ANNALES DE L'I. H. P., SECTION A

HENRYK GZYL

An unified presentation of equilibrium distributions in classical and quantum statistical mechanics

Annales de l'I. H. P., section A, tome 32, nº 2 (1980), p. 175-183
<http://www.numdam.org/item?id=AIHPA_1980__32_2_175_0>

Ann. Inst. Henri Poincaré Section A :
Vol. XXXII, n° 2, 1980, p. 175-183, *Physique théorique.*

An unified presentation of equilibrium distributions in classical and quantum statistical mechanics

by

Henryk GZYL

Escuela de fisica y matematicas, Facultad de Ciencias,
Universidad central de Venezuela, Caracas

ABSTRACT. — We show that in the usual description of equilibrium states in classical and quantum statistical mechanics it suffices to postulate the grandcanonical distributions and that the other two, the canonical and the microcanonical, can be obtained by taking conditional expectations with respect to the number of particles or with respect to the number of particles and the other macroscopical observables.

RÉSUMÉ. — Nous démontrons, que dans la description ordinaire pour les états d'équilibre dans le cadre de la mécanique statistique (classique et quantique) il suffit de postuler les distributions grand canoniques. Les distributions canoniques et microcanoniques, pouvant être obtenues si nous prenons des expectations conditionnels par rapport au nombre de particles, ou par rapport au nombre des particules et les observables macroscopiques restants.

1. INTRODUCTION

In the usual description of statistical systems (see [7], for example) one assumes postulates which tantamount to postulate three different types of « equilibrium distributions ». These are known respectively as micro-canonical, canonical and grand-canonical distributions, both in the case of classical and quantum statistical mechanics. They describe equilibrium states of systems with: fixed energy and number of particles, fixed number of particles and varying energy and varying energy and number of particles.

We are going to show that, both in the classical and quantum cases, it suffices to postulate the grand-canonical distribution and the other two can be obtained when we compute conditional expectations given the number of particles or given both the number of particles and the energy, obtaining respectively the canonical and the microcanonical distributions.

In section 2 we treat the case of classical statistical mechanics and in section 3 we treat the quantum case. Our approach will be general enough to allow for equilibrium distributions to depend on several macroscopical observables.

2. THE CLASSICAL CASE

In this section we shall postulate the equilibrium distribution for a system with a variable number of particles and obtain the canonical and microcanonical distributions by appropriate conditioning.

As *state space* for a system with variable number of particles we take a set of the type

$$E = \bigcup_{n=0}^{\infty} (n, E_n)$$

where $E_0 = \phi$ stands for the state space of a system with no particles and E_n, $n \geq 1$, will be a finite dimensional manifold (usually $E_n = \mathbb{R}^{6n}$) describing the states of a system of n particles.

The *observables* will be functions $f : E \to \mathbb{R}$ obtained from collections $\{ f_n \mid f_n : E_n \to \mathbb{R} \}$. An important observable is the « number of particles » N, which is such that $N|_{E_n} = n$.

If $\{ f_n \} = f$ is an observable, we shall put f_N for the function $f_N : E \to \mathbb{R}$ such that $f_N(x) = f_{N(x)}(x) = f_n(x)$ if $N(x) = n$.

The *states* of our system will be probability measures (whose structure will be specified more below) on $(E, \mathscr{B}(E))$. $\mathscr{B}(E)$ denotes the Borel sets of E. If μ is a state, we shall denote by μ_n the restriction of μ to E_n.

If is an observable, we define the *expected value of f in the state μ* by

$$E_\mu(f) = \int f d\mu = \sum_{n=0}^{\infty} \int_{E_n} f_n d\mu_n$$

For the observable N we have $E_\mu(N) = \sum_{n=0}^{\infty} nP_n$ where $P_n = \mu_n(E_n) = \mu(E_n)$

is the probability of having n particles in the state μ.

Before continuing, let us mention that this type of descriptions of classical mechanics can be found in [6] or [2]. Let us now make the nature of the μ_n more precise. In each E_n we pick up a differential from of degree equal to the dim E_n (a volume element) which we denote by dV_n, and we assume that there exists a finite independent family $F_N^1, \ldots, F_N^k$ of observables such that

$$d\mu_n = a_n \rho(F_n^1, \ldots, F_n^k) dv_n \tag{2.1}$$

where the a_n are constants and $\sum_{n=0}^{\infty} \mu_n(E_n) = 1$, and certainly the ρ_n are positive measurable functions. Now we are going to compute some conditional expectations. The reader should check [4] for all necessary definitions.

