International Journal of Theoretical Physics, Vol. 29, No. 1, 1990

Underlying Probability Distributions of the Canonical Ensemble

B. H. Lavenda[1] **and J. Dunning-Davies**[2]

Received June 20, 1989

The canonical distributions are chi-square distributions which are derived from parent distributions for nonconjugate fluctuating thermodynamic variables. The probability distributions are generated by discrete random variables which are the number of degrees of freedom and the number of particles. Randomized sampling of the total number of degrees of freedom and total number of particles gives rise, respectively, to fluctuations in the energy and volume.

1. INTRODUCTION

Statistical distributions describe chance phenomena such as the throw of a die or the toss of a coin. Allocation experiments are of this kind, where, for example, a certain number of balls are to be distributed among a given number of cells. Such types of experiments provide a convenient way of deriving quantum statistics within the framework of the grand canonical ensemble, but do not seem to be pertinent to the continuous distribution of extensive thermodynamic variables in the canonical ensemble. Boltzmann himself always worked with discontinuous random variables, only to take the continuous limit at the end of the calculation. This is undoubtedly the reason why his methodology could be used by Planck in the derivation of his radiation formula. But the question still remains: What is responsible for the fluctuations in the extensive thermodynamic variables, such as the energy and volume, which must be admitted even in the microcanonical ensemble in order to obtain a probability distribution?

The term "distributions" is used in a peculiar sense in physics since they are not probability distributions at all, but rather expressions for the average number of particles as functions of temperature at a given energy

---

[1] Università degli Studi, Camerino 62032 (MC), Italy.

[2] Department of Applied Mathematics, University of Hull, Hull HU6 7RX, United Kingdom.

or frequency. They can, however, be derived from proper probability distributions by equating the derivative of the entropy of the probability distribution with the second law. The entropy of the probability distribution is determined by casting the probability distribution as a Gaussian law of error (Lavenda, 1988). Classical or Maxwell-Boltzmann statistics arises, in this context, as a limiting distribution to which the quantum distributions tend as the ratio of the average number of particles to the number of cells tends to zero. The classical probability distributions always appear as *error laws* which identify the average value as the most probable value. In this paper we will show that such continuous probability densities are generated by discrete random events that are governed by the bi- or multinomial distributions in an somewhat analogous way to that in which Bose-Einstein and Fermi-Dirac statistics are derived from the negative binomial and binomial distributions (Lavenda, 1988).

But the relation appears to be more subtle, since the bi- or multinomial distributions are not the probability distributions for thermodynamic variables. The multinomial distributions for the occupation numbers are converted into independent Poisson distributions by randomized sampling. The presence of scale parameters, proportional to intensive thermodynamic variables, in these distributions indicates that there exist other random variables for which these parameters are the true scale parameters. The density of the cumulative distribution function is the probability density of the extensive thermodynamic variables. This is entirely distinct from the uncertainty relations between conjugate thermodynamic variables (Lavenda, 1987); random occupation numbers give rise to continuous random thermodynamic variables with probability densities all having the same form when the constraint that their sum be fixed is waived. That the various probability densities are all chi-square distributions seems to imply that the law of large numbers is actually behind the success of thermodynamics.

All this can be distilled out of the cryptic Chapter II of Gibbs' (1902) *Elementary Principles in Statistical Mechanics.* Here Gibbs actually derives the canonical distribution in a much more rigorous way than the heuristic arguments he offers in Chapter IV. Although he notes the analogy with the theory of errors and that the probability distribution depends only on the number of degrees of freedom, "being in other respects independent of its dynamical nature," Gibbs stops short of asking what are the random variables for the probability distribution that he has derived.

2. TEMPERATURE ENSEMBLES

Consider a phase space with $2m$ degrees of freedom $x_1, \ldots, x_{2m}$ without giving them any mechanical interpretation. A system can be represented by a point with coordinates $x_1, x_2, \ldots, x_{2m}$ in this phase space. The volume element in the phase space is (Khinchin, 1949)

$$d\mathcal{V}(;\, m) = dx_1 \cdots dx_{2m} = \Omega(E;\, m-1)\, dE \tag{1}$$

where $\Omega(E;\, m-1)$ is the surface "area" of the volume $\mathcal{V}(E;\, m)$, viz. (Gibbs, 1902)

$$\Omega(E;\, m-1) = \frac{d\mathcal{V}}{dE} = A^m \frac{E^{m-1}}{\Gamma(m)} \tag{2}$$

and A is a numerical coefficient that depends upon the nature of the physical system under consideration. $\Omega(E;\, m-1)$ is commonly referred to as the "structure function" and it contains all the relevant thermodynamic information about the system (Khinchin, 1949).

