THE

PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 97, NO. 6 MARCH 15, 1955

Relation Between Canonical and Microcanonical Ensembles*

MELVIN LAX
Department of Physics, Syracuse University, Syracuse, New York
(Received July 20, 1954)

The equivalence of averages calculated in canonical and microcanonical ensembles is shown to depend on the validity of a steepest-descent approximation. It is demonstrated that the microcanonical and canonical procedures yield different values for the order parameter below the Curie temperature for spherical model dipole lattices.

WE propose to show that averages calculated in the microcanonical[1] ensemble reduce to corresponding values obtained in the canonical ensemble *providing* the microcanonical calculation can be evaluated by a steepest descent method. When this condition is not satisfied, we shall demonstrate by means of an example—the spherical model of a ferromagnet[2,3] that the canonical ensemble *can* yield incorrect results. This note was stimulated by the astute observation of Lewis and Wannier[4] that the integration in the complex plane in the Berlin-Kac spherical model calculation can apparently be avoided by using a canonical treatment of the spherical constraint. We say apparently because Lewis and Wannier[5] have since then discovered a discrepancy between their canonical treatment and the corresponding microcanonical treatment of Berlin and Kac in evaluating a fluctuation in the spherical constraint.

Consider a phase or configuration space described by the set of variables $\epsilon=\{\epsilon_1, \epsilon_2, \cdots \epsilon_N\}$ and an extensive phase function $\phi(\epsilon)$. By extensive we mean that the value of $\phi(\epsilon)$ is proportional to N, the size of the system An ensemble canonical in the energy $U(\epsilon)$ and in $\phi(\epsilon)$ has the partition function:

$$Q(t)=\int \exp[-\beta U(\epsilon)-t\phi(\epsilon)]d\epsilon, \tag{1}$$

where $\beta=(kT)^{-1}$, and t is the variable conjugate to $\phi(\epsilon)$. The condition that $\phi(\epsilon)$ possess the mean value KN leads to

$$\langle\phi(\epsilon)\rangle=-(\partial/\partial t)(\ln Q)=KN. \tag{2}$$

The mean value of any other observable $H(\epsilon)$ is given by $H(t_s)$, where

$$H(t)=[Q(t)]^{-1}\int H(\epsilon)\exp[-\beta U(\epsilon)-t\phi(\epsilon)]d\epsilon, \tag{3}$$

and t_s is the value of t determined by (2).

The corresponding partition function in an ensemble microcanonical with respect to ϕ is

$$Q=\int \exp(-\beta U)d\epsilon\delta(KN-\phi). \tag{4}$$

Using the usual integral representation for a delta function

$$\delta(KN-\phi)=(2\pi i)^{-1}\int_{-i\infty}^{i\infty} dt \exp[t(KN-\phi)] \tag{5}$$

we can express Q in terms of the canonical partition function

$$Q=(2\pi i)^{-1}\int_{-i\infty}^{i\infty} \exp[NKt+\ln Q(t)]dt. \tag{6}$$

---

* Supported in part by the Office of Naval Research and in part by the U. S. Air Force (Air Research and Development Command).

[1] An ensemble is said here to be microcanonical or canonical *with respect to* a given extensive variable according to whether that variable is constrained to take a fixed value, or merely required to have a given average value. (The usual classification is based on whether the *energy* is constrained or not. Ensembles in which variables in addition to the energy are canonical are usually referred to as grand canonical.)

[2] T. H. Berlin and M. Kac, Phys. Rev. **86**, 821 (1952).

[3] M. Lax, J. Chem. Phys. **20**, 1351 (1952), hereafter referred to as I.

[4] H. W. Lewis and G. H. Wannier, Phys. Rev. **88**, 682 (1952).

[5] H. W. Lewis and G. H. Wannier, Phys. Rev. **90**, 1131(E) (1953).

Thus the microcanonical partition function is a weighted superposition of canonical partition functions over various values of the intensive parameter t. If, however, $\phi(\epsilon)$ is a suitably chosen (i.e., "macroscopic") variable, we may expect from (1) that for large N, $Q(t)$ has the form $[Z(t)]^N$. The exponent in (6) then possesses only terms of order N and a steepest descent evaluation is permissible. The saddle point $t=t_s$ may be located by setting the derivative of the exponent in (6) equal to zero, *which leads identically to Eq. (2)!* Thus the major contribution to Q comes from canonical ensembles in a small interval around t_s, the canonical value of the intensive variable t.

