SCHOOL and CONFERENCE
on
COMPLEX SYSTEMS
and
NONEXTENSIVE STATISTICAL MECHANICS

31 July - 8 August 2006

Entropy and Statistical Mechanics:
Fundamental principles

M. NAUENBERG

Department of Physics
University of California
Santa Cruz, USA

Entropy and Statistical Mechanics: Fundamental principles

Michael Nauenberg
Department of Physics
University of California, Santa Cruz, CA 95064

July 31, 2006

I. Elementary principles of mathematical statistics

Let p_i denote the probability for the occurence of states of a system labelled by the integers $i = 1, 2...n$. Then

$$\sum_i^n p_i = 1,$$ (1)

and the average or mean value $\bar{f}$ of any variable f_i of the system is

$$\bar{f} = \sum_i^n p_i f_i.$$ (2)

A simple way to understand these two fundamental relations of statistics is to suppose that we make N independent measurements of the variable f_i, and that the number of times that we obtain the value f_i is n_i. Then

$$N = \sum_1^n n_i,$$ (3)

and the observed average value f of f_i is

$$f = \frac{1}{N} \sum_i^n n_i f_i$$ (4)

Setting the probability p_i equal of the ratio n_i/N in the limit that N becomes arbitrarily large, i.e.

$$p_i \equiv (\frac{n_i}{N})_{limit N \to \infty},$$ (5)

$f \to \bar{f}$, and we recover the fundamental equations of statistics, Eqs.1 and 2.

In addition to these two equations , we will need the joint probability for two systems a and b which are statistically independent of each other. This probability is given by

$$p_{i,j}^{ab} = p_i^a p_j^b$$ (6)

where p_i^a and p_j^b are the probabilities for system a and b respectively.

II. Elementary principles of statistical mechanics

In quantum mechanics the stationary states of a system are characterized by its energy eigenvalues E_i. Hence, any equilibrium state of the system can be characterized by probabilities p_i determined by a function f which depends *only* on E_i, and a parameter β with dimensions of inverse energy. On dimensional grounds [1]

$$p_i = \frac{f(\beta E_i)}{Z(\beta)},$$ (7)

where $Z(\beta)$ is determined by the normalization requirement, Eq.1,

$$Z(\beta) = \sum_i f(\beta E_i).$$ (8)

Evidently, $Z(\beta)$ depends on *all* the energy eigenvalues of the system, and as we shall demonstrate, determines the thermodynamical functions of the system. For systems in equilibrium, we now will give a mathematical proof that the function. $f(x)$ is a *universal* function which is independent of the physical nature of the system under consideration.

Consider two systems a and b that are *weakly* coupled, i.e., the interaction energy can be neglected. This is the case for short range force, but not, for example, for gravitational force. Then their total energy $E_{i,j}^{ab}$ is *additive*,

$$E_{i,j}^{ab} = E_i^a + E_j^b,$$ (9)

---

[1] A mathematical function $f(x)$ and its argument x must be dimensionless variables

and the probability $p_{i,j}^{ab}$ for the state of the coupled system is given by Eq. 6. When the systems a and b are *subsystems* of the same physical system separated by a wall which allows the exchange of energy, and otherwise is under the same physical conditions, e.g. equal pressure, etc., we have $f_a(x) = f_b(x) = f_{ab}(x) = f(x)$. Substituting Eq. 7 in Eq. 6 we obtain the relation

$$\frac{f(\beta^{ab}E_{i,j}^{ab})}{Z^{ab}(\beta^{ab})} = \frac{f(\beta^a E_i^a)}{Z^a(\beta^a)}\frac{f(\beta^b E_i^b)}{Z^b(\beta^b)} \tag{10}$$

where β^a, β^b, and β^{ab} are the respective parameters β. This relation is a nonlinear equation which must be solved for $f(x)$. But this function, if it exists, is *independent* of the physical nature of the system, which is described here by the energy eigenvalues E_i. Hence f(x) must be a *universal* function, and the same equation must hold when a and b are *different* weakly coupled systems, i.e. provided that $E_{i,j}^{ab}$ satisfies the *additivity* condition Eq. 9.

