Lower Bounds for the Helmholtz Function*

SIDNEY GOLDEN
Department of Chemistry, Brandeis University, Waltham, Massachusetts
(Received 1 October 1964)

A mathematical theorem is established for traces of products of bounded Hermitian and definite operators, **a** and **b**: $\mathrm{Tr}(\mathbf{ab})^{2^{p+1}} \leqslant \mathrm{Tr}(\mathbf{a}^2\mathbf{b}^2)^{2^p}$ for p a non-negative integer. This theorem is applied to the equilibrium partition function by exploiting an infinite-product representation of the exponential function of the sum of two operators. As a result, a set of inequalities is established which yields a set of upper bounds for the partition function. This result is invariant to the particle statistics of the system. A general argument yields the result that the classical Helmholtz free-energy function serves as a lower bound to the corresponding quantum result.

1. INTRODUCTION

THE existence of several variational principles which furnish upper bounds for the Helmholtz free-energy function is relatively well known.[1] As a result, it is possible to effect reasonably good approximations to the statistical thermodynamic behavior of systems in practical computational terms. Whatever the variational formulation may be, however, its effectiveness as a minimal principle can be enhanced considerably if there is available a means for determining lower bounds to the Helmholtz free energy function. It is to that end that the present paper is directed.

In the following section, a mathematical theorem relating to the traces of products of bounded, Hermitian, definite (i.e., non-negative) operators is established. This theorem is employed in the succeeding section to establish a set of lower bounds for the Helmholtz free-energy function. In particular, the inequalities are unaffected by symmetry restrictions relating to identical particles. An important result is the demonstration that the classically evaluated Helmholtz free-energy function serves as a lower bound to the correctly evaluated quantum mechanical quantity.

2. A MATHEMATICAL THEOREM

In the present section the following theorem will be proved.

For any two bounded,[2] Hermitian, definite operators **a** and **b**, which are otherwise arbitrary,

$$\mathrm{Tr}(\mathbf{ab})^{2^{p+1}} \leqslant \mathrm{Tr}(\mathbf{a}^2\mathbf{b}^2)^{2^p},\ p \text{ integral and } \geqslant 0. \quad (1)$$

Since, by hypothesis, the operators are definite there exists a (nonunique) square root of the operator **a**, say. Hence, since the trace is invariant to cyclic permutation of its factors, we may write

$$\mathrm{Tr}(\mathbf{ab})^2 = \mathrm{Tr}\,\mathbf{a}^{1/2}\mathbf{baba}^{1/2} \geqslant 0\,,$$

where we have chosen $\mathbf{a}^{1/2}$ to be Hermitian. Then, by the Cauchy-Schwartz inequality,[3]

$$\mathrm{Tr}(\mathbf{ab})^2 \leqslant \mathrm{Tr}(\mathbf{ab})(\mathbf{ab})^\dagger = \mathrm{Tr}(\mathbf{ab})(\mathbf{ba}) = \mathrm{Tr}\,\mathbf{a}^2\mathbf{b}^2\,,$$

so that the theorem evidently holds for $p=0$.

Now we assume that the theorem holds for all $0 \leqslant p \leqslant m-1$, $m \geqslant 1$ and show that it then holds for $p=m$, thereby establishing the theorem by induction.[4] To that end, consider that for integral m

$$\mathrm{Tr}(\mathbf{ab})^{2^{m+1}} \leqslant \mathrm{Tr}(\mathbf{ab})^{2^m}(\mathbf{ba})^{2^m} \quad (2)$$

by the Cauchy-Schwartz inequality. Let

$$\boldsymbol{\alpha}_n \equiv (\mathbf{ab})^{2^{(m-n)}}(\mathbf{ba})^{2^{(m-n)}} \equiv \boldsymbol{\alpha}_n^\dagger\,, \quad (3)$$

$$\boldsymbol{\beta}_n \equiv (\mathbf{ba})^{2^{(m-n)}}(\mathbf{ab})^{2^{(m-n)}} \equiv \boldsymbol{\beta}_n^\dagger\,, \quad (4)$$

for $m \geqslant n \geqslant 0$, n integral. We note that, for arbitrary non-negative integral N,

$$\mathrm{Tr}(\boldsymbol{\alpha}_n)^N \equiv \mathrm{Tr}(\boldsymbol{\beta}_n)^N\,, \quad (5)$$

while

$$\mathrm{Tr}(\boldsymbol{\alpha}_n)^N \equiv \mathrm{Tr}(\boldsymbol{\alpha}_{n+1}\boldsymbol{\beta}_{n+1})^N\,,\quad m>n\,. \quad (6)$$

In these terms we have from Eq. (2)

$$\mathrm{Tr}(\mathbf{ab})^{2^{m+1}} \leqslant \mathrm{Tr}(\boldsymbol{\alpha}_1\boldsymbol{\beta}_1) \leqslant \mathrm{Tr}(\boldsymbol{\alpha}_1)^2\,, \quad (7)$$

the last inequality resulting from a further application of the Cauchy-Schwartz inequality.

