Title
PERTURBATION THEORY IN STATISTICAL MECHANICS AND THE THEORY OF SUPERCONDUCTIVITY

Author
Thouless, David J.

Publication Date
1959-08-31

UCRL-8884

UNIVERSITY OF CALIFORNIA

Ernest O. Lawrence Radiation Laboratory

BERKELEY, CALIFORNIA

DISCLAIMER

UCRL-8884

UNIVERSITY OF CALIFORNIA
Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

PERTURBATION THEORY IN STATISTICAL MECHANICS AND THE THEORY OF SUPERCONDUCTIVITY

David J. Thouless

August 31, 1959

Printed for the U. S. Atomic Energy Commission

UCRL-8884

PERTURBATION THEORY IN STATISTICAL MECHANICS AND THE THEORY OF SUPERCONDUCTIVITY*

David J. Thouless

Lawrence Radiation Laboratory
University of California
Berkeley, California

August 31, 1959

ABSTRACT

The connection between formal perturbation theory and the modern theory of superconductivity is investigated. It is found that the condition for ladder diagrams to give a convergent sum is identical with the condition for the temperature to be above the critical temperature. The effect of the residual terms of the Hamiltonian is investigated and found to be small. They give rise to a correlation between electrons in the normal state and to a $|T - T_C|^{-1/2}$ singularity in the specific heat, but with a very small coefficient in both the normal and superconducting states. It is found that, below the critical temperature, most of the divergence is removed by using the BCS Hamiltonian as the unperturbed Hamiltonian, but that ladder diagrams with momentum exactly zero still diverge. The convergence of the ladder diagrams is suggested as a criterion which the BCS solution must satisfy, and this criterion is used to investigate some more complicated interactions. It is found that there is an interaction for which pairing of particles with opposite spin or with the same spin is not possible, and a more complicated trial wave function must be used.

* This work was performed under the auspices of the U.S. Atomic Energy Commission.

PERTURBATION THEORY IN STATISTICAL MECHANICS AND THE THEORY OF SUPERCONDUCTIVITY

David J. Thouless

Lawrence Radiation Laboratory
University of California
Berkeley, California

August 31, 1959

I. INTRODUCTION

In the past few years, powerful methods have been developed for treating the statistical mechanics of quantum systems.[1,2,3] These methods are based on the fact that the grand canonical partition function is

$$\mathcal{Z} = \mathrm{Tr}\,\psi = \mathrm{Tr}\,\exp(\alpha N - \beta H)\ , \tag{1}$$

and the distribution function ψ satisfies the Bloch equation

$$\frac{\partial \psi}{\partial \beta} = - H \psi\ , \tag{2}$$

since the Hamiltonian H commutes with the number operator N. Here β is $(kT)^{-1}$, where k is the Boltzmann constant and T is the temperature. This equation looks like the time-dependent Schrodinger equation, and the whole apparatus of formal perturbation theory--originally developed for studying field theories, but later used for a more closely related problem, the determination of the ground-state energy of a many-fermion system[4]--can be taken over with a few modifications. The main differences are that we assume the solution known for $\beta = 0$, infinite temperature, instead of for $t = -\infty$, before the interaction was switched on, and that the derivative on the left of Eq. (2) is not multiplied by i as it would be in the Schrodinger equation. A particularly important feature of the methods is that cumbrous expressions in the perturbation expansion can be represented by comparatively simple Feynman diagrams, and this notation allows some otherwise complicated formal manipulations to be carried out quite easily.

The theory of superconductivity developed by Bardeen, Cooper, and Schrieffer[5] (which is referred to throughout this paper as BCS) seems to account for most of the phenomena observed with superconductors. This theory is based on the discovery by Cooper that, if we have an extended system of n electrons (density kept constant as n varies) which interact by predominantly attractive forces, there is a wave function ϕ_C with a lower energy than the wave function ϕ_0 for a degenerate Fermi gas.[6] The wave function ϕ_C differs from ϕ_0 by the coherent excitation of pairs of particles, Cooper pairs. The expectation value of H for ϕ_C differs from the expectation for ϕ_0 by an amount which varies as $\exp(-1/g)$, where g is the strength of the attractive interaction, so that no expansion in powers of g can give ϕ_C. The difference between $(\phi_C, H\phi_C)$ and $(\phi_0, H\phi_0)$ is almost entirely due to the interaction between particles of exactly opposite momentum and spin, and the rest of the interaction could be thrown away without altering Cooper's result. The interaction terms in the Hamiltonian which are significant are a fraction n^{-1} of the total (each electron interacts with only one other instead of with all others), and any finite order of perturbation theory, starting with ϕ_0 as the unperturbed wave function, gives a contribution to the energy which is, at best, independent of n. Cooper's result is, however, that there is a contribution to the energy proportional to n, and this result, being based on a variational argument, cannot be doubted.

The unusual properties of this wave function ϕ_0 are basic in the BCS theory of superconductivity. A variational principle can also be used to derive the free energy at finite temperatures, and very good agreement with the qualitative thermodynamic properties of superconductors is found. The excitation spectrum obtained by considering a simple class of excitations from the ground state has the important property that the first excited state is separated from the ground state by an energy gap independent of n ; this is essential for an explanation of the stability of supercurrents. A different approach, used by Anderson, based on the random phase approximation, shows that there are other excitations, longitudinal sound waves, which are not separated by an energy gap unless the Coulomb repulsion is taken into account.[7] The problem here is that there is no variational principle to give a rigorous proof of the existence of a gap.