Suppose that the phase space is divided into two parts, one with volume

$$\mathcal{V}_1(E_1;\, k) = \int \cdots \int dx_1 \cdots dx_{2k} = A^k \frac{E_1^k}{k!} \tag{3}$$

where the integration extends over all phases for which the energy is not greater than E_1, and another with volume

$$\mathcal{V}_2(E_2;\, m-k) = \int \cdots \int dx_{2k+1} \cdots dx_{2m} = A^{m-k} \frac{E_2^{m-k}}{(m-k)!} \tag{4}$$

The integration in (4) extends over all phases for which the energy is not superior to E_2. The total energy of the composite system is

$$E = E_1 + E_2 \tag{5}$$

The energies E_1 and E_2 are not assumed to be functions of the instantaneous values of the degrees of freedom, and hence are not random quantities; rather, they are related to the expected number of degrees of freedom in each of the subspaces.

We cannot be exactly sure that there are $2k$ degrees of freedom in subsystem "1" for there is no way of controlling precisely the microscopic coordinates. Rather, we must ask for the probability that out of the entire phase space of $2m$ degrees of freedom one of the two subspaces contains exactly $2k$ degrees of freedom. This is given by the binomial distribution

$$P(k \mid m) = \binom{m}{k} p^k q^{m-k} = \frac{\mathcal{V}_1(E_1;\, k)\mathcal{V}_2(E-E_1;\, m-k)}{\mathcal{V}(E;\, m)} \tag{6}$$

where the *a priori* probability $p = E_1/E$ and $q = 1-p$. On the strength of Newton's binomial theorem, the composition rule for the phase space volume elements is easily seen to be

$$\mathcal{V}(E) = \sum_{k=0}^{m} \mathcal{V}_1(E_1;\, k)\mathcal{V}_2(E-E_1;\, m-k) \tag{7}$$

At first sight, (7) may appear to be an unusual composition law for the phase volumes. Yet, it may be remembered that it is the degrees of freedom of the subspaces and not the energies which are the random variables. When we transfer from degrees of freedom to the energies of the subsystems as the relevant random variables, the surface areas Ω_1 and Ω_2 will be seen to obey the usual composition law [cf. (16) below].

Suppose now that we release the constraint that the total number of degrees of freedom $2m$ be fixed and replace it by a random number having the Poisson distribution

$$P(m|\beta)=\frac{(\beta E)^m}{m!}e^{-\beta E} \tag{8}$$

The parameter β is a physical constant which determines how many degrees of freedom will be expected in each energy interval. In mathematical statistics, the replacement of a fixed number by a random number is known as "randomized sampling."

Then "total probability" is the product of binomial distribution (6) and the Poisson distribution (8) summed over all $m\geq k$ (Feller, 1968)

$$P(k|\beta)=\sum_{m=k}^{\infty}P(k|m)P(m|\beta)=\frac{(\beta E_1)^k}{k!}e^{-\beta E_1} \tag{9}$$

Under randomized sampling, the random variable k becomes an independent Poisson variable with mean value βE_1. The quantity β appears as a scale parameter but it is *not* a scale parameter in the distribution (9). This seems to indicate that there is another random variable for which β is the scale parameter.

