

The random-variable canonical distribution

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2001 J. Phys. A: Math. Gen. 34 2913

(<http://iopscience.iop.org/0305-4470/34/14/301>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.95

The article was downloaded on 02/06/2010 at 08:55

Please note that [terms and conditions apply](#).

The random-variable canonical distribution

L E Beghian^{1,2,3} and E Sheldon^{2,4}

¹ Magdalen College, University of Oxford, Oxford, UK

² University of Massachusetts Lowell, Lowell, MA 01854, USA

Received 22 August 2000, in final form 12 January 2001

Abstract

An alternative interpretation to Gibbs' concept of the canonical distribution for an ensemble of systems in statistical equilibrium is proposed. Whereas Gibbs' theory is based upon a consideration of systems subject to dynamical law, the present analysis relies neither on the classical equations of motion nor makes use of any *a priori* probability of a complexion; rather, it makes avail of the basic algebra of random variables and, specifically, invokes the law of large numbers. Thereby, a canonical distribution is derived which describes a macrosystem in probabilistic, rather than deterministic, terms, and facilitates the understanding of energy fluctuations which occur in macrosystems at an overall constant ensemble temperature. A discussion is given of a modified form of the Gibbs canonical distribution which takes full account of the effects of random energy fluctuations. It is demonstrated that the results from this modified analysis are entirely consonant with those derived from the random-variable approach.

PACS numbers: 0520, 0250, 0530, 0570

1. Introduction

A tantalizing feature of statistical mechanics and thermal dynamics that has come to the fore in the course of the past decades is the recognition that the basic tenets of statistical mechanics—namely, the microcanonical and canonical ensembles—do not readily lend themselves to rigorous, definitive proof by the equations of motion inherent in classical or quantum physics. A thorough discussion of the underlying problems has been given in numerous comprehensive texts, such as those by Jancel [1] and Balescu [2].

The object of the present study is to demonstrate that the development of the concept of the canonical distribution can be achieved through the use of basic principles of random-variable algebra, in contrast to the deterministic character of classical or quantum mechanics.

Traditionally, the theory of ensembles considered by Gibbs deals with systems consonant with dynamical law, which follow Liouville's law. This approach, under certain conditions,

³ Waynflete Fellow. Retired. Please direct all correspondence to home address: 401 Briny Avenue, Apt 302, Pompano Beach, FL 33062, USA.

⁴ Retired. Home address: 56 Cunliffe Close, Oxford, Oxon. OX2 7BL, UK.

also implies a condition of statistical equilibrium, devoid of fluctuations. Any difficulties of this kind can, as shown below, be removed by the use of random-variable theory. The canonical distribution thereby no longer constitutes a law of distribution in phase space but one which, instead, applies to an individual system in statistical equilibrium, possessing a high state probability.

2. The basic ensemble

The conventional canonical ensemble can be considered to be a collection of microcanonical ensembles [3], that are independent but subject to small variations in temperature associated with any changes in mutually neighbouring conditions of equilibrium. This condition is required in the conventional ensemble in order to take account of energy fluctuations that occur in the corresponding *macrosystems*. For these underlying aspects, the reader is referred to the text by Klimontovich [4], who provides a thoroughgoing consideration of fluctuations and of the dynamical evolution of systems, while that by Balian [5] deals with the distinction between microsystems and macrosystems in this context. By way of contrast with the *conventional* canonical ensemble made up of microsystems, the present canonical ensemble considered below is deemed to consist of a collection of independent *macrosystems*, each such macrosystem interacting with a heat reservoir maintained at a common temperature, but each system possessing a random total energy. Acting non-collectively, the systems are taken to be identical in all respects. In the first instance, it is assumed that the random variables characterizing each system are all subject to the same distribution although, as will be shown later, this condition can, in fact, be relaxed, giving rise to different possible distributions.

The present study concerns itself with macrosystems in statistical equilibrium, all in mutual contact and all at the *same constant temperature*. This is at the very heart of the new approach. The condition for the existence of statistical equilibrium requires the mean energies of the individual members of the ensemble to be equal and each to be a constant of the motion.

