Proof of the Linked-Cluster Expansion in Quantum Statistical Mechanics*

DAVID J. THOULESS†
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received April 6, 1959)

In order to go over from a perturbation expansion of the grand partition function (the unlinked-cluster expansion) to an expansion of the thermodynamic potential (the linked-cluster expansion) in powers of the interaction, it is necessary to treat carefully those terms in which creation (or annihilation) operators for the same state occur twice or more. The unlinked- and linked-cluster expansions for a system of fermions are here shown to be equivalent by a direct comparison of the terms which occur in each. The relation between the two expansions is illustrated by the example of a system of fermions interacting only with a single-particle potential.

THE expansion of the grand partition function for a system of interacting fermions obtained by Glassgold, Heckrotte, and Watson[1] can be expressed as a sum over graphs which have the following structure and properties:

(a) All lines are continuous and go from right to left of the graph. Therefore the number of lines crossing a vertical section of the graph is constant.

(b) Interactions are denoted by dots on the lines. For a two-body interaction, a dot will occur at the intersection of two lines.

(c) Each line is labelled as a single-particle eigenstate of the unperturbed Hamiltonian; the label can only change at an interaction point. Generally, the spin and momentum will be specified.

(d) The external lines on the right of the graph are the same as the external lines on the left. We will refer to the external lines on the right and left as the initial and final states, and to the lines crossing a vertical section between two interactions as an intermediate state. The number of intermediate states is then one less than the number of interactions in a graph.

(e) No two lines are the same in any intermediate state, nor in the initial or final states: *all states satisfy the exclusion principle.*

(f) Each line must have at least one interaction along its length.

No other restriction is made about the connectedness of a graph, and Figs. 1(a) and 1(b) are both possible graphs. Each interaction in a graph contributes a matrix element of the interaction as a factor, and each intermediate state contributes an energy denominator determined by the difference between the unperturbed energies of the initial and intermediate states. It is convenient to regard a line as continuing through an interaction above the other line if it was above it before the interaction, and below it if it was originally below. Because of the antisymmetry of the particles, there is a factor $(-1)^p$, where p is the number of permutations required to go from the order of the final states to which the initial states are connected, back to the order of the initial states. This factor is $+1$ for both parts of Fig. 1. It is, however, the statistical factors, that come from taking the trace of $\exp(\alpha N-\beta H_0)$, with which we are mainly concerned.

A particular graph labelled λ has lines $k_1, k_2, \cdots, k_n$ in its initial state, and further lines $s_1, s_2, \cdots, s_m$ in its intermediate states. The graph makes a contribution to a particular diagonal element $\langle\psi|\exp(\alpha N-\beta H)|\psi\rangle$ if and only if all the states $k_1, k_2, \cdots, k_n$ are included in ψ and none of the states $s_1, s_2, \cdots, s_m$ are included in ψ. We can write the contribution as

$$\langle\psi|\exp(\alpha N-\beta H_0)|\psi\rangle \times a_\lambda(k_1, k_2, \cdots, k_n; s_1, s_2, , \cdots s_m), \quad (1)$$

if this condition is satisfied. The contribution of the graph λ to the grand partition function is then

$$\vartheta_0 a_\lambda(k_1\cdots k_n; s_1\cdots s_m) f_{k_1} f_{k_2}\cdots f_{k_n} \times(1-f_{s_1})(1-f_{s_2})\cdots(1-f_{s_m}), \quad (2)$$

where

$$\vartheta_0=\mathrm{Tr}[\exp(\alpha N-\beta H_0)], \quad (3)$$

$$f_k=[\exp(-\alpha+\beta\epsilon_k)+1]^{-1}. \quad (4)$$

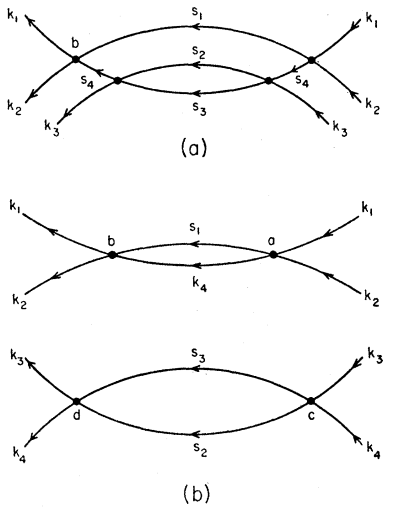

FIG. 1. Some typical graphs in Method I.