Perturbation theory is used to examine certain important questions in the BCS theory. The question of whether ϕ_C is an exact solution of the Schrodinger equation when there is interaction only between particles of opposite spin and momentum has been discussed by several authors. Bogoliubov has demonstrated that every term in the perturbation series for the energy (using ϕ_C as the unperturbed wave function) remains constant as n increases,[8] and he interprets this as showing that the solution is asymptotically exact in the limit of large n . This conclusion is by no means certain, since any contribution to the energy varies as $\exp(-1/g)$ gives zero in all orders of perturbation theory. The selection of a proportion n^{-1} of the terms in the interaction makes the application of perturbation theory uncertain. A concrete illustration of this is that Bogoliubov, Zubarev, and Tserkovnikov[9] find, by applying perturbation theory to the Bloch equation, that both the BCS free energy and the free energy for a noninteracting Fermi gas are exact solutions in the limit of large n , for temperatures below the critical temperature. They interpret this as meaning that the system can be either in a "superconducting" or in a "normal" state, with the normal state metastable below the critical temperature. It has been shown by the author that, in the strong coupling limit of this model, the normal state is certainly not metastable below the critical temperature,[10] and this interpretation of the apparent existence of two solutions for the equations for the partition function is open to doubt.

Perturbation theory has also been used to calculate the effect of the "residual terms" in the Hamiltonian, the terms describing the interaction of a particle with any particle other than the one with exactly opposite spin and momentum. The small size of the low-order terms in this perturbation series is not adequate justification of the neglect of these terms, since the same peculiar features of the problem which led to the existence of Cooper pairs may lead to more complicated wave functions, with lower energies than ϕ_C , differing from ϕ_0 by the coherent excitation of groups of four or more particles. A concrete example of this is given in this paper, although, fortunately, the interaction assumed does not seem to occur in metals. It is possible for there to be competition between the formation of Cooper pairs with spin zero and with spin one. As the ratio of the interactions leading to spin-one pairs and to spin-zero pairs varies, there may be an intermediate region in which coherent "quartets" are formed. Furthermore, it has been suggested by Heine and Pippard that, if the interaction between particles with momentum not exactly opposite is taken into account properly, it may be possible to explain the paramagnetism of small samples of superconductors which seems to be observed even at very low temperatures.[11] This conjecture is not supported by the work described here.

Anderson's approach to such problems seems more satisfactory than the perturbation methods, but it has not yet answered many of the questions raised here. An attempt has been made by Prange to use modern perturbation theory in this problem, and he concludes that there is no gap between the ground state and the first excited state;[12] this would probably mean that the actual phenomenon of superconductivity remains unexplained. The difficulty seems to be that it is not known how the simple BCS theory can be expressed in the language of formal perturbation theory. It is widely believed that there is an infinite class of terms in the perturbation series which can be formally summed to give the BCS results, but it is not known what this class of terms is.

For these varied reasons it seems desirable to examine the BCS theory by using the language of formal perturbation theory, and this is the principal aim of this paper. The clue to the connection between BCS theory and formal perturbation theory was discovered in the course of work on the foundations of Brueckner's theory of the energy of nuclear matter.[13] It was found that Brueckner's equation for the effective interaction between two particles has divergences if the force is on the whole attractive. These divergences occur for two particles with almost opposite momentum and with energies close to the Fermi energy, and its occurrence does not depend on the strength of the interaction. The weaker the interaction is, the nearer to the Fermi surface the particles have to be, and the closer to zero their total momentum has to be.

The terms in zero-temperature perturbation theory which Brueckner's equation takes into account are represented by "ladder diagrams"; in these, the two particles of a pair scatter each other any number of times in such a way that their intermediate states are always outside the Fermi sea. The "ladder diagrams" used in this paper are an obvious generalization of the Brueckner type of ladder diagrams, in which states below the Fermi surface are given the same weight as states above the Fermi surface. This is clearly necessary in any theory of the behavior of particles very close to the Fermi surface. At zero temperature the divergence of the expansion is still there, for the same reason as in the Brueckner theory. At finite temperature the expansion of the thermodynamic potential (minus pressure times volume, equal to -kT log $\mathcal{Z}$), which is the analog of the expansion of energy at zero temperature, may converge if only the ladder diagrams are taken into account. The condition for convergence is found to be that the temperature should be greater than the greatest temperature for which a solution of the BCS variational problem is possible. The equation derived for the critical temperature by Bogoliubov, Zubarev, and Tserkovnikov,[9] is identical with the equation derived and the temperature above which the ladder diagrams converge will therefore be referred to as the critical temperature. This is the main result of Sec. II. The sum of the ladder diagrams is similar in form to the sum of ring diagrams derived by Montroll and Ward,[2] and their result is rederived in a slightly simpler form in Appendix A to show the relation between the two sums.

Each ladder diagram is characterized by its total momentum, the sum of the momenta of the two particles involved. It is the sum of those ladders which have total momentum zero which diverges at the critical temperature. The sum of those ladder diagrams which have a particular momentum 2K not equal to zero is convergent even at the critical temperature, but this sum goes to infinity like -log K as K approaches zero. In Sec. III the properties of the system in the "normal state" just above the critical temperature are examined in the ladder approximation. It is found that the ladder diagrams with K close to zero could produce some interesting physical properties. The most striking result is that the specific heat should behave like $(T - T_C)^{-1/2}$, which is similar to the behavior of specific heats near a critical point at which a second-order transition becomes first-order, in the Landau theory.[14] All the effects discussed seem to be far too small for observation because of the very small ratio of the critical temperature to the Fermi temperature of the electrons, and the properties discussed are of merely theoretical interest. Another result of this nature is that there is a correlation between electrons, of the same nature as the correlation in the superconducting state of the BCS theory, but this correlation (either in the BCS theory or here) seems to be too small to observe.

In Sec. IV an attempt is made to apply similar methods to the superconducting state. The perturbation series diverges below the critical temperature in the ladder approximation, and there is always a value of the total momentum $2K$ for which the sum of ladders with that momentum is infinite, so that it does not seem reasonable to try to make an analytic continuation. Instead, we make a canonical transformation of the Bogoliubov type, and <u>then</u> apply perturbation theory in the ladder approximation. This is just using perturbation theory to estimate the effect of the residual terms of the Hamiltonian, but it is a consistent approximation. The ladder diagrams give the divergence in the perturbation series which indicates the phase transition, and we want to find if the canonical transformation can remove this divergence.