Computation of $E_\mu[f \mid N]$: i. e. of the conditional distribution of f given N in state μ. For any bounded measurable g it must hold that

$$E_\mu[fg(N)] = E_\mu[g(N)E_\mu[f \mid N]] = E_\mu(h(N)g(N)) \tag{2.2}$$

where we put $E_\mu[f \mid N] = h(N)$. Let us compute both sides of (2.2).

$$E_\mu[fg(N)] = \sum_{n=0}^{\infty} g(n) \int_{E_n} f d\mu_n$$

$$E_\mu[h(N)g(N)] = \sum_{n=0}^{\infty} h(n)g(n)p_n$$

Since 2.2 holds for any g, it follows that

$$h(n) = E_\mu[f \mid N = n] = \frac{1}{p_n} \int_{E_n} f_n d\mu_n = \frac{\int f_n \rho_n(F_n^1, \ldots, F_n^k) dV_n}{\int \rho_n(F_n^1, \ldots, F_n^k) dV_n} \tag{2.3}$$

Certainly, when $\rho_n(F_n^1, \ldots, F_n^k) = \exp(-\beta H_n)$ we have that

$$E_\mu[f \mid N = n] = \int f_n \exp(-\beta H_n) dV_n \Big/ \int \exp(-\beta H_n) dV_n$$

which corresponds to the canonical distribution. We also see that $E_\mu[f \mid N]$ does not depend on the choice of the a_n.

Computation of $E_\mu[f \mid N, F_N^1, \ldots, F_N^j]$: i. e. of the conditional expectation of f in the state μ given N and $F_N^1, \ldots, F_N^j$, $j < k$. This will result in a distribution that lies « between » the canonical and the microcanonical distributions. The later is obtained when $j = k$.

Let us put $h(\mathrm{N}, \mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j) = \mathrm{E}_\mu[f \mid \mathrm{N}, \mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j]$; then for any function of the type $g_1(\mathrm{N})g_2(\mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)(g_1, g_2$ and bounded and measurable) it is true that

$$(2.4)\quad \mathrm{E}_\mu[fg_1(\mathrm{N})g_2(\mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)] = \mathrm{E}_\mu[h(\mathrm{N}, \mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)g_1(\mathrm{N})g_2(\mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)]$$

Let us now compute each side of (2.4). Obviously

$$\mathrm{E}_\mu[fg_1(\mathrm{N})g_2(\mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)] = \sum_{n=0}^{\infty} g_1(n) \int_{\mathrm{E}_n} f_n g_2(\mathrm{F}_n^1, \ldots, \mathrm{F}_n^j) d\mu_n$$

$$\mathrm{E}_\mu[h(\mathrm{N}, \mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)g_1(\mathrm{N})g_2(\mathrm{F}_{\mathrm{N}}^1, \ldots, \mathrm{F}_{\mathrm{N}}^j)]$$
$$= \sum_n g_1(n) \int_{\mathrm{F}_n} h(n, \mathrm{F}_n^1, \ldots, \mathrm{F}_n^j) g_2(\mathrm{F}_n^1, \ldots, \mathrm{F}_n^j) d\mu_n$$

and since $g_1(n)$ is arbitrary

$$(2.5)\quad \mathrm{E}_{\mu_n}[f_n g_2(\mathrm{F}_n^1, \ldots, \mathrm{F}_n^j)] = \mathrm{E}_{\mu_n}[h_n(n, \mathrm{F}_n^1, \ldots, \mathrm{F}_n^j)g_2(\mathrm{F}_n^1, \ldots, \mathrm{F}_n^j)]$$

which, being valid for any g_2, implies that

$$(2.6)\qquad \mathrm{E}_{\mu_n}[f_n \mid \mathrm{F}_n^1, \ldots, \mathrm{F}_n^j] = h(n, \mathrm{F}_n^1, \ldots, \mathrm{F}_n^j).$$

We shall now make use of the independence of the functions F_n^i (in the sense that their differentials $d\mathrm{F}_n^i$ are independent cotangent vectors) in order to compute the left hand side of (2.6).

We begin by forming the j-form $d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j$ and the $d(n)$-j-form $*(d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j)$. Here $d(n)$ is the dimension of E_n and $*$ is the Hodge star operator (see [*3*]). We then have

$$d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \wedge *(d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j) = |\, d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \,|^2 d\mathrm{V}_n$$

($\|\ \|$ stands for the norm of the form given by the obvious scalar product).