Suppose we are able to increase the phase volume $\mathscr{V}_1$ *continuously*, by raising the energy E_1, until it contains exactly $2k$ degrees of freedom. This implies that the energy is the required random variable. We designate $\tilde{E}_1$ as that energy for which the volume $\mathscr{V}_1$ contains exactly $2k$ degrees of freedom. Then the probability that the volume $\mathscr{V}_1(E_1)$ will contain less than $2k$ degrees of freedom is

$$\Pr(\tilde{E}_1>E_1)=\sum_{n=0}^{k-1}\frac{(\beta E_1)^n}{n!}e^{-\beta E_1} \tag{10}$$

Thus, $\Pr(\tilde{E}_1\leq E_1)=1-\Pr(E_1>\tilde{E}_1)$ will coincide with the cumulative distribution function $F(E_1|\beta)$ (Blackwell and Girshick, 1954). This is precisely expression (68) in Gibbs (1902). Then

$$f_1(E_1|\beta)=\frac{\partial F}{\partial E_1}=\frac{(\beta E_1)^{k-1}\beta}{\Gamma(k)}e^{-\beta E_1} \tag{11}$$

is the probability density which, as Gibbs notes, depends upon the number of degrees of freedom and is independent of the particular nature of the physical system. Actually, Gibbs integrates (11) to get the cumulative distribution function, for which he assumes that there exists a value of the energy such that $F=\frac{1}{2}$. He then concludes that the probability of the phase falling between these limits is greater than the probability of falling in any other limits enclosing an equal-extension-in-phase. It happens, as he notes, to be equal to the probability of the phase falling outside these limits.

The chi-square distribution possesses the important "reproductive" property that the sum of two independent chi-square variates with $2k$ and $2l$ degrees of freedom is also a chi-square variate with $2k+2l=2m$ degrees of freedom, viz.,

$$f(E\,|\,\beta)=\frac{\beta(\beta E)^{m-1}}{\Gamma(m)}\,e^{-\beta E} \tag{12}$$

where E is the sum of the energies. This property is responsible for statistical equilibrium and the fundamental composition law for the structure function, equation (16).

From a statistical inference point of view (Lavenda and Scherer, 1988), the energies E_j can be interpreted as sample data which can be used to estimate the conjugate intensive parameter β. The maximum likelihood estimate is

$$\bar{E}=m/\beta \tag{13}$$

where $\bar{E}$ coincides with the mean of the chi-square distribution (12). A thermodynamic argument can be used to establish that β is the inverse temperature in energy units where Boltzmann's constant is equal to unity. Equation (13) may then be appreciated as the law of equipartition for a system with $2m$ degrees of freedom. According to the chi-square density, (12), it represents the expected value of the energy, while according to randomized sampling hypothesis, characterized by the fact the random number m has a Poisson distribution, it represents half the mean number of degrees of freedom. This is a fundamental complementarity relation whereby the relaxation of the constraint that the total number of degrees of freedom is a constant induces fluctuations in the energy. In the presence of the constraint, the energy appears as a parameter determining the mean number of degrees of freedom; the number of degrees of freedom is the random variable.

The chi-square distribution (12) can be written in the canonical form

$$f(E\,|\,\beta)=\frac{\Omega(E)}{\mathscr{Z}(\beta)}\,e^{-\beta E} \tag{14}$$

where Ω is the structure function (2) with $2m$ degrees of freedom. The norming constant or "partition function" is

$$\mathscr{Z}(\beta)=(A/\beta)^m \tag{15}$$

Consequently, the ratio $\Omega/\mathscr{Z}$, and hence the probability density (14), is independent of the particular nature of the physical system (i.e., independent of the numerical constant A), while it depends on the number of degrees of freedom of the system.

The probability density (14) can be written as a ratio of products of structures functions by dividing our system into a small subsystem, which we label as "1" with $2m_1$ degrees of freedom, the a larger (reservoir) subsystem, labeled by "2" with $2m_2$ degrees of freedom. The composite system will be index free and has $2m_1+2m_2=2m$ degrees of freedom. The random quantity $\tilde{E}_1$ is governed by the probability density (Khinchin, 1949)

$$\frac{\Omega_1(E_1;\, m_1-1)\Omega_2(E-E_1;\, m-m_1-1)}{\Omega(E;\, m-1)} \tag{16}$$

since $E_1+E_2=E=\text{const}$. Expanding Ω_2 in a Taylor series about the total energy E of the system, using $\partial \ln \Omega_2/\partial E=\beta$ and $\Omega_2(E)/\Omega(E)\simeq \mathscr{Z}^{-1}(\beta)$ (Blanc-Lapierre and Tortat, 1954), we obtain the canonical expression (14) for the probability density.