To solve Eq.10, it turns out that it is sufficient to satisfy the equality of the numerators on both sides of this equation, because then the equality of the denominators is automatically satisfied, as can be easily verified. Hence, we require that

$$f(\beta^a(x+y)) = f(\beta^a x)f(\beta^b y), \tag{11}$$

where x and y are independent variables. Setting $y = 0$ gives

$$f(\beta^{ab}x) = f(\beta^a x)f(0), \tag{12}$$

which requires that $f(0) = 1$, and $\beta^{ab} = \beta^a$. Likewise, setting $x = 0$,

$$f(\beta^{ab}y) = f(\beta^b)f(0) \tag{13}$$

again requires that $f(0) = 1$, and $\beta^{ab} = \beta^b$. Hence, a solution of Eq. 11 is possible provided that $\beta^a = \beta^b = \beta^{ab} = \beta$.

Setting $\beta = 1$, we have

$$f(x+y) = f(x)f(y), \tag{14}$$

and taking derivatives with respect to y on both sides of this equation, setting *afterwards* $y = 0$, gives the first order differential equation

$$f'(x) = f'(0)f(x), \tag{15}$$

3

where $f'(x) = df(x)/dx$ [2]. The solution is an exponential function

$$f(x) = Ce^{f'(0)x} \tag{16}$$

where C is a constant which is equal to unity by the condition $f(0) = 1$ The value of $f'(0)$, is arbitrary, but the finitness of the sum defining $Z(\beta)$, Eq. 8, requires that $f'(0) < 0$. For convenience we set $f'(0) = -1$, and obtain for the *universal* form of $f(x)$ the exponential function

$$f(x) = e^{-x}. \tag{17}$$

Substituting this exponential form in Eq. 8 for $Z(\beta)$, we obtain [3]

$$Z(\beta) = \sum_i e^{-\beta E_i}. \tag{18}$$

Of course, $Z(\beta)$ depends on the system under consideration, because it is determined by the energy eigenvalues E_i of this system. It can be readily verified that the condition for equality of the denominators required to satisfy Eq. 10,

$$Z_{ab}(\beta(E^a + E^b)) = Z_a(\beta E^a)Z_b(\beta E^b), \tag{19}$$

is now automatically satisfied. Q.E.D.

It is important to stress the *uniqueness* of the exponential form for the probability p_i,

$$p_i = \frac{e^{-\beta E_i}}{Z(\beta),} \tag{20}$$

as a function of the energy eigenvalues E_i. [4] This exponential form is known

---

[2]The same result would have been obtained by first taking derivates of both sides of Eq. 11 with respect x, and afterwards setting $x = 0$.

[3]In statistical mechanics $Z(\beta)$ is called the partition function,

[4]In numerous papers on q-entropy, it has been *claimed* that a different form for the function $f(x)$ introduced by Tsallis,

$$f_q(x) = (1 - (1 - q)x)^{1/(1-q)}, \tag{21}$$

also satisfies the additivity condition for the energy, Eq. 9, and the multiplicativiy condition for probabilities of statistical independent systems, Eq. 6. It can be readily verified, however, as expected from the *uniqueness* of the solution of Eq.14, rigorously demonstrated here, that this is not the case. In fact

$$f_q(x + y) \neq f_q(x)f_q(y), \tag{22}$$

unless we take the limit $q \to 1$ in Eq. 21, in which case $f_1(x) = e^{-x} = f(x)$.

as the *canonical distribution* and in thermodynamics it applies to systems in thermal equilibrium, i.e in contact with a *thermal* bath with which it can exchange any required amount of energy in the form of heat. This will be demonstrated below. It was first introduced in statistical mechanics by Maxwell, by arguments similar to those presented here [5], to describe the distribution of velocities of a molecular gas in thermal equilibrium; subsequently it was applied in greater generality by Boltzmann [1], and then for all weakly coupled equilibrium systems by Gibbs [2] and by Einstein [3] [6].