After making the temperature-dependent transformation given by Bogoliubov, Zubarev, and Tserkovnikov,[9] we can add up the ladder diagrams, if we use the simplified interaction (δ function in coordinate space, with a cutoff in momentum space) assumed by BCS. It is found that the sum of ladders with momentum zero is infinite, and we have to neglect these on the grounds that they should contribute only a negligible amount for an extended system. The sum of ladders with momentum $2K$ is <u>finite</u> in the transformed system, but goes to infinity like $-\log K$ as K goes to zero. There is a distinction between K close to zero and $K = 0$, since some of the interaction between a pair of particles with total momentum zero is included in the unperturbed Hamiltonian by the Bogoliubov transformation, and the more difficult problem of adding up ladders with $K = 0$ is treated in Appendix B.

The value of pairing particles with exactly opposite momentum in the BCS theory can be understood in these terms; if particles with total momentum $2K_0$ were paired, ladder diagrams with total momentum $2K < 2K_0$ would give a divergent sum.

Once again an attempt is made to find what physical effects the ladders with momentum almost zero might have. The specific heat behaves like $(T_C - T)^{-1/2}$ in the neighborhood of the transition, but the coefficient of this term is again very small. Other effects considered also seem to be too small for detection, and, in particular, the hypothesis of Heine and Pippard[11] that there should be a finite paramagnetism at zero temperature is rejected.

In Sec. V the methods developed in the earlier parts of the paper are applied to interactions with a more complicated form than the one originally used in BCS. The equations for the canonical transformation usually have more than one solution, and the usual criterion for distinguishing between these solutions is that the right one should give the lowest free energy. The methods of this paper suggest the alternative criterion that the ladder diagrams with nonzero momentum should have a convergent sum in the transformed system. The value of this criterion is illustrated by considering a problem in which neither of the two possible BCS solutions makes the ladder diagrams converge. In this case, there is in fact a lower energy state in which four particles are excited at a time from the Fermi sea. Application of this criterion to an angular-dependent interaction leads to conditions for the ground state to be not spherically symmetrical.

II. THE LADDER APPROXIMATION AT FINITE TEMPERATURES

We wish to work out the equation of state of a system of fermions interacting by two-body forces. The crystal structure of the superconducting metal is ignored, except in so far as it changes the effective mass of the electrons and the Fermi energy. There is a predominantly attractive interaction between electrons mediated by phonons, as well as the repulsive Coulomb force, but we include both of these in a two-body potential acting between the particles. The Hamiltonian must conserve momentum and spin, and we can write

$$
\begin{aligned}
H - \mu N = {} & \sum_k \epsilon_k (a^\dagger_{k,+} a_{k,+} + a^\dagger_{k,-} a_{k,-}) \\
& + \sum_K \sum_p \sum_q \{ V_{pq}^K a^\dagger_{K+p,+} a^\dagger_{K-p,-} a_{K-q,-} a_{K+q,+} \\
& + \frac{1}{2} W_{pq}^K (a^\dagger_{K+p,+} a^\dagger_{K-p,+} a_{K-q,+} a_{K+q,+} \\
& + a^\dagger_{K+p,-} a^\dagger_{K-p,-} a_{K-q,-} a_{K+q,-}) \} \quad . \qquad (3)
\end{aligned}
$$

The operators $a^\dagger_{k,+}$ and $a_{k,+}$ are the operators which create and destroy particles with momentum k and spin up (any specified direction), and the operators $a^\dagger_{k,-}$ and $a_{k,-}$ refer to particles with spin down. ϵ_k is the single-particle energy of an electron in the state k, measured from the chemical potential μ, while V and W give the interactions of particles with opposite spin and the same spin, respectively. The sums go over all states in momentum space.

We use the expansion of the thermodynamic potential derived by Bloch and De Dominicis.[3] Graphs, consisting of vertices (denoted by dots) and directed lines, are drawn with the following properties. Each line joins two vertices, or joins a vertex to itself. Two directed lines go into each vertex, and two come out of each vertex. Each vertex i is labeled with a coordinate t_i where $0 < t_i < \beta$; we draw the graph so that a vertex has a greater value of the coordinate t than all vertices to its right. We refer to a line going from left to right as a hole line, and a line going from right to left as a particle line. The graphs must be connected; that is, they must not consist of two or more parts unconnected by any lines. Each line is labeled with a momentum k and a spin up or down; there is no restriction on the label except that it must correspond to an eigenstate of the unperturbed Hamiltonian. It is convenient to regard a line as continuing through an interaction above the other line if it was originally above, and below it if it was originally below. The directed lines then form a series of closed loops whose number is well defined.

The contribution of a graph to the thermodynamic potential is found by getting a factor from each vertex, a factor from each line, and a factor -1 from each closed loop. The factor from a vertex is -1 times the matrix element of the interaction which annihilates particles in the states that label the lines entering the vertex and creates particles in the states that label the lines leaving the vertex. We use the convention that the upper line entering a vertex corresponds to the annihilation operator on the right, and the upper line leaving the vertex corresponds to the creation operator on the left of a term in Eq. (3). The factor from a line with momentum k going from a vertex at t to a vertex at t' is the propagator

$$S(k, t - t') \begin{cases} = -f_k \exp\left[\epsilon_k(t - t')\right] , & t \geqslant t' \\ = (1 - f_k) \exp\left[\epsilon_k(t - t')\right] , & t < t' , \end{cases} \qquad (4)$$

where

$$f_k = [\exp(\beta\epsilon_k) + 1]^{-1} . \tag{5}$$

We take the product of all these factors for a particular graph, integrate over all the coordinates from 0 to β , and divide by β . We must add the contributions from all distinct graphs to the expression for the thermodynamic potential of noninteracting fermions,[15] which we would get if V and W in Eq. (3) were zero.