However, there is a flaw in the demonstration, because even if the successive terms in the Taylor series of $\Omega_2(E-E_1)$ about E are small compared with the first-order term, a nonnegligible error is introduced when the first two terms are used to approximate an exponential. This can be remedied by observing that (16) is the beta probability density

$$f_1(x\,|\,\beta)=\frac{x^{m_1-1}(1-x)^{m-m_1-1}}{B(m_1,\, m-m_1)} \tag{17}$$

where the beta function is $B(m_1;\, m-m_1)=\Gamma(m_1)\Gamma(m-m_1)/\Gamma(m)$ and $x=E_1/(E_1+E_2)$. It is well known (Zellner, 1971) that the beta probability density can be derived from a product of chi-square densities (12) for the random variables $\tilde{E}_1$ and $\tilde{E}_2$ by a change of variables. Allowing the number of degrees of freedom of the reservoir, or subsystem "2", to increase without limit, will affect a transformation of the beta into the chi-square distribution for the random energy variable $\tilde{E}_1$. This is tantamount to letting $m\to\infty$ with m_1 kept small and fixed. In this limit,

$$\lim_{m\to\infty}\left(1-\frac{\beta E_1}{m}\right)^m=e^{-\beta E_1} \tag{18}$$

and consequently the beta distribution goes over into the chi-square distribution for the random variable $\tilde{E}_1$.

3. PRESSURE ENSEMBLES

The type of complementarity between nonconjugate thermodynamic variables also appears between the number of particles and the partial volumes which they occupy. The probability of finding a molecule in a partial volume V_j of total volume V is

$$p_j = \frac{V_j}{V} \tag{19}$$

Dividing the entire volume into r subvolumes, the probability of finding n_1 molecules in V_1, n_2 molecules in $V_2, \ldots,$ is given by the multinomial distribution

$$P(n_1, n_2, \ldots, n_r \mid n) = \frac{n!}{n_1!\, n_2! \cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r} \tag{20}$$

subject to the condition that

$$\sum_{j=1}^{r} n_j = n \tag{21}$$

be a constant.

Now suppose that we relax the condition that (21) be constant. If the particles are well mixed, their distribution will be Poisson. Let λ be the expected number of particles in a unit volume. The distribution of the total number of particles n is given by

$$P(n \mid \lambda) = \frac{(\lambda V)^n}{n!} e^{-\lambda V} \tag{22}$$

The "total probability" is the product of the multinomial distribution (20) and the Poisson distribution (22) summed over all n (Feller, 1968)

$$\begin{aligned} P(n_1, n_2, \ldots, n_{r-1} | \lambda) &= \sum_n P(n_1, n_2, \ldots, n_r \mid n) P(n \mid \lambda) \\ &= \prod_{j=1}^{r-1} \frac{(\lambda V_j)^{n_j}}{n_j!} e^{-\lambda V_j} \end{aligned} \tag{23}$$

which again shows that randomized sampling from a multinomial distribution where the total number is a Poisson random variable produces *independent Poisson variables.*

In the same way that β is not the scale parameter for the total number of degrees of freedom, λ is not the scale parameter for the total number of particles. We now show that by varying the volume V so that for some value we have exactly n particles located in it, the size of this volume is the required random variable. Constructing the cumulative distribution function as above, we find the density to be given by

$$f_n(V|\lambda) = \frac{(\lambda V)^{n-1}\lambda}{\Gamma(n)} e^{-\lambda V} \tag{24}$$

and analogous expressions for each of the partial volumes. Thus, λ is a scale parameter in the distribution of V which is a chi-square distribution with $2n$ degrees of freedom.

The parameter λ can estimated from observations made on the volume. The maximum likelihood estimate is obtained from the likelihood equation

$$\frac{\partial \ln f_n}{\partial \lambda} = \frac{n}{\lambda} - V = 0 \tag{25}$$

We now show that the most likely value of V, determined from the likelihood equation (25), coincides with the average value $\bar{V}$. The chi-square density (24) can be written in the canonical form

$$f_n(V|\lambda) = \frac{\Omega(V;n)}{\mathscr{Z}(\lambda)} e^{-\lambda V} \tag{26}$$

where $\Omega(V;n) = V^{n-1}/\Gamma(n)$ is the "configurational" phase space structure function whose integral is

$$\mathscr{V}(V;n) = \frac{V^n}{n!} \tag{27}$$

The numerator of (27) results from integrating over all the $3n$ positional coordinates and it is what Gibbs (1902) referred to as *extension-in-configuration.* The denominator had to be introduced in an *ad hoc* manner, prior to the advent of quantum theory, in order that the entropy turn out to be extensive. Although we shall refer to (27) as the "configurational" phase space volume, in order to distinguish it from the energy phase space volume (3), we will not require any mechanical interpretation of the $\mathscr{V}$'s.