The mean value of $H(\epsilon)$ in the microcanonical ensemble can be shown by a similar argument to be:

$$\langle H(\epsilon)\rangle=(2\pi iQ)^{-1}\int H(t)\exp[NKt+\ln Q(t)]dt, \quad (7)$$

or

$$\langle H(\epsilon)\rangle\simeq H(t_s). \quad (8)$$

Passage from (7) to (8) however requires (a) the existence of a saddle point which imposes requirements on the nature of $\phi(\epsilon)$, (b) that $\ln H(t)$ is not of order N, so that the saddle condition in (7) does not differ from that for (6) and (c) that $H(t)$ does not possess a singularity in the immediate neighborhood of t_s. Conditions (b) and (c) impose requirements in the nature of $H(\epsilon)$. All of these conditions must be met before (8) is valid, i.e., before the canonical and microcanonical ensembles yield the same mean value for $H(\epsilon)$.

It may now be of interest to illustrate these remarks by applying them to the spherical model of a dipole lattice, for which $\phi(\epsilon)=K\sum\epsilon_j^2$, where $K=n\mu^2/(2kT)$. Lewis and Wannier found[5] that the mean of the variable $H(\epsilon)=\sum\epsilon_j^4$ is not given correctly by the canonical approach. They point out, however, that the sum $\sum\epsilon_j^4$ determines the fluctuations in the constraint $\phi(\epsilon)$ and state that "discrepancies are particularly apt to occur in those averages which are connected with fluctuations in the assumed constraints." They say however that the canonical method is adequate for the derivation of "all thermodynamic properties."

We shall show that the last remark is not always valid by demonstrating that the order parameter (a thermodynamic variable proportional to the magnetization in the ferromagnetic case) and its moments are not given correctly by the canonical approach. According to I(7.4) the order parameter can be defined as

$$S(\epsilon)=N^{-\frac{1}{2}}|y_m|=|\sum\pm\epsilon_j|/N, \quad (9)$$

where plus signs are to be used in the ferromagnetic case, and alternating signs if the mode y_m with the largest eigenvalue λ_m (i.e., lowest energy) corresponds to antiferromagnetic order. The mean value of $[S(\epsilon)]^n$ in the canonical ensemble using a slight modification of I(7.5) is given by

$$S_n(t)=[NK(t-\lambda_m)]^{-n/2}\Gamma((n+1)/2)/\Gamma(1/2). \quad (10)$$

After the saddle value for t is inserted,[6] (10) becomes

$$S_n(t_s)=[1-(T/T_c)]^{n/2}2^n\Gamma((n+1)/2)/\Gamma(1/2). \quad (11)$$

It is clear from (10) that no choice of t_s will make $S_n(t_s)=[S_1(t_s)]^n$. This condition should be obeyed, however, in the limit $N\rightarrow\infty$ since the order parameter S_1 is a macroscopic thermodynamic variable below the Curie temperature (i.e., the relative fluctuations in S_1 are of order $N^{-\frac{1}{2}}$).

The above discrepancy arises because of the singularity in $S_n(t)$ at $t=\lambda_m$. Below the Curie temperature, the saddle point[6] t_s differs from λ_m only by terms of order $(1/N)$. Thus condition (c) is violated and the saddle point or canonical method is invalid. With the help of (6), (7), I(7.1), and I(7.6) the correct nth moment is given by

$$S_n=\frac{\int_C S_n(t)(t-\lambda_m)^{-\frac{1}{2}}\exp[NK(1-T/T_c)(t-\lambda_m)]dt}{\int_C (t-\lambda_m)^{-\frac{1}{2}}\exp[NK(1-T/T_c)(t-\lambda_m)]dt}, \quad (12)$$

where the contour C extends from $-\infty$ to λ_m below the real axis, counterclockwise around λ_m, and back to $-\infty$ above the real axis. With the help of the Hankel integral formula,[7] we obtain

$$S_n=[1-(T/T_c)]^{n/2}, \quad (13)$$

the correct result for the spherical model.

The spherical model need not be an adequate representation of a dipole lattice. But it forms a perfectly valid example for comparing the canonical and microcanonical procedures. We may conclude that the complex integrations required by the microcanonical procedure can only be avoided (using a canonical procedure) when these integrations are easy to perform by a saddle point method.

There is no guarantee, furthermore, that the canonical procedure is valid for all thermodynamic properties. However, once the canonical value $H(t)$ has been obtained for a given thermodynamic variable, it should be easy to verify whether conditions (b) and (c) for the validity of the canonical method are satisfied for the variable H.

[6] For the saddle condition see for example Eqs. (2.12) and (2.18) of M. Lax, Phys. Rev. **97**, 629 (1955).

[7] E. T. Copson, *Functions of a Complex Variable* (Oxford University Press, London, 1935), Sec. 9.6, p. 225.