* This work was done under the auspices of the U. S. Atomic Energy Commission.
† Present address: Department of Mathematical Physics, University of Birmingham, Birmingham, England.
[1] Glassgold, Heckrotte, and Watson, Phys. Rev. **115**, 1374 (1959).

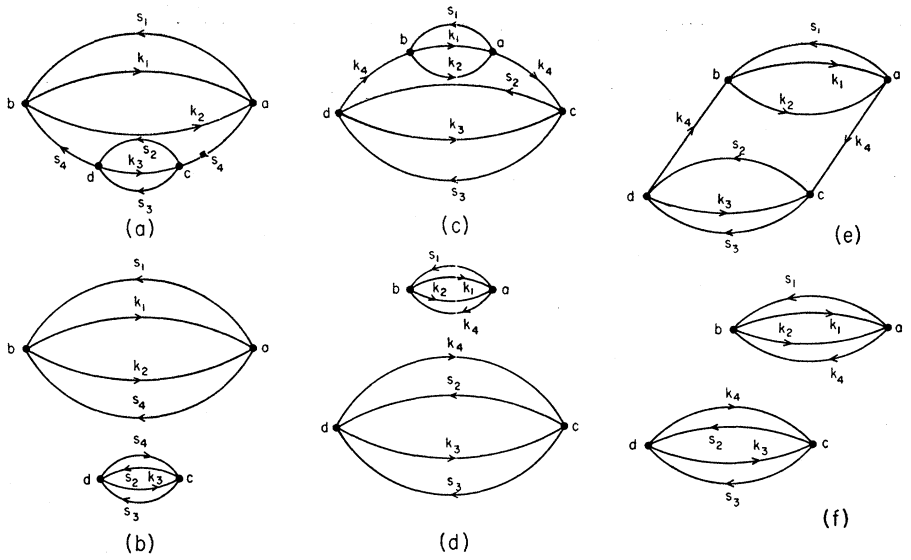

FIG. 2. Some typical graphs in Method II.

Here H_0 is the unperturbed Hamiltonian, given by the sum of the single-particle energies ϵ_k.

This way of evaluating the partition function will be referred to as Method I. It must be noticed that, if the same line occurs twice as an intermediate state, like s_4 in Fig. 1(a), the factor $1-f_{s4}$ only occurs once. Similarly, if a line occurs both as an initial state and as an intermediate one, like k_4 in Fig. 1(b), we have just a factor f_{k4}. Both these graphs can be seen to have five degrees of freedom and four interactions, so that each gives a contribution proportional to the extent of the system. According to the analysis of Hugenholtz,[2] the contribution of a graph is proportional to the volume raised to a power equal to the number of unlinked clusters less the number of additional constraints. Graphs with two clusters and one constraint, like Fig. 1(b), contribute to the same order as graphs with one cluster and no constraint, like Fig. 1(a). The grand partition function cannot yet be expressed as an exponential of contributions from linked clusters, because a graph like Fig. 1(b) gives a contribution which is not simply the product of factors from its two linked clusters. Firstly, there is no factor $1-f_{k4}$ from the intermediate state k_4, and, secondly, the points a and b are constrained to lie between c and d by the exclusion principle.