The moment generating function in (26) is

$$\mathscr{Z}(\lambda) = \lambda^{-n} \tag{28}$$

which, upon taking the logarithm and differentiating, gives

$$-\frac{\partial \ln \mathscr{Z}}{\partial \lambda} = \frac{n}{\lambda} = \bar{V} \tag{29}$$

A comparison of (29) with the likelihood equation (25) shows that the most probable value of the volume, determined from that equation, coincides with the average value, determined from the moment generating function.

The equation of state (29) allows for a complementary interpretation. In terms of the chi-square density (24), we would interpret it as the expression for the average volume, while in terms of the Poisson distribution (22), it gives the expected number of particles in terms of the volume, which is a mere parameter.

In order to attribute a thermodynamic significance to the scale parameter λ, we have to assume that the partition function (28) is also a function of the scale parameter β, the inverse temperature. Otherwise, we would come out with a meaningless relation of converting work into heat and vice versa without taking into account changes in the internal energy of the system. This points to the fact that there is a definite hierarchy in securing different types of equilibria, with the highest priority given to thermal equilibrium where all subsystems arrive at a common value of β.

If each particle has $2m$ degrees of freedom and there are n particles, the energy phase space volume

$$\mathcal{V}(E;\, mn) = A^{mn}\frac{E^{nm}}{(mn)!} \tag{30}$$

has $2mn$ degrees of freedom, where E is a sum of n terms. The total element of volume of the phase space is

$$\begin{aligned} d\tilde{\mathcal{V}}(E, V) &= d\mathcal{V}(E;\, mn)\cdot d\mathcal{V}(V;\, n)\\ &= \tilde{\Omega}(E, V;\, n)\, dE\, dV\\ &= \frac{A^{mn}E^{mn-1}V^{n-1}}{\Gamma(mn)\Gamma(n)}\, dE\, dV \end{aligned} \tag{31}$$

which can be taken as the composition law of the phase space volumes. The internal energy is

$$-\left(\frac{\partial \ln \tilde{\mathcal{Z}}}{\partial \beta}\right)_{\bar{V}} = U \tag{32}$$

where

$$\ln \tilde{\mathcal{Z}} = n \ln\left(\frac{A^m}{\beta^m \lambda}\right) \tag{33}$$

and the internal energy $U = n\bar{E}$ is n times the average energy per particle. The moment equation (29) can now be written as

$$d \ln \tilde{\mathcal{Z}} + U\, d\beta = -\bar{V}\, d\lambda \tag{34}$$

Adding the differential of $\beta U + \lambda \bar{V}$ to both sides of (34), we get

$$d\{\ln \tilde{\mathcal{Z}} + \beta U + \lambda \bar{V}\} = \beta\left\{dU + \frac{\lambda}{\beta}\, d\bar{V}\right\} \tag{35}$$

With the identification $\lambda = \mathscr{P}\beta$, where $\mathscr{P}$ is the pressure, the scale parameter β multiplies the difference between the increment in the internal energy and the work done by the system. According to the first law, this is the "amount" of heat δQ received by the system during an elementary transition between neighboring equilibrium states, while according to the second law, the product $\beta\,\delta Q$ is equal to a total differential, given by the left side of (35). The integrating factor β, for the quantity of heat δQ, in energy units where Boltzmann's constant is unity, is the inverse of the absolute temperature.

The expression for the first moment, (29), is simply the equation of state for an ideal gas, since $\lambda = \beta\mathscr{P}$. And just like the law of equipartition of energy, (13), it, too, admits a dual interpretation, depending on whether the volume is a random variable and the particle number a parameter or *vice versa.* It therefore appears that this complementary is at the root of all constitutive relations relating nonconjugate thermodynamic variables.