3. The law of large numbers

Each system in the ensemble is in statistical equilibrium at a fixed definite temperature. The application of the law of large numbers makes it possible to define an overall fixed, definite, average system energy. In the present environment, this can be accomplished by designating the mean of the system energy distribution, $\langle E \rangle$, as this average energy. It should be emphasized that the actual energy distribution, although identical for all systems, has yet to be identified. In this respect, an important theorem, known as the law of large numbers [6], relates $\langle E \rangle$ to the random variables E_1, \dots, E_r, \dots . This law is pivotal to the development of the random-variable canonical distribution. It states that the distribution of the sum of a large number of random variables X will asymptotically approach a *normal distribution*. Consequently, it can be shown that the arithmetic mean of a large number of measurements of the values of a random variable will with utmost probability approach the mean of the random variable, namely

$$\sum_{r=1}^N E_r / N \rightarrow \langle E \rangle \quad \text{as } N \rightarrow \infty$$

or

$$\sum_{r=1}^N E_r / N = \langle E \rangle + \varepsilon(N) \quad \text{wherein } \varepsilon(N) \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

4. The random-variable canonical distribution

It bears emphasizing that, in contrast to the Gibbs-ensemble treatment in which a number of non-interacting microcanonical systems, each obeying the classical laws of dynamics, are considered to establish the behaviour of the macro-ensemble, the present approach avoids the inherent constraints of classic dynamics and ergodic theory by employing a treatment that applies the law of large numbers to a set of identical independent macrosystems, each interacting with a heat source and each possessing random total energy at any one time. Thereby, the ergodic principle or related principles are no longer featured in establishing the distribution. The law of large numbers is valid for any distribution of the variables E_r provided that the distribution is identically the same for all systems of the ensemble. However, even this condition can be relaxed; a short discussion of this aspect will be presented later in the text.

The present objective is to determine the most probable distribution P of the ensemble, subject to the constraints of the law of large numbers and of the fixed number of systems comprising the ensemble.

The probability for any member of the overall macrosystem to be in a given (random) state r will be designated by a *random* variable as p_r . Such random probabilities $p_1, p_2, \dots, p_r, \dots$, etc are, accordingly, unrelated to any prior probabilities for the components of the system or of the system as a whole, therein bypassing the whole question of *a priori* probability. If we suppose there to be n_1 such identical systems with probability p_1 and energy E_1 , and n_2 identical systems with probability p_2 and energy E_2 , etc, and a total number of N systems making up the ensemble (bearing in mind that the members of the ensemble are all independent of one another) we find the net state probability to be

$$P = \prod_{r=1}^N p_r^{n_r} \quad (1)$$

or

$$\ln P = \sum_{r=1}^N n_r \ln p_r. \quad (2)$$

The condition for the law of large numbers is given by the following two expressions:

$$\sum_{r=1}^N n_r E_r = N \langle E \rangle \pm N \varepsilon(N) \quad (3)$$

and

$$\sum_{r=1}^N n_r = N. \quad (4)$$

From equations (3) and (4) it follows that

$$\sum_{r=1}^N n_r [E_r \pm \varepsilon(N)] = N \langle E \rangle. \quad (5)$$

Any deviations from the uniform energy distribution would be inconsistent with the original premise that the systems are identical in all respects. It should further be noted that the consistency requirement also mandates that out of all the possible choices of systems, only the most probable configuration can be the sole option, consonant with equation (6) below. Accordingly, the most probable condition for $\ln P$ consistent with conditions (3) and (4) is determined in the conventional variational manner by using the method of undetermined multipliers. The variational principle, as used in the present context, thus has two salient functions, namely

- (a) it enables one to predict the most likely situation prevailing for the total ensemble; and
 (b) it substantiates the basic assumption that the random-variable distributions are indeed the same for all systems, i.e. the macrosystems are just identical copies of one another.

In this, it can be contrasted with the Gibbs approach, in which a collection of microcanonical systems made up of indistinguishable particles in an incompressible mélange comprise the overall system. One can therefore argue that in the current treatment one *must* use the variational principle in order to avoid inconsistency within this approach and vindicate the validity of the development. The method of undetermined multipliers gives, accordingly,

$$\sum_{r=1}^N \delta n_r \ln p_r = 0 \quad (6)$$

namely

$$\sum_{r=1}^N \delta n_r [E_r \pm \varepsilon(N)] = 0 \quad (7)$$

wherein

$$\sum_{r=1}^N \delta n_r = 0. \quad (8)$$

Consequently,

$$\ln p_r + \lambda + \mu [E_r \pm \varepsilon(N)] = 0 \quad (9)$$

where λ and μ are the undetermined multipliers. Hence as $N \rightarrow \infty$,

$$\ln p_r + \lambda + \mu [E_r] = 0 \quad (10)$$

or, in conventional form,

$$p_r = \exp[-\lambda - \mu E_r]. \quad (11)$$

It bears emphasizing at this juncture that E_r will not in general be a time-independent function. Each value of E_r is a random quantity and, as will be discussed in the next section, each member of the ensemble will be subject to random energy variation, as a result of energy exchange with the heat reservoir.