Since it is impossible to make an exact calculation, we try to take into account a large class of graphs, in the hope that these will give the most important properties of the system. The usual first step in such an approximation is to redefine the single-particle energies so that they include some of the effects of the interaction; this is the procedure used to obtain the Hartree-Fock approximation, for example.[16] It will be assumed that this has been done already, since the single-particle energy spectrum used in Eq. (3) is the experimental one, and we cannot conveniently separate effects due to the lattice structure from the effects of the interaction of conduction electrons with one another. We therefore ignore "self-energy parts" of graphs. It can be seen from the results of Sec. III that the effects we consider here do not much influence the single-particle energies.

The graphs we consider are the ladder graphs, of which a few are shown in Fig. 1. These have the property that the two lines coming out of one vertex both go to one other vertex. We do not distinguish between particle lines and hole lines, which makes this definition of ladder diagrams different from the definition used in some other papers.[3,4] This class of diagrams, like the ring diagrams of Montroll and Ward,[2] is one of the simplest infinite classes. We will now derive a linear integral equation which gives the sum of these graphs, and show how an explicit formula can be derived for simple forms of the interaction.

Since momentum and spin are conserved in the interaction, the total momentum and total spin of a pair of lines going from one vertex to another is constant within a ladder diagram. We call these the momentum and spin of the ladder, and add together all ladder diagrams with the same momentum and spin. We consider only ladders with spin zero, since ladders with spin up or spin down can be summed in the same manner, replacing V by W. We define the ladder propagator $L_{nm}(K; t', t)$ to be the propagator which carries a pair of lines $[K + m, +]$ and $[K - m, -]$ at coordinate t to a pair of lines $[K + n, +]$ and $[K - n, -]$ at coordinate t' with any number of vertices in between, provided that the two lines which leave one vertex both go on to the next. Some diagrams contributing to the ladder propagator are shown in Fig. 2. To a first approximation, it is just the product of two single-particle propagators as defined by Eq. (4), for $m = n$. It satisfies the integral equation

$$L_{nm}(K; t', t) = S(K + m, t - t') S(K - m, t - t')\delta_{nm}$$

$$- \sum_p \int_0^\beta S(K + n, t'' - t') S(K - n, t'' - t') V_{np}^K L_{pm}(K; t'', t)dt''. \tag{6}$$

Unless this equation is singular, we can show that $L_{nm}(K; t', t)$ is periodic with period β, and that it is a function only of $t - t'$.[17] We use the property of S which follows from Eqs. (4) and (5), that, for $0 < t < \beta$,

$$S(k, t) = -S(k, t - \beta) . \tag{7}$$

This gives, with Eq. (6),

$$L_{nm}(K; 0, t) = L_{nm}(K; \beta, t) . \tag{8}$$

We can now operate on Eq. (6) with $\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial t'} \right)$ to get

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial t'} \right) L_{nm}(K; t', t) = \sum_p \left[S(K + n, t'' - t') S(K - n, t'' - t') V_{np}{}^K L_{pm}(K; t'', t) \right]_0^\beta$$

$$- \sum_p \int_0^\beta S(K + n, t'' - t') S(K - n, t'' - t') V_{np}{}^K \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial t''} \right) L_{pm}(K; t'', t) dt''$$

$$= - \sum_p \int_0^\beta S(K + n, t'' - t') S(K - n, t'' - t') V_{np}{}^K \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial t''} \right) L_{pm}(K; t'', t) dt'' . \tag{9}$$

Since this is a homogeneous integral equation for $\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial t'} \right) L_{nm}(K; t', t)$ with the same kernel as Eq. (6), which we have assumed to be nonsingular, we know that $L_{nm}(K; t', t)$ must be independent of $t + t'$, and must be a function only of $t - t'$.

Since Eq. (8) shows that L is a periodic function, we can write it as a Fourier series,

$$L_{nm}(K; t', t) = \sum_{\nu=-\infty}^{\infty} L_{nm}(K, \nu) \exp \left[2\pi i \nu (t - t') / \beta \right] . \tag{10}$$

Substitution of this in Eq. (6) gives the matrix equation

$$L_{nm}(K, \nu) = \frac{\tanh(\frac{1}{2}\beta\, \epsilon_{K+n}) + \tanh(\frac{1}{2}\beta\, \epsilon_{K-n})}{2\beta(\epsilon_{K+n} + \epsilon_{K-n}) - 4\pi\, i\, \nu} \left[\delta_{nm} - \beta \sum_p V_{np}^K L_{pm}(K, \nu)\right] . \tag{11}$$

The solution of this equation does not immediately give the contribution of the ladder graphs to the thermodynamic potential Ω . If we close the graphs of Fig. 2 with a vertex, getting a series for $-\sum_m \sum_n V_{mn}^K L_{nm}(K; t, t)$, we have closed-ladder graphs with one vertex singled out, and we obtain each distinct ladder graph a number of times equal to the number of its vertices. This means that $-\sum_m \sum_n V_{mn}^K L_{nm}(K; t, t)$ gives the contribution of the ladder graphs to $g\, d\Omega/dg$, where g is a coupling constant measuring the strength of the interaction V , rather than to Ω itself. There is also a slight difficulty due to the discontinuity in $L_{nm}(K; t', t)$ at $t' = t$, which comes from the discontinuity in S given by Eq. (4). This discontinuity is entirely in the first-order term, and we get around the difficulty by subtracting out the first-order term, represented by Fig. 1(a), and treating it separately if necessary.