Comparing (35) with the Gibbs equation identifies the logarithm of the partition function as

$$\ln \mathscr{Z}(\beta, \beta\mathscr{P}, n) = -\beta(\mathscr{F} + \mathscr{P}\bar{V}) = -\beta\mathscr{G} \tag{36}$$

where $\mathscr{F}$ and $\mathscr{G}$ are the Helmholtz and Gibbs free energies, respectively. The commonly accepted expression for the partition function is

$$\tilde{\mathscr{Z}} e^{\beta\mathscr{P}\bar{V}} = \left(\frac{A}{\beta}\right)^{mn} \frac{\bar{V}^n}{n!} \tag{37}$$

where we have used (29) and Stirling's approximation $n! = n^n e^{-n}$. Ordinarily, the $n!$ in the denominator of (37) is written in by hand to take into account the indistinguishability of the particles. In other words, in writing the partition function as a product of single-particle partition functions, it is claimed that the number of configurations have been overcounted $n!$ times. But expressing the partition function as a product does not mean that we are considering only one out of the $n!$ permutations. Rather, the $n!$ enters naturally into the expression for the configurational phase space volume (27) and not into the expression for the partition function.

For a subsystem in contact with both a thermal and pressure reservoir, the probability to occupy a state of energy E and volume V is

$$f_n(e, V \mid \beta, \lambda) = \tilde{\Omega}(E, V; n) e^{\beta\mathscr{G} - \beta E - \lambda V} \tag{38}$$

The product $\beta\mathscr{G}$ is the Massieu transform of the entropy with respect to the internal energy U and the average volume $\bar{V}$, viz.

$$-\beta\mathscr{G} = \mathscr{S}[\beta, \beta\mathscr{P}, n] = \mathscr{S}(U, \bar{V}, n) - \left(\frac{\partial\mathscr{S}}{\partial U}\right)_{\bar{V},n} U - \left(\frac{\partial\mathscr{S}}{\partial \bar{V}}\right)_{U,n} \bar{V} \tag{39}$$

Introducing this expression into (38) and taking the logarithm casts the probability density in the form of a Gaussian law of error (Lavenda, 1988)

$$\ln f_n = -\beta(E-U) - \lambda(V-\bar{V}) - \mathcal{S}(U, \bar{V}, n) + s(E, V, n) \tag{40}$$

where the stochastic entropy has been defined as

$$s(E, V, n) = \ln \tilde{\Omega}(E, V; n) \tag{41}$$

in accordance with Boltzmann's principle.

The Gaussian error law (40) gives the probability that a measurement will give the joint values E and V whose true values are known to be U and $\bar{V}$, respectively. Since the probability density is a maximum when the average and most probable values coincide, the entropy must be the same function of the average values that the stochastic entropy is of the random variables. Hence,

$$\mathcal{S}(U, \bar{V}, n) = n \ln\left[\left(\frac{AU}{mn}\right)^m \frac{\bar{V}}{n}\right] + n(m+1) \tag{42}$$

Expression (42) has the form of the entropy for an ideal monatomic gas, where $A = 2\pi M/h^2$, with M the mass and h Planck's constant, with $2m = 3$ degrees of freedom. In deriving (42), we have assumed that n is sufficiently large so as to apply Stirling's approximation. In this approximation,

$$\mathcal{S}(U, \bar{V}, n) = \ln \tilde{\mathcal{V}}(U, \bar{V}; n) \tag{43}$$

which means that there is a logarithmic equivalence between total phase volume (30) and total surface area (31). Since the former is proportional to the number of states below a given energy and volume, these states will be concentrated in a thin layer near the surface. This is the justification for Ehrenfest's use of "surface ensembles" (Ehrenfest and Ehrenfest, 1959).

According to Greene and Callen (1959), the high dimensionality of the phase spaces is "responsible for the fact that there is a single general thermodynamics, rather than a 'microcanonical thermodynamics' and a separate 'canonical thermodynamics'." For systems with a large number of degrees of freedom, the structure function actually specifies the entropy function when the random variables are replaced by their average, or most probable, values. This explains why the structure function completely determines the structure of the corresponding physical system (Khinchin, 1949).