As already mentioned, the law of large numbers may also be valid under less stringent conditions than those stipulated above. Specifically, the sum of a large number of random variables may also obey normal conditions for $N \rightarrow \infty$ even if the variables are subject to other than normal distributions [7]. The necessary condition for this to occur can (approximately) be expressed as follows:

$$\lim_{N \rightarrow \infty} \sigma_j / s_n = 0 \quad \text{for all } j$$

where σ_j^2 is the variance of the j th term, and $s_n^2 = \sum_{j=1}^N \sigma_n^2$. However, since it has been demonstrated that the condition for a maximum value of P requires a *uniform* energy distribution for the ensemble members, any such less stringent condition is unlikely to be advantageous in practice.

5. The Liouville equation and the random-variable canonical ensemble

At first sight, there would seem to be a marked similarity between Gibbs' formulation of the canonical distribution and the random-variable distribution. However, this, in fact, is not the case. The variable distribution has, built into the system (wherein, henceforth, each member of the ensemble will be designated as a 'system'), the contribution of random energy fluctuations resulting from the interaction between each system and the heat reservoir. On the other hand, the Gibbs formulation does *not a priori* include the effect of such an interaction. However, in a recent analysis [8] involving a modification of the standard Liouville equation, it was shown that it was possible to include the effects of random energy fluctuations within a comprehensive Gibbs-type formalism.

The relevant time-dependent Liouville equation giving the partial time derivative of the probability density ρ_s (per unit volume) for a system s in statistical equilibrium herein takes the form

$$-\frac{\partial \rho_s}{\partial t} = \sum_j \left(\frac{\partial H_s(t)}{\partial p_j} \times \frac{\partial \rho_s}{\partial q_j} - \frac{\partial H_s(t)}{\partial q_j} \times \frac{\partial \rho_s}{\partial p_j} \right) + \sum_j \frac{\partial}{\partial p_j} (\rho_s f_j) \quad (12)$$

where the Hamiltonian, as we stress, is time-dependent in character, and its time-average $\langle H_s(t) \rangle$ is, by definition, a constant of the motion. The sums in the second term for the j th particle within the system s extend over the contribution f_j of time-dependent forces from external sources upon j . The canonical variables p_j and q_j for the j th particle obey the conditions

$$\dot{p}_j = -\frac{\partial H_s(t)}{\partial q_j} + f_j \quad \text{and} \quad \dot{q}_j = \frac{\partial H_s(t)}{\partial p_j}. \quad (13)$$

If one considers the *stationary* case and sets $\partial \rho_s / \partial t = 0$ one obtains, after iteration using Pascal's method, the solution of the Liouville equation (12) as the density

$$\rho_s = \text{constant} \times \exp[-H_s/kT] \quad (14)$$

where H_s represents the *stationary* Hamiltonian. Modifying this to allow for *time dependence* in a statistical equilibrium situation, one obtains *formally* the same solution as (14) but with a *time-dependent Hamiltonian* [8]:

$$\rho_s = (1/W) \exp[-H_s(t)/kT] \quad (15)$$

where W is the sum over states. This is in conformity with the customary definition of the temperature T through the relation

$$\frac{1}{T} = \frac{d}{dE} \ln[\rho(E)]. \quad (16)$$

The basic criterion for statistical equilibrium is taken to be the same as for the random-variable solution, namely that the average energy of an incompressible system, the members of which are physically indistinguishable, is effectively a constant of the motion. Energy fluctuations about the average system energy are inherently random; the temperature of the overall system (namely, ensemble system plus reservoir) must accordingly also be a constant of the motion, as otherwise heat diffusion would occur throughout the system. At the local level, the canonical temperature T_s is thus a constant of the motion, and only fluctuations of the *kinetic* degrees of freedom are an issue here. One notes, however, that fluctuations of *entropy* may nevertheless take place, since transfer of free energy in the kinetic phase can and does occur.

Since it has also been demonstrated in an earlier analysis [9] for the dynamics of a single particle that the momentum coordinates in the momentum space of a system obey the Maxwell

distribution and, further, that such parameters are statistically independent of one another, it also follows that the kinetic energy fluctuations of the system are random. From the Maxwellian distribution it can readily be shown that a system energy variance is $\sim(n)^{1/2}kT$, where n is the total number of molecules comprising the system, k is Boltzmann's constant and T is the reservoir temperature.