III. The entropy in statistical mechanics

Both Gibbs [2] and Einstein[3] also demonstrated that the entropy [7] of a thermodynamic system is *uniquely* determined by the exponential form of the probability distribution, Eq. 20. For this purpose, assume that the energy eigenvalues E_i depend also on an external variable which could be the volume V, an external magnetic field, etc.. In this case, the mean energy E [8] is given by

$$E = -\frac{\partial ln(Z)}{\partial \beta} \tag{23}$$

where the partial derivatives are taken keeping the external variable V fixed. Likewise, if the external variable is the volume V, we obtain the average pressure P by the relation

$$P = -\frac{1}{\beta}\frac{\partial ln(Z)}{\partial V}. \tag{24}$$

Setting

$$F = -\frac{1}{\beta}ln(Z) \tag{25}$$

---

[5]Maxwell was inspired to his solution by the mathematics of gambling

[6]The independent development of the canonical ensemble in statistical mechanics by Einstein is not well known, as evidence by the fact that the canonical distribution is often called the Boltzmann-Gibbs (BG) distribution

[7]The concept of entropy was introduced by Rudolf Clausius in 1856, on purely thermodynamic grounds: the principles of conservation of energy, and Carnot's maximum for the work efficieny of reversible engines

[8]In this section we drop the bar to refer to mean quantities

We have

$$\frac{\partial F}{\partial \beta} = \frac{E - F}{\beta},\qquad(26)$$

which suggests defining a new function S by the relation

$$S = k\beta(E - F) = k\left(-\beta\frac{\partial lnZ}{\partial \beta} + lnZ\right)\qquad(27)$$

where k is a constant. Then

$$\frac{\partial S}{\partial \beta} = k\beta\frac{\partial E}{\partial \beta},\qquad(28)$$

and therefore we can identify

$$T = \frac{1}{k\beta}\qquad(29)$$

with the thermodynamic *absolute temperature* [9], S with the entropy and F with the free energy. Indeed, Eq. 28 implies that

$$\frac{\partial S}{\partial E} = \frac{1}{T},\qquad(30)$$

and we also find that

$$\frac{\partial F}{\partial \beta} = \frac{1}{k\beta^2}S.\qquad(31)$$

Hence, changing from β to T as the independent variable, we obtain the familiar thermodynamic relations

$$F = E - TS,\qquad(32)$$

$$\frac{\partial F}{\partial T} = -S,\qquad(33)$$

$$\frac{\partial F}{\partial V} = P,\qquad(34)$$

and

$$\frac{\partial S}{\partial V} = \frac{\partial P}{\partial T}.\qquad(35)$$

---

[9]The constant k is known as Boltzmann's constant. Originally, however, it was introduced by Planck in his celebrated paper on black-body radiation. It is required because the temperature is given in degree units, while the entropy is in energy units /degree units

According to Eq. 20

$$ln(p_i) = -\beta E_i - ln(Z), \tag{36}$$

and it follows from the definition of the entropy function S given in Eq. 27 that

$$S = k(-\sum_i p_i ln(p_i)). \tag{37}$$

This relation for the entropy was first obtained by Boltzmann [1] who obtained the maximum number of configurations W of a classical system of n molecules having energies $E_j = j\epsilon$, where ϵ is a small energy unit which he set equal to zero in the limit corresponding to classical mechanics[10] . In that case

$$W = \frac{n!}{n_0! n_1! \dots} \tag{38}$$

where n_i are the number of molecules having energy E_i, and

$$n = \sum_i n_i \tag{39}$$

For large values of n_i he applied the Stirling aproximation $ln(n!) = n(ln(n) - 1)$ and obtained the relation

$$ln(W) = -n \sum_i p_i ln(p_i) \tag{40}$$

where $p_i = n_i/n$. For the distribution that leads to the maximum value of W subject to the condition of fixed total energy, he obtained the canonical distribution, Eq. 20, and demonstrated his celebrated relation for the entropy S of the molecules in terms of W,

$$S = kln(W). \tag{41}$$

Boltzmann's variational method gives an alternative derivation for the canonical probability distribution p_i, and is presented in general form in the next section.