The behavior of Eq. (11) is best illustrated by choosing V_{np}^K to be "separable." The most general separable potential is

$$V_{mn}^K = g\, v^*_m\, v_n , \tag{12}$$

where the coupling constant g is a real number. If this is substituted in Eq. (11), we can immediately solve the equation to get

$$\sum_m \sum_n g v_n\, L_{nm}(K, \nu) v^*_m = \beta^{-1}\, gQ(K, \nu)/[1 + gQ(K, \nu)] , \tag{13}$$

where

$$Q(K, \nu) = \sum_n \frac{\frac{1}{2}\beta \mid v_n \mid^2}{\beta(\epsilon_{K+n} + \epsilon_{K-n}) - 2\pi i \nu} [\tanh(\tfrac{1}{2}\beta\, \epsilon_{K+n}) + \tanh (\tfrac{1}{2}\beta\, \epsilon_{K-n})] . \tag{14}$$

We can also solve for $L_{nm}(K, \nu)$ alone, and get

$$L_{nm}(K, \nu) = F(K, n, \nu) \left\{ \delta_{nm} - \frac{g\, \beta\, v^*_n\, v_m\, F(K, m, \nu)}{1 + g\, Q(K, \nu)} \right\} , \tag{15}$$

where

$$F(K, n, \nu) = \frac{1}{2\beta(\epsilon_{K+n} + \epsilon_{K-n}) - 4\pi i \nu} [\tanh(\tfrac{1}{2}\beta\, \epsilon_{K+n}) + \tanh (\tfrac{1}{2}\beta\, \epsilon_{K-n})] . \tag{16}$$

Equation (13) gives the contribution of the ladder diagrams to $g\ d\Omega/dg$, and so we must integrate from zero to g in order to get the contribution to Ω. The integral is $\beta^{-1} \log[1 + gQ(K, \nu)]$. If we subtract the first-order term from this to remove the discontinuity in $L(K; t', t)$, we can substitute the expression for $L(K, \nu)$ into Eq. (8) and take the limit $t' = t$. The contribution of the ladder diagrams to the thermodynamic potential is now

$$\Omega_L = \sum_K \sum_{\nu=-\infty}^{\infty} \beta^{-1} \left\{ \log[1 + gQ(K, \nu)] - gQ(K, \nu) \right\} . \tag{17}$$

This expression is similar in structure to the expression obtained by Montroll and Ward for the sum of the ring diagrams,[2] although there the final sum (integration) is over the momentum transfer q rather than the total momentum 2K. In Appendix A, the result of Montroll and Ward is rederived by the methods used here, in order to emphasize the closeness of the analogy.

Equation (17) was derived by formally summing a power series in g, and this derivation will be suspect unless the sum has a convergent power series in g; the condition for this is $| gQ(K, \nu) | < 1$. Every term in the sum of Eq. (14) is positive for $\nu = 0$, and therefore $Q(K, 0)$ is greater than the sum of the moduli of the terms in the sum for any $Q(K, \nu)$, $\nu \neq 0$. Therefore $Q(K, 0)$ determines the radius of convergence of the power-series expansion of the summand of Eq. (17). It also seems probable that $Q(K, 0)$ has a maximum for $K = 0$, although this depends on the behavior of v_n, so that $Q(0, 0)$ determines the radius of convergence. The question of the maximum of $Q(K, 0)$ will be examined more carefully in Sec. III. The condition for convergence is $| gQ(0, 0) | < 1$, and, from Eq. (14), this gives

$$| g | \sum_n (|v_n|^2 / 2 \epsilon_n) \tanh(\tfrac{1}{2} \beta \epsilon_n) < 1 . \qquad (18)$$

The failure of convergence would be of no particular interest for positive g (repulsive forces), since one could then regard Eq. (17) as an analytic continuation, similar to that made in the papers of Gell-Mann and Brueckner[18] and Montroll and Ward.[2] If, however, g is negative (attractive forces) and the inequality (18) is not satisfied, then there will be some value of K for which $gQ(K, 0) = -1$, since $Q(K, 0)$ falls off smoothly to zero for very large K, with a finite potential. This means that there are infinite terms in the sum of Eq. (17). For fixed g, which we take equal to -1, Eq. (18) can be regarded as a condition on the inverse of the temperature, and can be written as $\beta < \beta_C$, where

$$\sum_n \left(|v_n|^2 / 2\,\epsilon_n \right) \tanh\left(\tfrac{1}{2} \beta_C\, \epsilon_n \right) = 1 \quad . \tag{19}$$

The potential used for the calculations of BCS is of the form (12), with $g = -1$ and $|v_n|^2$ a constant for n within a certain distance of the Fermi surface and zero otherwise. Equation (19) is identical with the BCS equation for the critical temperature. The condition for convergence of the sum of the ladder diagrams is simply that the temperature be greater than the critical temperature, or that the equilibrium state of the metal should be the normal state.

It is not only with this simple interaction that our condition for convergence is equivalent to the BCS condition for the normal state to be the stable state. We can replace V_{np} in Eq. (11) by $-gV_{np}$, and then get a power series in g by iterating the equation. This power series will converge for $g = 1$ if L_{nm} , regarded as a function of g , has no singularities for $|g| < 1$. The condition for a singularity is the existence of a nonzero solution of the homogeneous equation

$$C_n = \sum_p \beta\, g\, F(K, n, \nu) V_{np}\, C_p \quad , \tag{20}$$

where F is defined by Eq. (16). The condition that there should be no solution of this for $K = 0$, $\nu = 0$, $|g| < 1$, can be regarded as a condition on the temperature. Since $F(K, n, 0)$ is real and positive, Eq. (20) can be written as an eigenvalue equation for the Hermitian matrix $[F(K, n, 0)F(K, p, 0)]^{1/2}\, V_{np}$, and the eigenvalues must be real. For a repulsive potential, the eigenvalues are negative and do not correspond to infinite terms in the ladder approximation. For an attractive potential, the existence of solutions of Eq. (20) with $g < 1$ is equivalent to $\beta < \beta_C$, where β_C is the smallest solution of

$$C_n = \sum_p (1/2\epsilon_n)\tanh(\tfrac{1}{2}\beta_C\,\epsilon_n)V_{np}\,C_p \quad . \tag{21}$$