If we define (with an obvious abuse of notation)

$$(2.7)\quad \frac{d\Sigma_{\mathrm{F}_n^1,\ldots,\mathrm{F}_n^j}}{\| d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \|} = \frac{\rho(\mathrm{F}_n^1, \ldots, \mathrm{F}_n^k) * (d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j)}{\| d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \|^2}$$

then

$$d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \wedge \frac{d\Sigma_{\mathrm{F}_n^1,\ldots,\mathrm{F}_n^j}}{\| d\mathrm{F}_n^1 \wedge \ldots \wedge d\mathrm{F}_n^j \|} = d\mu_n$$

and we see that $d\Sigma / \| \,.\, \|$ plays the role of the surface element on each hyper-surface $\{ \mathrm{F}_n^1 = a_1, \ldots, \mathrm{F}_n^j = a_n \}$.

Let us now compute both sides of (2.5).

$$E_{\mu_n}[h(n, F_n^1, \ldots, F_n^j)g_2(F_n^1, \ldots, F_n^j)]$$

$$= \int_{\mathbb{R}^j} h(n, a_1, \ldots, a_j)g_2(a_1, \ldots, a_j)\Omega(a_1, \ldots, a_j)da_1, \ldots, da_j$$

$$E_\mu[f_n g_2(F_n^1, \ldots, F_n^j)]$$

$$= \int_{\mathbb{R}^j} g_2(a_1, \ldots, a_j) \int_{\{F_n^1 = a_1, \ldots, F_n^j = a_j\}} f_n \rho_n(a_1, \ldots, a_j, F_n^{j+1}, \ldots, F_n^k)$$

$$= \int_{\mathbb{R}^j} g_2(a_1, \ldots, a_j)$$

$$\left(\int_{\{F_n^1 = a_1, \ldots, F_n^j = a_j\}} f_n \rho_n(a_1, \ldots, a_j, F_n^{i+1}, \ldots, F_n^k) \frac{d\Sigma_{a_1, \ldots, a_j}}{\| dF_n^1 \wedge \ldots \wedge dF_n^j \|}\right) da_1, \ldots, da_n$$

where

$$\Omega(a_1, \ldots, a_j)$$

$$= \int_{\{F_n^1 = a_1, \ldots, F_n^j = a_j\}} \rho_n(a_1, \ldots, a_j, F_n^{j+1}, \ldots, F_n^k) \frac{d\Sigma_{a_1, \ldots, a_j}}{\| dF_n^1 \wedge \ldots \wedge dF_n^j \|}$$

is the area of the hypersurface $\{ F_n^1 = a_1, \ldots, F_n^j = a_j \}$.

Since (2.5) is valid for any g_2 it follows that

(2.8) $$E_\mu[f \mid N = n, F_n^1 = a_1, \ldots, F_n^j = a_j]$$

$$= h(n, a_1, \ldots, a_j) = \frac{1}{\Omega'(a_1, \ldots, a_j)}$$

$$\int_{\{F_n^1 = a_1, \ldots, F_n^j = a_j\}} f_n \rho_n(a_1, \ldots, a_j, F_n^{j+1}, \ldots, F_n^k) \frac{d\Sigma_{a_1, \ldots, a_j}}{\| dF_n^1 \wedge \ldots \wedge dF_n^j \|}$$

This is a distribution which is « between » the canonical and the microcanonical. When $j = k$ we obtain

$$E_\mu[f \mid N = n, F_n^1 = a_1, \ldots, F_n^k = a_k]$$

$$= \frac{1}{\Omega'(a_1, \ldots, a_k)} \int_{\{F_n^1 = a_1, \ldots, F_n^k = a_n\}} f_n d\Sigma_{a_1, \ldots, a_k} / \| dF_n^1 \wedge \ldots \wedge dF_n^k \|$$

with

$$\Omega'(a_1, \ldots, a_k) = \int_{\{F_n^1 = a_1, \ldots, F_n^j = a_k\}} d\Sigma_{a_1, \ldots, a_k} / \| dF_n^1 \wedge \ldots \wedge dF_n^k \|$$

and again the result does not depend on $\rho_n(F_n^1, \ldots, F_n^k)$ and that we obtain the usual result see [7] for $j = k = 1$ and $F_n^1 = H_n$ is obvious.

3. QUANTUM CASE

As in the previous section, we shall see that it is enough to postulate the equilibrium distribution for a system with variable number of particles, and that in order to obtain the other two distributions it is enough to specify (or rather, to measure) the number of particles and the other macroscopic observables.