For this reason, we want to show that this expansion of the grand partition function is equivalent to a slightly different expansion, which is essentially the Laplace transform of the expansion obtained originally by Matsubara[3] and studied in more detail by Thouless[4] and by Bloch and De Dominicis.[5] This expansion, which we call Method II, has the advantage that it can be directly expressed as an exponential.

We therefore want to show that the results of Method I can be obtained by summing over graphs with different structures and properties, given below:

(a′) All lines are continuous, but can go from right to left or from left to right, changing direction only at an interaction point on the line; we call them particle lines and hole lines, respectively. There are no external lines, and the lines are continuous, so that the number of particle lines in an intermediate state is equal to the number of hole lines.

(d′) There are no lines in the initial or final state. The energy of an intermediate state is the energy of its particles less the energy of its holes. This is the quantity that determines the energy denominator.

(e′) A hole or particle line can occur any number of times in an intermediate state, and the same line can occur both as a particle and a hole in one state. *There is no exclusion principle.*

(f′) A line may join a point to itself, in which case it counts as a hole line. It can only do this if there is an interaction at the point.

The conditions (b) and (c) remain the same. Because of the antisymmetry of the matrix elements of the interaction, we now get a factor -1 for each closed loop of fermion lines in the graph. The matrix elements and energy denominators come in the same way as they did in Method I, and the only essential difference is that there is now a factor $-f_k$ for each hole line, and a factor $1-f_k$ for each particle line in the graph, irrespective of how many times the same line is repeated. For example, there is a factor $(1-f_{s4})^2$ in the contribution from Fig. 2(a), and a factor $-f_{s4}(1-f_{s4})$ in the contribution from Fig. 2(b). The grand partition function is still given by the sum of all graphs, connected or otherwise.

It is easy to show that the contributions of those graphs that do not have two different lines representing the same state are the same for the two methods. The hole states in Method II are equivalent to the initial

[2] N. M. Hugenholtz, Physica **23**, 481 (1957).
[3] T. Matsubara, Progr. Theoret. Phys. (Kyoto) **14**, 351 (1955).
[4] D. J. Thouless, Phys. Rev. **107**, 1162 (1957).
[5] C. Bloch and C. De Dominicis, Nuclear Phys. **7**, 459 (1958).

states in Method I, and the statistical factors are the same except for a factor -1 for each hole line. The factors -1 for each hole and -1 for each closed loop combine to give the factor $(-1)^p$ in Method I, where p is the number of permutations required to change the ordering of the final states to the ordering of the initial states to which each one is connected.

We therefore consider graphs in which one line, referring to the state k, occurs in several places in the same graph. Using Method II, we consider a particular graph. There are some points in the graph which the directed line k enters, which we call entry points, and an equal number from which the directed line k leaves, which we call exit points. In each part of Fig. 2, the points a and d are exit points for the line s_4 or k_4, and the points b and c are entry points for that line. It may happen that an exit point and an entry point coincide, in which case we displace the entry point a little to the right in order to keep the condition that a point may only be joined to itself by a hole line.

Now we observe that the energy denominators and the matrix elements of the interaction which are involved do not depend on the way in which the n entry points and n exit points are connected by the n lines k. The $n!$ ways of joining these points by the lines k differ by the number of closed loops and by the number of factors $-f_k$ and $1-f_k$ which occur. Figures 2(a) and 2(b), or Figs. 2(c) and 2(d), or Figs. 2(e) and 2(f), are graphs which differ only by which entry point is connected to which exit point. We can get these $n!$ different graphs from the initial one by making elementary permutations successively. An elementary permutation is obtained by taking two of the lines k, connecting the points a to c and the points d to b, and then making them connect the points a to b and d to c; we denote this permutation by (cb) or (ad). Each elementary permutation changes the number of closed loops by ± 1. This permutation is the one that carries Fig. 2(a) into Fig. 2(b), and it can be seen that Fig. 2(a) has three closed loops, while Fig. 2(b) has four.