This is equivalent to the condition for the critical temperature derived by Bogoliubov.[19]

It will be convenient here to generalize these results and consider a representation in which the propagators S defined by Eq. (4) are not diagonal. Writing the Fourier transform of the product of two propagators as $S_{nm}(K, \nu)$, where the index m defines the initial state and the index n defines the final state, we get an equation for the ladder propagator equivalent to Eq. (11),

$$L_{nm}(K, \nu) = S_{nm}(K, \nu) - \beta \sum_p \sum_q S_{np}(K, \nu)V_{pq}\,L_{qm}(K, \nu) \quad . \tag{22}$$

If we can write the potential as

$$V_{pq} = \sum_i g_i\, v_p^{i*}\, v_q^{i} \quad , \tag{23}$$

we can transform the matrices to

$$\begin{aligned} L^{jk}(K, \nu) &= (g_j g_k)^{1/2} \sum_n \sum_m v_n^{j}\, L_{nm}(K, \nu)v_m^{k*} \quad , \\ S^{jk}(K, \nu) &= (g_j g_k)^{1/2} \sum_n \sum_m v_n^{j}\, S_{nm}(K, \nu)v_m^{k*} \quad , \end{aligned} \tag{24}$$

and get the equation

$$L^{jk}(K, \nu) = S^{jk}(K, \nu) - \beta \sum_{\ell} S^{j\ell}(K, \nu)L^{\ell k}(K, \nu) \quad . \tag{25}$$

If the eigenvalues of $S^{jk}(K, \nu)$ are $\lambda_i(K, \nu)$, then this gives

$$\mathrm{Tr}\ L^{jk}(K, \nu) = \sum_i [1 + \beta\lambda_i(K, \nu)]^{-1} \lambda_i(K, \nu) \quad . \tag{26}$$

The condition for convergence of the series expansion of this is $|\beta\lambda_i| < 1$, although there is an infinity in the sum only if there is an i for which $\beta\lambda_i = -1$.

III. APPROACH TO THE CRITICAL TEMPERATURE FROM ABOVE

Evaluation of the Thermodynamic Potential.

When β is just below the value β_C defined by Eq. (19) or (21), the series expansion given by perturbation theory should converge, but the sum on the right of Eq. (17) has some very large terms in it, because $Q(0, 0)$ is very nearly unity. We use Eq. (14) to evaluate $Q(K, \nu)$ for β just less than β_C and K small. For simplicity, we take the potential used in BCS; this is like the potential of Eq. (12), with $|v_n|^2$ equal to a constant, $J/\mathcal{V}$, where $\mathcal{V}$ is the total volume of the system, for $k_F - w < n < k_F + w$, and with $|v_n|^2$ equal to zero otherwise. The Fermi momentum k_F is defined by

$$k_F^2/2M = \mu , \tag{27}$$

where μ is the chemical potential and M is the electron effective mass; we set $\hbar = 1$ throughout this paper. We have, then,

$$Q(K, \nu) = \frac{J\beta}{8\pi^2} \int_{k_F-w}^{k_F+w} n^2\, dn \int_{-1}^{1} d(\cos\theta)\, \frac{1}{\beta(K^2 + n^2 - k_F^2)/M - 2\pi i \nu}$$

$$\times \Big\{ \tanh\,[\beta(K^2 + n^2 - k_F^2 + 2n K \cos\theta)/4M]$$

$$+ \tanh[\beta(K^2 + n^2 - k_F^2 - 2n K \cos\theta)/4M] \Big\}$$

$$= \frac{JM}{2\pi^2 K} \int_{k_F-w}^{k_F+w} \frac{n\, dn}{\beta(K^2 + n^2 - k_F^2)/M - 2\pi i \nu} \log \frac{\cosh\{\beta[(n+K)^2 - k_F^2]/4M\}}{\cosh\{\beta[(n-K)^2 - k_F^2]/4M\}} . \tag{28}$$

To evaluate this integral, we must make some assumptions about the magnitudes of the quantities entering the expression. Our assumptions are $k_F \gg w \gg K$ and $\beta\, k_F w/M \gg 1$. We change the variable to

$$z = \beta\, k_F(n - k_F)/2M - \beta\, K^2/4M \tag{29}$$

and drop all but the highest powers of k_F. This gives

$$Q(K, \nu) = \frac{JM^2}{\pi^2 \beta\, K} \int_{-B+C}^{B+C} \frac{dz}{4z - 2\pi\, i\, \nu} \log \frac{\cosh(z + A)}{\cosh(z - A)} \;, \tag{30}$$

where

$$A = \beta\, K\, k_F/2M, \qquad B = \beta\, k_F\, w/2M\;, \qquad C = \beta\, K^2/4M\;. \tag{31}$$

The integral on the right of Eq. (30) can be evaluated by contour integration, because of the assumption that B is large. The integrand has branch points at $z = \frac{1}{2}(2\nu' + 1)\,\pi\, i \pm A$, where ν' is a whole number, and we define the integrand by making straight cuts of length $2A$ between them. There is a pole at $z = \frac{1}{2}\pi\, i\, \nu$ for ν odd, but we take the contour of integration in the lower half plane if ν is positive, and in the upper half plane if ν is negative, and thus we avoid contributions from the pole. We take the contour of integration to be three sides of a rectangle with vertices at $z = -B + C$, $z = -B + C + \pi\, i\, \nu''$, $z = B + C + \pi\, i\, \nu''$, and $z = B + C$, where ν'' is a whole number of opposite sign to ν. The logarithm in Eq. (30) tends to $-2A$ if the real part of z is negative, and to $2A$ if the real part of z is positive, with only exponentially small terms left over. The sum of the contributions to the integral of the first and third sides of the rectangle is

$$\frac{1}{2} A \log \frac{(B + C - \frac{1}{2}\pi i \nu)(-B + C - \frac{1}{2}\pi i \nu)}{(B + C + \pi i \nu'' - \frac{1}{2}\pi i \nu)(-B + C + \pi i \nu'' - \frac{1}{2}\pi i \nu)} . \tag{32}$$