Let us briefly review the description of equilibrium states in quantum statistical mechanics. The reader can check in [*2*] for more details. We start with a Hilbert space

$$H = \bigoplus_{n=0}^{\infty} H_n$$

where H_n, $n \geq 1$ describes the « pure » states of a_n n-particle system and H_0 is essentially the complex numbers. Let us recall that the elements in H are of the from $\psi = \psi^0 \oplus \psi' \oplus \ldots \oplus \psi^n \oplus \ldots$ with ψ^n in H_n, and only finitely many of them different from zero, or such that $\sum_n \langle \psi_n \mid \psi_n \rangle < \infty$.

The *observables* of the system are described by operators of the form

$$A = \bigoplus_{n=0}^{\infty} A_n$$

where A_n is a hermitian operator on H_n representing an observable of a system of n-particles.

The usual statistical description makes use of a generalization of the notion of state via the introduction of « density matrices ». We shall consider density matrices of the type

$$\rho = \bigoplus_{n=0}^{\infty} \rho_n$$

where the ρ_n are positive hermitian matrices.

The *expected value* of an observables A in the state ρ is given by

$$E_\rho(A) = \mathrm{Tr}(\rho A) = \sum_0^{\infty} \mathrm{Tr}(\rho_n A_n)_n$$

where $\mathrm{Tr}(\ )_n$ indicates trace relative to H_n.

In the quantum case the number of particles operator will be of the from

$$N = \sum_{n=0}^{\infty} n I_n$$

where I_n is the identity on H_n. Note that N commutes with any observable A and that its eigenstates are of the form $O \oplus \ldots \oplus O \oplus \psi^n \oplus O \oplus \ldots$; ψ^n arbitrary in H_n.

Also P_n, the projector on the n-particle states is of the form

$$P_n = O \oplus \ldots \oplus O \oplus I_n \oplus . O \oplus \ldots$$

In this section we shall assume that the ρ_n are of the form

$$\rho_n = a_n \sigma_n(F_n^1, \ldots, F_n^k)/Z$$

where

$$Z = \sum_{n=0}^{\infty} \mathrm{Tr}(\rho_n)_n .$$

We shall *assume that the* $F_n^1, \ldots, F_n^k$ *are a commuting set of hermitian operators and that their spectra are discrete and that their eigenvectors form a complete set. In order not to have domain problems we shall assume them bounded.*

Now we suppose we are given a ρ as above, and a_n observable A, and we assume we measure N (or N and $F^1, \ldots, F^k$) and we want to compute $E_\rho[A \mid N]$, $E_\rho[A \mid N, F^1, \ldots, F^k]$ the conditional expected values of A in the state ρ given N (or given $N, F^1, \ldots, F^k$).

Computation of $E_\rho[A \mid N]$: By definition (see [*1*] or [*2*]) $E_\rho[A \mid N]$ is the operator in the Von Neumann algebra generated by N satisfying

$$E_\rho[P_n E_\rho[A \mid N] P_n] = E_\rho[P_n A P_n] \tag{3.1}$$

for any P_n. Certainly

$$E_\rho[P_n A P_n] = E_{\rho_n}[A_n] = \mathrm{Tr}(\rho_n A_n)$$
$$E_\rho[P_n E_\rho[A \mid N] P_n] = E_\rho[A \mid N = n] \,\mathrm{Tr}(\rho_n)_n$$

where we wrote $E_\rho[A \mid N]$ as $\sum_{n=0}^{\infty} E_\rho[A \mid N = n] P_n$ which follows from the commutativity of N with any other observable.

Therefore, it follows from (3.1) that

$$\begin{aligned} E_\rho[A \mid N = n] &= \mathrm{Tr}(\rho_n A_n)_n / \mathrm{Tr}(\rho_n)_n \\ &= \mathrm{Tr}(A_n \sigma_n(F_n^1, \ldots, F_n^k)) / \mathrm{Tr}(\sigma_n(F_n^1, \ldots, F_n^k)) \end{aligned} \tag{3.2}$$

which in the particular case in which $F_n^1 = H_n$ and $\sigma_n(H_n) = \exp(-\beta H_n)$, we obtain

$$E_\rho[A \mid N = n] = \mathrm{Tr}(A_n e^{-\beta H_n})_n / \mathrm{Tr}(e^{-\beta H_n}) . \tag{3.3}$$

and we see that after measuring N in the state ρ, the expected value of A coincides with the expected value of A_n according to the canonical distribution.

Computation of $E_\rho[A \mid N, F^1, \ldots, F^k]$: If $f_{i_1}^m, \ldots, f_{i_k}^m$ are the eigenvalues of $F_m^1, \ldots, F_m^k$ and $P_{f_{i_1}^m,\ldots,f_{i_k}^m}$ is the projector on the subspace generated by the corresponding eigenvectors we have $P_n P_{f_{i_1}^m,\ldots,f_{i_k}^m} = 0$ unless $m = n$.