The foregoing considerations prompt an examination of the thermodynamic aspects of the random-variable canonical ensemble, contrasted against the conventional Gibbs microcanonical approach. This is undertaken in the next section via the detailed analysis of the (Helmholtz) free energy F and the Gibbs–Helmholtz equation, to arrive at a comparison of the (Gibbs) entropy S_G and the microcanonical entropy S_m , the latter being defined through the relation

$$\frac{dS_m}{dE} = \frac{1}{T}. \quad (17)$$

6. Energy, free energy and entropy in the thermodynamics of the classical and random-variable canonical ensemble

The treatment in this section follows conventional lines, as presented in most basic texts (e.g. [10]), but relies heavily upon the recent synopsis by Cole [11] and employs generally the same notation.

The conventional development involving discrete *microstates*, the s th state having energy E_s and probability p_s (not to be confused with the p_j of section 5), proceeds from the canonical partition function Z , defined through

$$Z = \sum_s \exp(-\beta E_s) \quad (18)$$

with the coefficient β (representing an inverse energy, equivalent to within a factor of k to an inverse temperature) given by the differential

$$\beta \equiv \frac{1}{T} = \frac{d}{dE} [\ln(\rho(E))] \quad \text{at } E = E_{\max} \quad (19)$$

where $\rho(E)$ is the energy-dependent canonical state density.

The microstate probability is, accordingly,

$$p_s = (1/Z) \exp(-\beta E_s) \quad (20)$$

and the average value of the energy (i.e. the internal energy of the system) is thus

$$U \equiv \langle E \rangle = \sum_s p_s E_s \quad (21)$$

namely

$$U = \frac{\sum_s E_s \exp(-\beta E_s)}{\sum_s \exp(-\beta E_s)} \equiv \frac{\sum_s E_s \exp(-\beta E_s)}{Z}. \quad (22)$$

With this, the (Helmholtz) free energy takes the form

$$F = -T \ln(Z) \quad (23)$$

and thus the microstate probability (20) can be re-expressed as

$$p_s = \exp[\beta(F - E_s)]. \quad (24)$$

The sum of the probabilities is unity for all values of β ; hence the differential of this sum with respect to β vanishes. Moreover, one sees that $\beta \partial / \partial \beta = -T \partial / \partial T$ and that the free energy F (at constant β and volume V) is a constant, whence

$$\sum_s F \exp[\beta(F - E_s)] \equiv \sum_s F p_s = F. \quad (25)$$

With this, and with the aid of equation (22), one obtains the Gibbs–Helmholtz equation for a given system volume V :

$$U(T) = \left(1 - \frac{\partial}{\partial \ln T}\right) F = F - T \frac{\partial F}{\partial T}. \quad (26)$$

The partial derivative in the final term ensues from the free-energy definition (23), after some algebra, as

$$-\frac{\partial F}{\partial T} = \frac{\partial}{\partial T} \left\{ T \ln \left[\sum_s \exp\left(-\frac{E_s}{T}\right) \right] \right\} = -\sum_s p_s \ln(p_s). \quad (27)$$

It is this which defines the (Gibbs) entropy:

$$S_G = -\partial F / \partial T = -\sum_s p_s \ln(p_s) \quad (28)$$

or, if the temperature T is given in Kelvin (rather than in energy units as Joule):

$$S_G = -k \sum_s p_s \ln(p_s) \quad (29)$$

where k is Boltzmann's constant. It is evident that this definition of the entropy S is consonant with the Gibbs–Helmholtz equation written in the usual way:

$$U = F + TS. \quad (30)$$

On the other hand, the microcanonical entropy S_m is defined as the natural logarithm of the microcanonical sum of substates s (the reciprocal of the microcanonical probability). A system which has only one microstate is deemed to have zero entropy ($S_m = 0$).

To compare this with the Gibbs entropy, we rewrite the Gibbs–Helmholtz equation (26) with the definition (28) as

$$-\frac{\partial F}{\partial T} \equiv S_G = \frac{U}{T} - \frac{F}{T} \quad (31)$$

so that, with (20) and (23), one obtains

$$S_G = \beta U + \ln(Z). \quad (32)$$

Therefore (in the $k = 1$ convention),

$$\frac{\partial S_G}{\partial U} = \beta = \frac{1}{T} \quad (33)$$

whereas, from equation (17) one finds

$$\frac{dS_m}{dE} = \frac{1}{T}. \quad (34)$$

It is thus the *average* energy U in the canonical ensemble which plays the role of the energy E in the microcanonical ensemble. This is entirely consistent for large systems in which the fluctuations away from average values (of energy or probability, wherein the averaging is carried out over all states of the combined system) can, at least to first order, be neglected. For further details, see [4, 5].