On the second side, the modulus of the integrand is not more than $A/2\pi \; |2\nu'' - \nu|$, so that the integral along a line of length $2B$ can be made to vanish by taking $|\nu''|$ large enough. The contribution from one of the cuts is

$$\mp 2\pi i \int_{\frac{1}{2}(2\nu'+1)\pi i - A}^{\frac{1}{2}(2\nu'+1)\pi i + A} \frac{dz}{4z - 2\pi i \nu} = \mp \frac{1}{2}\pi i \log \frac{\pi i(2\nu' - \nu + 1) + 2A}{\pi i(2\nu' - \nu + 1) - 2A} , \tag{33}$$

where the upper sign is for ν positive and the lower for ν negative. We can combine the expressions (32) and (33) to get the result

$$\int_{-B+C}^{B+C} \frac{dz}{4z - 2\pi i \nu} \log \frac{\cosh(z + A)}{\cosh(z - A)} = \frac{1}{2} A \log[(B + C - \frac{1}{2}\pi i \nu)(B - C + \frac{1}{2}\pi i \nu)/\pi^2]$$

$$+ \lim_{\nu'' \to \infty} \left\{ -A \log \nu'' + \frac{1}{2}\pi i \sum_{\nu'=0}^{\nu''-1} \log \frac{\pi i(2\nu' + |\nu| + 1) + 2A}{\pi i(2\nu' + |\nu| + 1) - 2A} \right\} . \tag{34}$$

The dependence on C is unimportant and is neglected. If A is much less than unity, we can expand Eq. (34) as a power series in K, and, keeping the leading terms, we get

$$Q(K, \nu) = \frac{JMk_F}{2\pi^2} \left\{ \frac{1}{2} \log \left(\frac{\beta^2 k_F^2 w^2}{4 M^2} + \frac{1}{4} \pi^2 \nu^2 \right) + \gamma + \log(4/\pi) \right.$$

$$\left. - 2 \sum_{\nu'=1}^{\frac{1}{2}|\nu|} \frac{1}{2\nu' - 1} - \frac{\beta^2 K^2 k_F^2}{12 \pi^2 M^2} \left[7 \zeta(3) - \sum_{\nu'=1}^{\frac{1}{2}|\nu|} \left(\frac{2}{2\nu' - 1} \right)^3 \right] \right\},$$

for ν even,

$$Q(K, \nu) = \frac{JMk_F}{2\pi^2} \left\{ \frac{1}{2} \log \left(\frac{\beta^2 k_F^2 w^2}{4 M^2} + \frac{1}{4} \pi^2 \nu^2 \right) + \gamma - \log \pi \right.$$

$$\left. - \sum_{\nu'=1}^{\frac{1}{2}(|\nu|-1)} \frac{1}{\nu'} - \frac{\beta^2 K^2 k_F^2}{12 \pi^2 M^2} \left[\zeta(3) - \sum_{\nu'=1}^{\frac{1}{2}(|\nu|-1)} \frac{1}{\nu'^3} \right] \right\}, \tag{35}$$

for ν odd.

Here γ is Euler's constant, equal to 0.577, and $\zeta(3)$ is the Riemann zeta function of argument three, equal to 1.202.

We can evaluate the integral in Eq. (30) for A large (but still smaller than B) by putting $\log[\cosh(z + A)] = | z + A |$. This gives, if we neglect C,

$$Q(K, \nu) = \frac{JMk_F}{2\pi^2} \left\{ \frac{1}{2} \log \frac{B^2 + \frac{1}{4}\pi^2 \nu^2}{A^2 + \frac{1}{4}\pi^2 \nu^2} + 1 - \frac{\pi \nu}{2 A} \tan^{-1} \left(\frac{2 A}{\pi \nu} \right) \right\} . \tag{36}$$

According to Eqs. (35) and (36), $Q(K, \nu)$ appears to have its largest value for $K = 0$ and $\nu = 0$, while $Q(0, 0) = 1$ substituted in Eq. (35) gives an equation for β_C in agreement with BCS Eq. (3.29). We wish to evaluate Eq. (17) for the case of $Q(0, 0)$ just less than unity, and we are interested mainly in the effect of the logarithmic singularity of the summand. Now, although the argument of the logarithm in Eq. (35) is a large number, essentially the square of the ratio of the Debye temperature to the critical temperature, the logarithm is only about 10, and $Q(0, \nu)$ falls off from its maximum value quite rapidly as ν increases from zero. For this reason, we shall consider only those terms in Eq. (17) with $\nu = 0$, since we should be able to use low-order perturbation to take account of the effects of the other terms. For the same reason, we consider the contribution only of terms with very low K, and we therefore cut off the summation over K at a value L, using Eq. (35) to determine $Q(K, 0)$ in this region. Comparison of Eqs. (35) and (36) shows that the cutoff L should be of the order of $2M/\beta k_F$. In this approximation, we get Eq. (17) as

$$\Omega_L = \frac{4\mathcal{V}}{\pi^2 \beta} \int_0^L K^2 \log \left\{ \frac{JMk_F}{2\pi^2} \left[\log \frac{\beta_C}{\beta} + \frac{7\,\zeta(3)\beta^2 k_F^{\,2}}{12\,\pi^2 M^2} K^2 \right] \right\} dK , \tag{37}$$

where we have made use of the equation for β_C. This expression can be evaluated by use of the equation

$$\int_0^L K^2 \log(a^2 + b^2 K^2) dK = -\frac{2}{9} L^3 + \frac{2}{3}\frac{a^2}{b^2} L + \frac{1}{3} L^3 \log(a^2 + b^2 L^2) - \frac{2}{3}\left(\frac{a}{b}\right)^3 \tan^{-1}\left(L\frac{b}{a}\right) . \tag{38}$$