4. PARTICLE DISTRIBUTIONS

Suppose we wish to retain the fact that it is the particles themselves which give rise to the fluctuations. We have assumed that since the particles are well mixed, their distribution is given by the Poisson distributions (8), which may be written generically as

$$P(n\,|\,\bar{n})=\frac{\bar{n}^{n}}{n!}e^{-\bar{n}} \tag{44}$$

Writing it in terms of a Gaussian law of error for which the mean value is also the most probable, we have

$$P(n\,|\,\bar{n})=\exp\left\{-(n-\bar{n})\left(\frac{\partial S}{\partial \bar{n}}\right)_{V}-S(\bar{n})+s(n)\right\} \tag{45}$$

where the entropy is given by $s=\bar{n}-\bar{n}\ln\bar{n}+\text{const}$. This expression for the entropy is not thermodynamically admissible because it lacks the property of extensivity (Lavenda, 1988). It can be made extensive by writing

$$S(\bar{n})=\bar{n}\left[\ln\left(\frac{C}{\bar{n}}\right)+1\right] \tag{46}$$

which leaves Gauss' principle invariant, since all terms involving the extensive quantity C cancel out. We will subsequently identify C as the number of "cells" in a given particle energy range $\Delta\varepsilon$.

On the strength of the second law, we equate the derivative of the expression of the statistical entropy (46) with the corresponding thermodynamic expression to obtain

$$\left(\frac{\partial S}{\partial \bar{n}}\right)_{V}=\ln\left(\frac{C}{\bar{n}}\right)=\beta(\varepsilon-\mu) \tag{47}$$

or, upon rearranging,

$$\bar{n}=C(\varepsilon)e^{-\beta(\varepsilon-\mu)} \tag{48}$$

where ε is the energy of the individual particles, or, in discrete form, the energy of a given cell, and μ is the chemical potential. Applying the second law in the form (47) means that the total energy is no longer an independent variable, for it depends on the average number of particles in the various cells.

Instead of the Massieu transform (39), the relevant thermodynamic potential is given by

$$\beta PV=S[\beta,V,\beta\mu]=S(\bar{n},V)-\left(\frac{\partial S}{\partial \bar{n}}\right)_{V}\bar{n} \tag{49}$$

where P is the "monochromatic" pressure or the pressure in the energy interval $\Delta\varepsilon$. The volume V plays the role of the mute extensive parameter in the Massieu transform (49), just as the number of particles n plays the same role in (39). In order not to confuse *n*, the total number of particles, with the fluctuating variable, we will denote it by *N*.

The total number of particles is obtained by summing (48) over regions of the phase volume or by requiring $\Delta\varepsilon$ to be small enough to allow summations over the different regions to be replaced by integration over ε. We then obtain

$$N = \int_0^\infty C(\varepsilon) e^{-\beta(\varepsilon-\mu)}\, d\varepsilon \tag{50}$$

which must be compatible with the expression we get on solving (33) for *N*, viz.,

$$N = \left(\frac{A}{\beta}\right)^m V e^{\beta\mu} \tag{51}$$

where we have used (36) and $\mathcal{G} = \mu N$. From this equivalence, it is seen that the integral is a gamma integral and

$$C(\varepsilon) = A^m \frac{\varepsilon^{m-1}}{\Gamma(m)} V \tag{52}$$

It is readily seen from the specific case of a monatomic gas with $2m = 3$ degrees of freedom and $A = 2\pi M/h^2$ that $C(\varepsilon)\Delta\varepsilon$ is the number of states of one particle whose energies lie between ε and $\varepsilon + \Delta\varepsilon$. Introducing (48) into the entropy expression (46) and integrating over all ε, we obtain the integrated entropy

$$\mathcal{S} = \beta U + N \ln\left[\left(\frac{A}{\beta}\right)^m \frac{V}{N}\right] + N \tag{53}$$

which is identical with (42) when the equipartition of energy (13) is introduced.