For $m=1$, it is evident that

$$\mathrm{Tr}(\boldsymbol{\alpha}_1)^2 = \mathrm{Tr}[(\mathbf{ab})(\mathbf{ba})(\mathbf{ab})(\mathbf{ba})] = \mathrm{Tr}(\mathbf{a}^2\mathbf{b}^2)^2\,,$$

in accord with the theorem. However, for $m \geqslant 2$, we have

$$\mathrm{Tr}(\boldsymbol{\alpha}_1)^2 = \mathrm{Tr}(\boldsymbol{\alpha}_2\boldsymbol{\beta}_2)^2\,,$$

* Supported in part by the U. S. Office of Naval Research.

[1] These seem first to have been examined systematically for classical systems by Gibbs. See, for example, J. W. Gibbs, *Collected Works, Statistical Mechanics* (Dover Publications, Inc., New York, 1961), Vol. II, Chap. 11. See also M. D. Girardeau, J. Chem. Phys. **40**, 899 (1964), where additional references to the subject may be found. See also W. Byers Brown, *ibid.* **41**, 2945 (1964).

[2] The restriction that the operators be bounded is necessary only to ensure the existence of the trace operation in what follows. Alternatively, the existence of the traces may be assumed instead of the explicit restriction upon the operators. See the end of the present section.

[3] When the trace operation involves discrete sums, it is known as the Cauchy inequality; when integration is implied, the inequality is the well known one due to Schwartz. See, for example, *NBS Applied Mathematics Series-55* (U. S. Government Printing Office, Washington, D. C., 1964), p. 11.

[4] We are employing here the so-called "Second Principle of Finite Induction." See, for example, G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (The Macmillan Company, New York, 1953), p. 13.

making use of Eq. (6). Since $\boldsymbol{\alpha}_2$ and $\boldsymbol{\beta}_2$ conform to the conditions of the theorem with $p=0$, we have

$$\operatorname{Tr}(\boldsymbol{\alpha}_2\boldsymbol{\beta}_2)^2 \leqslant \operatorname{Tr}(\boldsymbol{\alpha}_2)^2(\boldsymbol{\beta}_2)^2 \leqslant \operatorname{Tr}(\boldsymbol{\alpha}_2)^4, \qquad (8)$$

the last inequality resulting from a further application of the Cauchy-Schwartz inequality. Again if $m=2$, the theorem is satisfied explicitly; however, if $m\geqslant 3$, we have

$$\operatorname{Tr}(\boldsymbol{\alpha}_2)^4 = \operatorname{Tr}(\boldsymbol{\alpha}_3\boldsymbol{\beta}_3)^4,$$

in terms of which an application of the theorem for $p=1$ yields

$$\operatorname{Tr}(\mathbf{ab})^{2^{m+1}} \leqslant \operatorname{Tr}(\boldsymbol{\alpha}_3)^8.$$

Clearly, the procedure may be repeated under the stated assumptions until one obtains

$$\operatorname{Tr}(\mathbf{ab})^{2^{m+1}} \leqslant \operatorname{Tr}(\boldsymbol{\alpha}_m)^{2^m} = \operatorname{Tr}(\mathbf{a}^2\mathbf{b}^2)^{2^m}. \qquad (9)$$

The theorem stated in Eq. (1) now follows.

As an immediate consequence of Eq. (1) we can obtain

$$\operatorname{Tr}(\mathbf{ab})^{2^p} \leqslant \operatorname{Tr}[\mathbf{a}^{2^r}\mathbf{b}^{2^r}]^{2^{(p-r)}} \leqslant \operatorname{Tr}[\mathbf{a}^{2^q}\mathbf{b}^{2^q}]^{2^{(p-q)}}, \qquad (10)$$

$p\geqslant q\geqslant r\geqslant 0$. As a result, we see that the previous analysis applies if the right side of Eq. (10) exists for $q>1$. In such a circumstance any boundedness restrictions implicit upon **a** or **b** may be relaxed and replaced by a condition on the existence of the trace of the relevant product of these operators.