As Nussbaum [12] has pointed out, maximizing S with Lagrange multipliers leads to the appropriate overall statistical distribution: Maxwell–Boltzmann, Fermi–Dirac or Bose–Einstein.

In the light of the foregoing demonstration that the random-variable treatment with the law of large numbers is entirely equivalent in its outcome to the conventional (Gibbs-type) microcanonical treatment of systems with a large number of (fluctuating) substates, it is intuitively obvious that the above findings apply just as well in the $N \rightarrow \infty$ limit of the random-variable canonical distribution formalism. We can therefore quite justifiably conclude that the latter constitutes an equally valid approach to the development of all thermodynamic results, such as the above, even in the presence of fluctuations.

7. Concluding remarks

With the above-mentioned modifications, one finds a striking similarity between the two seemingly disparate distributions, inasmuch as:

- (a) both are based on an ensemble involving indistinguishable, non-interacting members;
- (b) the condition for statistical equilibrium is the same for either case, and has been shown to be based upon the law of large numbers;
- (c) the two distribution laws are similar in form, albeit when the formalism is generalized to involve a time-dependent Hamiltonian;
- (d) in both cases, the presence of random energy fluctuations is accommodated within the (time-dependent) Hamiltonian;
- (e) both approaches lead to the same, consistent outcome in thermodynamics.

Notwithstanding the random energy fluctuations within the system, the temperature of each member in the ensemble is a constant as thermal equilibrium prevails throughout the system; there is no net overall change in entropy and no violation of the second law of thermodynamics.

Finally, a brief statement is given below, reiterating the principal assumptions involved in the development of the random-variable canonical distribution in comparison with the standard Gibbs approach.

There are two main assumptions made in the present development:

- (a) the validity of the law of large numbers; and
- (b) the systems comprising the ensemble are taken to be identical in all respects (i.e. ‘clones’ of one another), and the random variables characterizing each system are all subject to the same distribution.

On the other hand, the Gibbs approach is based on the equilibrium solution of the Liouville equation. This equation is not founded on statistical arguments but on the conservation of transport of identical, but independent, members of an ensemble taken in conjunction with the classical Hamiltonian equations of motion.

Specifically, the statistical approach adopted in this paper has the advantage of avoiding any difficulties which might be encountered with problems of ergodicity or the validity of *a priori* probabilities.

From this work, one observes the remarkable kinship between the modified Liouville theorem based on the equations of classical mechanics, and the random-variable canonical distribution based on classical probability theory. Despite their disparate origins and seemingly distinct approaches, the findings reveal their intrinsic inter-consistency and serve to demonstrate an inherent relationship between the Gibbs result and the random-variable approach. Thereby, the large-number condition ($N \rightarrow \infty$) manifests *statistical* mechanics

in its very essence as a formalism that not only bridges but, in a fundamental sense, actually *transcends* the oftentimes artificial demarcation between the complementary domains of classical and quantum physics.

References

- [1] Jancel R 1969 *Foundations of Classical and Quantum Statistical Mechanics* (Oxford: Pergamon)
- [2] Balescu R 1975 *Equilibrium and Non-Equilibrium Statistical Mechanics* (New York: Wiley)
- [3] Morse P M 1960 *Thermal Physics* 2nd edn (New York: Benjamin) p 271
- [4] Klimontovich Yu L'v 1982 *Statisticheskaya Fizika* (Moscow: Nauka); Engl. transl. 1986 *Statistical Physics* (New York: Harwood)
- [5] Balian R 1991 *From Microphysics to Macrophysics* (Berlin: Springer)
- [6] Beckmann P 1967 *Probability in Communication Engineering* (New York: Harcourt Brace) p 100
- [7] Gnedenko B V and Kolmogorov A N 1954 *Limit Distributions of Sums of Independent Random Variables* (Reading, MA: Addison-Wesley) (Engl. transl.)
- [8] Beghian L E 1995 *Nuovo Cimento B* **110** 1369–74
- [9] Beghian L E 1992 *Nuovo Cimento B* **107** 141–52
- [10] Sears F W 1959 *Thermodynamics* (Reading, MA: Addison-Wesley)
- [11] Cole A J 2000 *Statistical Models for Nuclear Decay: From Evaporation to Vaporization* (Bristol: IOP Publishing) pp 5ff
- [12] Nussbaum A 2000 *Phys. Today* October pp 11–2