5. GENERALIZED STEFAN LAW

The law of equipartition of energy (13) can be combined with the equation of state of an ideal gas (29) to give a generalization of Stefan's law that was first proposed by Lord Rayleigh (1902) in his discussion on the pressure of vibrations. Eliminating β between (13) and (29) gives

$$\mathcal{P} = u/m \tag{54}$$

where the energy density $u = U/V$. *It is quite remarkable that the pressure-energy density relation* (54) *is valid whether or not equipartition of energy and the equation of state of an ideal gas hold.*

The Gibbs–Duhem relation for a system with a constant, or vanishing, chemical potential can be written as

$$\frac{d\mathcal{P}}{dT}=\frac{\mathcal{S}}{V}=\frac{u+\mathcal{P}}{T} \tag{55}$$

Introducing (54) and integrating gives

$$u \propto T^{m+1} \tag{56}$$

which is the generalization of Stefan's law that was first proposed by Lord Rayleigh (1902). For Stefan's law $u \propto T^4$, one has $m=3$, while for a monatomic gas, $\mathcal{P} \propto \frac{2}{3}u$, so that $m=\frac{3}{2}$ and Stefan's law is $u \propto T^{5/2}$.

We now show that the generalized Stefan law attributes the same physical significance to $2m$ as we have previously deduced from the chi-square distribution. In a quasistatic change, the heat added to the system in an infinitesimal transition to a neighboring equilibrium state is given by the first law

$$dQ=d(uV)+\mathcal{P}\,dV \tag{57}$$

Considering a quasistatistical adiabatic change and introducing the generalized equation of state (54), we have

$$\left(1+\frac{1}{m}\right)u\,dV+V\,du=0 \tag{58}$$

which integrates to give

$$uV^{1+1/m}=\text{const} \tag{59}$$

Introducing the generalized Stefan law (56) gives

$$TV^{1/m}=\text{const} \tag{60}$$

For an ideal gas we have

$$C_V\frac{dT}{T}+(C_P-C_V)\frac{dV}{V}=0 \tag{61}$$

in a quasistatic adiabatic change. Integrating (61), we obtain

$$TV^{\gamma-1}=\text{const} \tag{62}$$

where $\gamma=C_P/C_V$, the ratio of the specific heat at constant pressure to that at constant volume, and C_V was assumed to be independent of the temperature. Comparing (60) and (62), we conclude that

$$\gamma=1+\frac{1}{m} \tag{63}$$

Now according to kinetic theory, γ is given by (Boltzmann, 1964)

$$\gamma = \frac{\eta + 2}{\eta} \tag{64}$$

where η is the number of degrees of freedom of the molecule. A comparison of (63) and (64) shows that the numerical quantity appearing in the generalized Stefan law (56) is precisely $m = \frac{1}{2}\eta$, or one-half the number of degrees of freedom of the elementary constituents of the system. We have arrived at the same physical interpretation based on an entirely different approach leading to the derivation of the chi-square distribution governing the continuous distribution in the energy, (12).

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Public Education (MPI) and the National Research Council (CNR).

REFERENCES

Blackwell, D., and Girshick, M. A (1954). *Theory of Games and Statistical Decisions*, Wiley, New York, p. 319.

Blanc-Lapierre, A., and Tortat, A. (1956). In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. III, University of California Press, Berkeley, pp. 145–170.

Boltzmann, L. (1964). *Lectures on Gas Theory* (S. G. Brush, transl.), University of California Press, Berkeley, Part II, § 43ff.

Ehrenfest, P., and Ehrenfest, T. (1959). *The Conceptual Foundations of the Statistical Approach in Mechanics* (M. J. Moravcsik, transl.), Cornell University Press, Ithaca, New York.

Feller, W. (1968). *An Introduction to Probability Theory & Its Applications*, Vol. I, 3rd ed., Wiley, New York, p. 116.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics*, Yale University Press, New Haven, Connecticut.

Greene, R. F., and Callen, H. B. (1951). *Physical Review*, **83**, 1231.

Khinchin, A. I. (1949). *Mathematical Foundations of Statistical Mechanics* (G. Gamow, transl.), Dover, New York.

Lavenda, B. H. (1987). *International Journal of Theoretical Physics*, **26**, 1069–1084.

Lavenda, B. H. (1988). *International Journal of Theoretical Physics*, **27**, 1371–1381.

Lavenda, B. H., and Scherer, C. (1988). *Rivista del Nuovo Cimento* **11**(6).

Lord Rayleigh (1902). *Philosophical Magazine* **3**, 338–346.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York, pp. 373–376.