3. THE HELMHOLTZ FREE-ENERGY FUNCTION

The Helmholtz free-energy function for a system is defined by

$$F = -\Theta \ln \operatorname{Tr} e^{-\mathbf{H}/\Theta}, \qquad (11)$$

where $\Theta = kT$, k is Boltzmann's constant, T is the absolute temperature, and $\mathbf{H}$ is the Hamiltonian of the system in question. To exploit the mathematical theorem of the preceding section, we note that for any partition of the Hamiltonian

$$\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2 \qquad (12)$$

the exponential can be represented by[5]

$$e^{-H/\Theta} = \lim_{N\to\infty} [e^{-H_1/N\Theta} e^{-H_2/N\Theta}]^N. \qquad (13)$$

Clearly, an immediate identification of the factors in Eq. (13) may be made with the **a** and **b** operators of Eq. (1). As a result, we may transcribe Eq. (10) to yield

$$\begin{aligned}\operatorname{Tr}[e^{-\mathbf{H}_1/\Theta}e^{-\mathbf{H}_2/\Theta}] &\geqslant \operatorname{Tr}[e^{-\mathbf{H}_1/2\Theta}e^{-\mathbf{H}_2/2\Theta}]^2 \\ &\geqslant \operatorname{Tr}[e^{-\mathbf{H}_1/2^p\Theta}e^{-\mathbf{H}_2/2^p\Theta}]^{2^p} \qquad (14)\\ &\geqslant \operatorname{Tr}[e^{-\mathbf{H}_1/2^q\Theta}e^{-\mathbf{H}_2/2^q\Theta}]^{2^q} \\ &\geqslant \operatorname{Tr} e^{-\mathbf{H}/\Theta}, \quad p\leqslant q.\end{aligned}$$

[5] See, for example, S. T. Butler and M. H. Friedman, Phys. Rev. **98**, 287 (1955). A proof is given by S. Golden, Phys. Rev. **107**, 1283 (1957).

Then, if we define

$$F_q \equiv -\Theta \ln \operatorname{Tr}[e^{-\mathbf{H}_1/2^q\Theta}e^{-\mathbf{H}_2/2^q\Theta}]^{2^q}, \qquad (15)$$

we have

$$F\geqslant F_q \geqslant F_p, \quad p\leqslant q, \qquad (16)$$

since the logarithm is a monotonically increasing function of its argument. For any finite value of N in Eq. (13) we are able to take the resulting expression as an approximation to the canonical distribution formula. Regardless of the mode of partition of the Hamiltonian,[6] such an approximation yields one for the partition function which results in a value for the Helmholtz function which is not greater than the correct value.

It is evident that when the physical system consists of identical particles the trace operation may be restricted to bases of complete, orthonormal functions which are either symmetric or antisymmetric with respect to exchange of identical particles. The relations which have been derived are unaltered thereby. Hence Eq. (16) applies, as well, to systems satisfying Bose-Einstein or Fermi-Dirac statistics.[6]

When $q=0$ in Eq. (15), the resulting form of the partition function corresponds to what may be termed *pseudoclassical*.[7] In particular, when $\mathbf{H}_1$ is taken to be the kinetic-energy operator for the system, with $\mathbf{H}_2$ the potential energy of the system, the resulting partition function can be evaluated in such a basis (i.e., plane waves with appropriate boundary conditions) that provides for a separation into the usual factors for translation and configuration. Such a partition function yields a Helmholtz function which may be identified with the classical evaluation of that quantity.[8] Designating the latter by F_{cl}, we have the important result that

$$F\geqslant F_{\mathrm{cl}}, \qquad (17)$$

or that the classical Helmholtz function[8] provides a lower bound to the correct quantum-mechanical Helmholtz function.[9]

[6] One must ensure that the partition of the Hamiltonian does not introduce divergences which will yield unbounded operators as factors. Likewise, the separate factors may be supposed to have symmetries identical with those of the total system.

[7] The phraseology *pseudoclassical* used here is meant to emphasize that a choice has been made of one of the operators, say $\mathbf{H}_1$, which can be regarded as the "unperturbed" Hamiltonian of the system. In the basis which diagonalizes the latter, the energy of a state is then a sum of the unperturbed energy and the remainder. See, for example, the discussion by the present author in Ref. 5.

[8] Again we note that what has been termed here as classical must be understood as being restricted. Thus, the choice of the kinetic-energy operator for $\mathbf{H}_1$ still requires a proper quantum elevation of partition function for translation. Of course, for large finite containers one gets the usual classical statistical mechanical result obtained by integration over the appropriately measured momentum space.

[9] This result is known as an approximation. See, for example, L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1958), p. 100.