The conditional expectation $E_\rho[A \mid N, F_m^1, \ldots, F_m^k]$ is the element in the Von Neumann algebra generated by $N, F_m^1, \ldots, F_m^k$, such that

$$(3.4)\quad E_\rho[P_n P_{f_{i_1}^n,\ldots,f_{i_k}^n} E_\rho[A \mid N, F^1, \ldots, F^k] P_n P_{f_{i_1}^n,\ldots,f_{i_k}^n}] = E_\rho[P_n P_{f_{i_1}^n,\ldots,f_{i_n}^n} A P_n P_{f_{i_1}^n,\ldots,f_{i_k}^n}]$$

Now, the right hand side of (3.4) equals

$$E_\rho[P_{f_{i_1}^n,\ldots,f_{i_k}^n} A_n P_{f_{i_1}^n,\ldots,f_{i_k}^n}] = \mathrm{Tr}\,[\rho_n P_{f_{i_1}^n,\ldots,f_{i_k}^n} A_n] = a_n(\sigma_n(f_{i_1}^n, \ldots, f_{i_k}^n)\,\mathrm{Tr}\,(A_n P_{f_{i_1}^n,\ldots,f_{i_k}^n})$$

From the assumptions on the set $F_n^1, \ldots, F_n^k$, it follows that

$$E_\rho[A \mid N, F^1, \ldots, F^k] = \sum_{n, f_{i_1}^n, \ldots, f_{i_k}^n} h(n, f_{i_1}^n, \ldots, f_{i_n}^n) P_n P_{f_{i_1}^n,\ldots,f_{i_n}^n} = h(N, F^1, \ldots, F^k)$$

and taking this in to account it follows that the left hand side of (3.4) can be computed as follows

$$E_\rho[P_n P_{f_{i_1}^n,\ldots,f_{i_n}^n} h(N, F^1, \ldots, F^k) P_n P_{f_{i_1}^n,\ldots,f_{i_n}^n}] = h(n, f_{i_1}^n, \ldots, f_{i_k}^n) a_n \sigma_n(f_{i_1}^n, \ldots, f_{i_k}^n)\,\mathrm{Tr}\,(P_{f_{i_1}^n,\ldots,f_{i_k}^n})_n$$

Comparing both sides of (3.4) we obtain

$$(3.5)\quad E_\rho[A \mid N = n, F_n^1 = f_{i_1}^n, \ldots, F_n^k = f_{i_k}^n] = \mathrm{Tr}(A_n P_{f_{i_1}^n,\ldots,f_{i_n}^n})_n / \mathrm{Tr}\,(P_{f_{i_1}^n,\ldots,f_{i_k}^n})_n$$

the meaning of which is clear: $\mathrm{Tr}\,(A_n P_{f_{i_1}^n,\ldots,f_{i_n}^n})_n$ is the expected value of A_n over the set of states corresponding to eigenvalues $f_{i_1}^n, \ldots, f_{i_k}^n$ and $\mathrm{Tr}\,(P_{f_{i_1}^n,\ldots,f_{i_k}^n})_n$ is the degeneracy of that set of states.

With this we prove our contention, namely: that it is enough to postulate the grandcanonical distribution and that the other two can be obtained from it by specifying enough macroscopical observables.

REFERENCES

[*1*] E. B. Davies and J. T. Lewis, An Operational Apprach to Quantum Probability, *Comm. Math. Phys.*, t. 7, 1970, p. 239-260.

[*2*] J. M. Cook, Banach Algebras and Asymptotic Mechanics, Cargese, *Lect. Theor. Phys.*, Gordon Breach, 1967.

[*3*] H. Flanders, *Differential Forms*, Acad. Press, N. Y., 1963.

[*4*] I. I. Gihman and A. V. Skorohod, *The Theory of Stochastic Processes I*, Springer-Verlag, Berlin, 1974.

[5] S. GUDDER and I. P. MARCHAND, Non-communative Probability on Von Neumann Algebras. *Jour. Math. Phys.*, t. **13**, no. 5, 1972, p. 199-306.
[6] D. RUELLE, States of Classical Mechanics. *Jour. Math. Phys.*, t. **8**, 1968, p. 1657-1668.
[7] D. N. ZUBAREV, *Non-equilibrium Statistical Thermodynamics*. Consultants Bureau, N. Y., 1974.

(Manuscrit reçu le 12 juin 1978)