extending the equivalence between $\ln \bar{P}(F)$ with Gibbs entropy. This specific approach may be of some value in further attempts to settle the equivalence to more general classes of energies, an speculation which is kept to the future. Secondly, a more general proof for the compatibility between statistical physics and thermodynamics were also provided.

Above all that, the authors eager to dispel the fallacious argument that Gibbs' entropy, Boltzmann's entropy and Boltzmann's H function are all the same thing, thus the required generalization of statistical mechanics to tackle non-standard distributions must be a departure from all such important concepts. The failure to perceive the distinct role each one performs is the failure to understand what can be and what cannot be modified in the present theory.

Acknowledgments

References

[1] A. Figueiredo et. al., *Statistical Interpretation of Generalized Entropies*, *Physica A*, **367** (2006) 191-206.

[2] C. Tsallis, *Nonextensive Statistics: Theoretical, Experimental and Computational Evidences and Connections*, *Brazilian Journal of Physics*, **29** (1999) 1-35.

[3] P. H. Chavanis, *Coarse-grained distributions and superstatistics*, *Physica A*, **359** (2006) 177-212.

[4] E. Schrödinger, *Statistical Thermodynamics*, Dover Publications, Inc, 1989.

[5] E. T. Jaynes , *Information Theory and Statistical Mechanics*, *Physical Review*, **106** 4 (1957) 620-630.

[6] A. Y. Khinchin, *Mathematical Foundations of Information Theory*, Dover Publications, Inc, 1957.

[7] A. Caticha, *Maximum entropy and Bayesian data analysis: Entropic prior distributions*, *Physical Review E*, **70** 4 (1957) 046127(12).

[8] W. T. Grandy, *Principle of maximum entropy and irreversible process*, *Physics Reports*, **63** 3 (1980) 175-266.

[9] E. T. Jaynes , *Gibbs vs Boltzmann Entropies*, *American Journal of Physics*, **33** 5 (1965) 391-398.

[10] L. D. Landau, *Statistical Physics*, Butterworth Heinemann, 1980.