Now consider the way in which the $2n$ points are ordered, from right to left of the graph. Suppose that two entry points or two exit points c and b come successively, as they do in Fig. 2(e). Then we can make the elementary permutation (cb), which will not alter the direction of any line, and so will change the contribution of the graph by a factor -1 because of the change in the number of closed loops. This is the permutation which carries Fig. 2(e) into Fig. 2(f), and it can be seen that the contributions from these graphs differ merely by a factor -1. If we have two exit points or two entry points adjacent, then the total contribution of all the $n!$ arrangements is zero. This is in agreement with the results of Method I, since the corresponding graph would violate the exclusion principle just to the right of the two points if they are entry points, and just to the left of the two points if they are exit points.

It remains to consider the case in which entry points and exit points for the line k come alternately. In this case there is one and only one corresponding graph in Method I for which the exclusion principle is not violated. If the first point on the right is an entry point, the corresponding graph has k as an external line, and gives a factor f_k. Figures 2(c) and 2(d) are examples of this, and Fig. 1(b) is the corresponding graph. If the first point is an exit point, the corresponding graph has k as an internal line, and gives a factor $1-f_k$. Figures 2(a) and 2(b) are examples of this, and Fig. 1(a) is the corresponding graph.

The statistical factor which we get by summing over all $n!$ permutations is called $\pm F_n^{(+)}$ if the first point is an exit point and $\pm F_n^{(-)}$ if the first point is an entry point. The sign is determined by the sign of the arrangement which is allowed in Method I (which depends on the number of permutations needed to make the final state correspond to the initial state). Explicitly, we have

$$\begin{aligned} F_n^{(+)} &= (1-f_k)^n + \cdots, \\ F_n^{(-)} &= -f_k(1-f_k)^{n-1} + \cdots. \end{aligned} \tag{5}$$

Now, if the point 1 (on the right) is connected to the point 2 adjacent to it, and we sum over the $(n-1)!$ permutations for the other points, we get a contribution to $F_n^{(+)}$ of $(1-f_k)F_{n-1}^{(+)}$. If the point 1 is connected to the point $2n$, on the left, we get a contribution to $F_n^{(+)}$ of $-(1-f_k)F_{n-1}^{(-)}$. If the point 1 is connected to some other entry point, the summation over the $(n-1)!$ permutations for the remaining points gives zero, since there are two adjacent exit points in the remainder. Therefore, we obtain

$$F_n^{(+)} = (1-f_k)[F_{n-1}^{(+)} - F_{n-1}^{(-)}]. \tag{6}$$

In the same way we get

$$F_n^{(-)} = -f_k[F_{n-1}^{(+)} - F_{n-1}^{(-)}], \tag{7}$$

and so we have

$$F_n^{(+)} - F_n^{(-)} = F_{n-1}^{(+)} - F_{n-1}^{(-)}. \tag{8}$$

Considering the case $n=1$, we find that the expression in Eq. (8) is equal to one, and so we have

$$\begin{aligned} F_n^{(+)} &= 1-f_k, \\ F_n^{(-)} &= -f_k. \end{aligned} \tag{9}$$

Therefore we have shown that, by summing over the $n!$ permutations, we get just the result found by Method I. The proof can easily be extended to the case where more than one line occurs in several different parts of the graph, and so we have shown that Method I and Method II are equivalent. Figures 2(a) and 2(b) add to give Fig. 1(a), because we have $(1-f_{s4})^2 + f_{s4}(1-f_{s4}) = 1-f_{s4}$. In the same way, Figs. 2(c) and 2(d) add to give Fig. 1(b), since we have $f_{k4}^2 + f_{k4}(1-f_{k4}) = f_{k4}$.

