On a Minimum Property of the Free Energy

R. PEIERLS
The University, Birmingham, England
(Received September 22, 1938)

IT is well known that the eigenfunction ψ_1 of the lowest state of any system has the property of making the integral

$$\int \psi^* \mathbf{H} \psi d\tau \qquad (1)$$

a minimum; the value of the integral is the corresponding eigenvalue E_1 of the Hamiltonian $\mathbf{H}$. It is also well known that this property leads to a powerful method of approximating, at least qualitatively, ψ_1 and E_1, by minimizing (1) among a restricted class of functions.

No similarly simple minimum property exists for the higher eigenvalues. One can obtain the nth eigenvalue by minimizing (1) with the subsidiary condition that ψ be orthogonal to $\psi_1, \psi_2, \cdots, \psi_{n-1}$. However, this procedure becomes progressively more cumbersome as n increases and besides it is of no use unless $\psi_1, \psi_2, \cdots, \psi_{n-1}$ are very exactly known.

For problems of thermal equilibrium one is often concerned with the free energy, rather than with the individual eigenvalues. It is the purpose of this note to draw attention to a simple minimum property of the free energy which may be considered as a generalization of the variation principle for the lowest eigenvalue.

The free energy has the following property: If $\varphi_1, \varphi_2, \cdots, \varphi_n, \cdots$ are an arbitrary set of orthogonal and normalized functions, and

$$H_{nn} = \int \varphi_n^* \mathbf{H} \varphi_n d\tau \qquad (2)$$

is the expectation value of the Hamiltonian for the nth of them, then for any temperature T the function

$$\bar{F} = -kT \log \bar{P} = -kT \log \sum_n e^{-H_{nn}/kT}, \qquad (3)$$

which *would* represent the free energy if the H_{nn} *were* the true eigenvalues, is higher than the true free energy[1]

$$F_0 = -kT \log P_0 = -kT \log \sum_n e^{-E_n/kT}. \qquad (4)$$

[1] For $T=0$ this theorem evidently reduces to the variation principle for the lowest eigenvalue E_1.

This is equivalent to saying that the partition function, as formed by means of the expectation values H_{nn}:

$$\bar{P} = \sum_n e^{-H_{nn}/kT} \qquad (5)$$

is less than the true partition function

$$P_0 = \sum_n e^{-E_n/kT}.$$

In this form our theorem is a special case of the more general statement that if $f(E)$ is a function with the properties

$$df/dE < 0, \quad d^2f/dE^2 > 0, \qquad (6)$$

the expression

$$\bar{f} = \sum_n f(H_{nn})$$

is less than

$$f_0 = \sum_n f(E_n).$$

(We shall always assume that the latter is finite.)

We prove this first for the partial sum consisting of the first N terms. This partial sum,

$$\bar{f}_N = \sum_{n=1}^{N} f(H_{nn}) \qquad (7)$$

is bounded, as it cannot exceed the value $N \cdot f(E_1)$, according to (6). There must therefore be a greatest value of $\bar{f}_N$ and we shall show that this cannot be assumed for any set of functions other than $\psi_1, \psi_2, \cdots, \psi_N$.

Suppose that the set $\varphi_1, \varphi_2, \cdots, \varphi_N$ yields the maximum value of $\bar{f}_N$ and suppose that these functions are not all eigenfunctions. Then they do not make the energy diagonal, and hence at least one of the nondiagonal elements, say H_{mn}, with $m \leqq N$, must differ from zero. We must then distinguish three cases,

(*a*) $n > N$,

(*b*) $n \leqq N, \quad H_{mm} \neq H_{nn}$,

(*c*) $n \leqq N, \quad H_{mm} = H_{nn}$.

Case a. In this case we do not spoil the orthogonality of our set if we replace φ_m by

$$\varphi_m' = \varphi_m + \epsilon\varphi_n \tag{8}$$

with infiitely small ϵ. This leads to replacing the expectation value H_{mm}, up to first order, by

$$H_{mm}' = H_{mm} + (\epsilon H_{mn} + \epsilon^* H_{nm})$$

and the corresponding change in $\bar{f}_N$ is:

$$f_N' - \bar{f}_N = (df/dE)_{H_{mm}}(\epsilon H_{mn} + \epsilon^* H_{nm}).$$

It is obvious that this can be made positive by suitably choosing the sign of ϵ, and this is in contradiction with the assumption that our original set gave the maximum possible value of $\bar{f}_m$.

Case b. Here both φ_m and φ_n occur among the first N functions, and we replace them by

$$\varphi_m' = \varphi_m + \epsilon\varphi_n, \quad \varphi_n' = \varphi_n - \epsilon^*\varphi_n. \tag{9}$$

This changes H_{mm} and H_{nn} into

$$\begin{aligned} H_{mm}' &= H_{mm} + (\epsilon H_{mn} + \epsilon^* H_{nm}); \\ H_{nn}' &= H_{nn} - (\epsilon H_{mn} + \epsilon^* H_{nm}), \end{aligned} \tag{10}$$

except for second-order terms, and hence the change in $\bar{f}_N$ is, to first order:

$$f_N' - \bar{f}_N = \{(df/dE)_{H_{mm}} - (df/dE)_{H_{nn}}\} \times (\epsilon H_{mn} + \epsilon^* H_{nm}).$$

Since because of (6) the condition $H_{mm} \neq H_{nn}$ implies

$$(df/dE)_{H_{mm}} - (df/dE)_{H_{nn}} \neq 0,$$

(11) can be made positive by suitably choosing ϵ, and we arrive at the same contradiction.

Case c. If $H_{mm} = H_{nn}$, our substitution (9) produces no first-order change in $\bar{f}_N$. Proceeding to terms of the second order (for which the change of normalization has to be taken into account) one easily verifies that because of the equality of the diagonal elements, (10) gives the change of the expectation values correctly to second order, and hence the change in f_N becomes

$$f_N' - \bar{f}_N = (d^2f/dE^2)_{H_{mm}}(\epsilon H_{mn} + \epsilon^* H_{nm})^2.$$

This, as the square of a real quantity, is always positive.

We have thus, in all three cases, arrived at a contradiction to the assumption that a set of functions which are not all eigenfunctions gave the maximum value of $\bar{f}_N$. Hence the maximum must belong to a set consisting of N eigenfunctions, and it will then obviously be reached for the first N of them, and hence be equal to f_{N0}.

Consider now the infinite sum $\bar{f}$. If, for a suitable set of functions, it could reach a value exceeding f_0, there would have to exist a number N such that

$$f_N > f_{N0},$$

since for $N \to \infty$, $\bar{f}_N \to \bar{f}$ and $f_{N0} \to f_0$. This, however, we have just shown to be impossible.

So far we have assumed our set of functions to be complete. Our theorem holds *a fortiori* for an incomplete orthogonal set, since an incomplete set can always be thought of as formed by omitting some of the functions from a complete set; and by omitting terms the value of the sum (7) is reduced.

If we insert for the general function f:

$$f(E) = e^{-E/kT},$$

we obtain the above theorem about the free energy, but by inserting other functions it is easy to see that the theorem also applies to the free energy of an assembly of independent particles satisfying Bose-Einstein or Fermi-Dirac statistics.

It is hoped that an illustration of the application of this theorem will soon be published by Mr. F. Hoyle.