---

[10]Later on, Planck took over Boltzmann's method for an ensemble of charge oscillators in equilibrium with electromagnetic radiation in a cavity, but he kept the energy unit ϵ finite setting $\epsilon = h\nu$, where h is a constant now known as Planck's constant. In this way he obtain his famous blackbody formula and ushered in quantum mechanics.

IV. Boltzmann's variational derivation for the canonical probability distribution and entropy

Consider a system with energy levels E_i, and suppose we have a very large number N of replicas such that n_i have energy E_i. Then

$$N = \sum n_i \tag{42}$$

and

$$E = \sum n_i E_i \tag{43}$$

is the total energy. In the limit that $N \to \infty$, $p_i = n_i/N$ is the probability that a system chosen at random has energy E_i, and

$$\frac{E}{N} = \sum p_i E_i \tag{44}$$

is the mean energy of the ensemble. The total number W of posible configurations of this ensemble is

$$W = \frac{N!}{n_1! n_2! \dots} \tag{45}$$

We now evaluate the maximum value of W subject to the constraints that N and E are fixed [11]. For large N we can approximate $ln(W)$ by

$$ln(W) \approx N(ln(N) - 1) - \sum n_i(ln(n_i) - 1) = -N \sum p_i ln(p_i) \tag{47}$$

and determine p_i by the variational equation

$$\delta W - \beta \delta E - \alpha N \sum \delta p_i = 0, \tag{48}$$

---

[11]We like to stress the important point that the maximum of the entropy $S = kln(W)$ should not be subject to other constraints that have absolutely no physical meaning. Unfortunately, this has been done in the vast literature on q-entropy. There the quantity that is kept constant is an unphysical variable U_q defined by

$$U_q = \sum \frac{p_i^q}{\sum_j p_j^q} E_i, \tag{46}$$

where q is an undetermined constant. Moreover, to add to this error, U_q is claimed to be the energy of the system, but if p_i is the fundamental probability distribution, this identification, according to the principles of statistics (see section I), is *wrong* unless $q = 1$

where α and β are Lagrange parameters, and

$$\delta ln(W) = -N \sum \delta p_i ln(p_i), \tag{49}$$

and

$$\delta E = N \sum \delta p_i E_i. \tag{50}$$

Hence, we obtain the canonical exponential form for p_i,

$$p_i = e^{-(\alpha+\beta E_i)}, \tag{51}$$

where the normalization condition for the probabilies p_i, $\sum_i p_i = 1$, Eq. 1, determines α as a function of β,

$$\alpha = ln(Z(\beta)), \tag{52}$$

and

$$Z(\beta) = \sum e^{-\beta E_i}. \tag{53}$$

References

[1] Ludwig Boltzmann, *Über die Beziehung zwischen den zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewitch*,
Wien, Ber. 76. (1877) 373-435. Reprinted in *Wissenschaftliche Abhandlunger von Ludwig Boltzmann*, volume 1 (Chelsea, New York, 1968)

[2] J.W. Gibbs, *Elementary Principles in Statistical Mechanics* (1902) Reprinted by OX BoW Press, 1981.

[3] A.Einstein, *Kinetische Theorie des Wärmegleichgewichtes und des zweiten Hauptsatzes der Thermodynamik*
Annalen der Physik 9 (1902) 417-433
Zur allgemeinen molekularen Theorie der Wärme
Annalen der Physik 14 (1904) 354-362. Reprinted in *The Collected Papers of Albert Einstein* vol. 1 edited by J. Stachel (Princeton, 1989)