If the grand partition function is expanded by Method II, it is easy to find its logarithm. If a graph

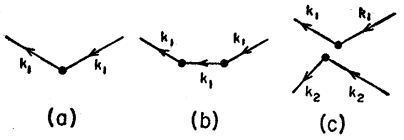

FIG. 3. Some graphs in the expansion of the grand partition function with a one-particle potential.

consists of several clusters, its contribution is equal to the product of the contributions of each cluster, if we take all possible orderings of the vertices of different clusters. This means that in order to get the product we would have to add together Figs. 2(b), 2(d), and 2(f), and three other graphs not shown. Therefore the thermodynamic potential is given by the sum of all linked graphs, evaluated with the statistical factors of Method II. A linked graph like Fig. 2(e) has no counterpart in Method I.

As a simple example of the use of the linked-cluster expansion, we take the case of particles interacting only with a diagonal one-particle potential, whose matrix elements are b_k. This was also considered by Glassgold, Heckrotte, and Watson,[1] and they showed that the partition function is exactly given by a sum of graphs like those in Fig. 3. We will take the exact answer, and write down the first four terms of its expansion in powers of b_k. We have

$$\begin{aligned}\Omega-\Omega_0&=-\beta^{-1}\textstyle\sum_k \log\{1+f_k[\exp(-\beta b_k)-1]\}\\&=-\beta^{-1}\textstyle\sum_k\{-\beta b_k f_k+\frac{1}{2}\beta^2 b_k^2(f_k-f_k^2)\\&\quad-\tfrac{1}{6}\beta^3 b_k^3(f_k-3f_k^2+2f_k^3)\\&\quad+(1/24)\beta^4 b_k^4(f_k-7f_k^2+12f_k^3-6f_k^4)-\cdots\}\\&=-\beta^{-1}\textstyle\sum_k\{-\beta b_k[f_k]+\frac{1}{2}\beta^2 b_k^2[f_k(1-f_k)]\\&\quad-\tfrac{1}{6}\beta^3 b_k^3[f_k(1-f_k)^2-f_k^2(1-f_k)]\\&\quad+(1/24)\beta^4 b_k^4[f_k(1-f_k)^3\\&\qquad-4f_k^2(1-f_k)^2+f_k^3(1-f_k)]-\cdots\}.\end{aligned}\tag{10}$$

The graphs which denote the terms written down here are all shown in Fig. 4. All the lines in these connected graphs represent the state k. It can be seen that the graphs of Fig. 4 are very different from those of Fig. 3.

Method I gives the expansion of the grand partition function which arises naturally in perturbation theory, but does not lead directly to an expansion of the thermodynamic potential. We have shown that it is equivalent to Method II without using the dependence of various terms on the total volume, so that the proof is good even for a finite system. Method II differs from Method I in its neglect of the exclusion principle and in the statistical factors used, and leads naturally to an expansion of the thermodynamic potential as a sum of linked clusters. In the limit of zero temperature, all the statistical factors are one or zero, and so the two methods differ only in the treatment of the exclusion principle. Thus we have here a generalization of the

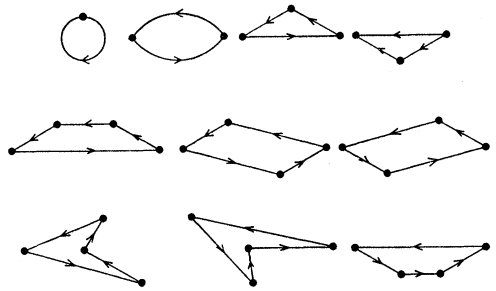

FIG. 4. Graphs that represent the terms of Eq. (10) up to fourth order in the potential.

theorem that the exclusion principle need not be used for intermediate states in the calculation of the ground-state energy by perturbation theory.[6] The difference between the two methods will be very important at nonzero temperatures for the calculation of the expectation value of a diagonal operator, or if some self-consistent potential is introduced.

ACKNOWLEDGMENTS

I wish to thank Dr. A. E. Glassgold and Professor K. M. Watson for several helpful discussions of this matter.

[6] J. Goldstone, Proc. Roy. Soc. (London) **A239**, 267 (1957).