Specific Heat Anomaly

Thermodynamic quantities can be calculated by taking derivatives of Eq. (37) with respect to β and k_F. The terms on the right of Eq. (38) are all regular near $a = 0$ except for the last one. This last term is finite and has finite first derivatives, but its second derivative with respect to a^2 behaves like $-\frac{1}{4}\pi\, a^{-1}\, b^{-3}$ near $a = 0$. It follows that the specific heat of the electron gas should become infinite at the critical temperature T_C. If we write this anomalous part of the specific heat per unit volume as C_a and write $\theta = (T - T_C)/T_C$, we get, from Eqs. (37) and (38),

$$C_a = \pi^2\, [12/7\, \zeta(3)]^{3/2}\, M^3\, k_F^{-3}\, k^4\, T_C^{\,3}\, \theta^{-1/2} \quad . \tag{39}$$

This must be compared with the contribution from the unperturbed system at the temperature T_C, which is[20]

$$C = M\, k_F\, k^2\, T_C/3 \quad , \tag{40}$$

and the ratio of the two is

$$C_a/C = 12.6\, (T_C/T_F)^2\, \theta^{-1/2} \quad , \tag{41}$$

where the Fermi temperature T_F is defined by $k\, T_F = k_F^{\,2}/2M$.

Equation (41) shows that this specific-heat anomaly is far beyond the range of observation for usual superconducting substances. A quite favorable case would be niobium,[21] which has $T_C = 8.8^\circ$, $C = \gamma\, T$, where $\gamma = 19 \times 10^{-4}$ cal/mole deg.^{2}, and, presumably, five electrons per atom. This would give $C_a/C = 1.3 \times 10^{-6}\, \theta^{-1/2}$, so that it would be necessary to go within 10^{-11} degrees of the critical temperature for the anomalous part of the specific heat to be comparable to the normal part of the electronic specific heat.

Although this specific-heat anomaly is so small, it is of some theoretical interest, since the theory presented here suggests that the specific heat appears finite but discontinuous at the critical temperature only because of the small ratio of the critical temperature to the Fermi temperature. This behavior of the specific heat should be contrasted with the logarithmic behavior of the specific heat of He^4 observed near the lambda point.[22] The specific heat of a substance near a critical point at which a second-order transition becomes first-order behaves like Eq. (39), according to Landau's theory.[14]

Electron Correlation Function

The electron correlation function $g(x+, x'-)$ is defined as the probability per unit volume of finding an electron with spin up at x and an electron with spin down at x', minus the probability of finding an electron at x times the probability of finding an electron at x'. Making use of the conservation of momentum, we can write this as

$$g(x+, x'-) = \mathcal{V}^{-2} \sum_K \sum_m \sum_n \langle a^\dagger_{K+n,+} a^\dagger_{K-n,-} a_{K-m,-} a_{K+m,+} \rangle e^{i(n-m)(x-x')} - \mathcal{V}^{-2} [\sum_p \langle a^\dagger_{p,+} a_{p,+} \rangle]^2 , \tag{42}$$

where the averages are taken over the statistical ensemble. For noninteracting particles, this function would be zero. In the ladder approximation, the first average is just $\lim_{t \to +0} L_{nm}(K; 0, t)$. Use of Eqs. (10) and (15) shows

$$g(x+, x'-) = \beta \mathcal{V}^{-2} \sum_\nu \sum_K \frac{\left| \sum_m v_m F(K, m, \nu) \exp[i\, m(x - x')] \right|^2}{1 - Q(K, \nu)} . \tag{43}$$

This function vanishes for x and x' far apart. From Eqs. (16) and (14) and from the form of interaction we have chosen, with $|v_m|^2$ equal either to J/γ or zero, we deduce

$$g(x+, x-) = (\beta J \gamma)^{-1} \sum_{\nu} \sum_{K} [Q(K, \nu)]^2 / [1 - Q(K, \nu)] . \qquad (44)$$

Again we are interested in the effect of the smallness of the denominator, we take only the $\nu = 0$ term, and integrate over K within L of zero, putting $Q(K, 0) = 1$ in the numerator and using the same approximation for the denominator as was used in Eq. (37). This leads to

$$g(x+, x-) = 8(\beta J^2 M k_F)^{-1} [L b^{-2} - \theta^{1/2} b^{-3} \tan^{-1}(L b \theta^{-1/2})] , \qquad (45)$$

where θ again denotes $(T - T_C)/T_C$ and

$$b^2 = 7 \zeta(3) \beta^2 k_F^2 / 12 \pi^2 M^2 . \qquad (46)$$

The right-hand side of Eq. (45) goes to zero if θ is large, but goes to $8L/\beta J^2 M k_F b^2$ if θ is much less than unity, since $L b$ is of the order of unity. Therefore, as θ goes to zero, the probability of finding two electrons of opposite spin close together increases. The range of this correlation is of the order of b^{-1}, since the main contribution to the sum over m in Eq. (43) comes from the region in which m is less than b. Thus, there is a correlation between electrons of opposite spin in the normal state just above the critical temperature very similar to the correlation in the superconducting state of the BCS theory, with the same range, and of the same order of magnitude. This correlation is, however, very small, and it can be described by saying that the electrons are distributed in such a way that there are about T_C/T_F more electrons within a distance b^{-1} of a given electron than there would be in a random distribution.

There are various other effects predicted by this theory which should also be very small. There is a slight enhancement of the spin paramagnetism, and an enhancement of the number of high-energy phonons present as the temperature approaches the critical temperature. No effect has been found which might be experimentally detectable.

IV. THE SUPERCONDUCTING STATE

The Ladder Approximation in the BCS Theory

Below the critical temperature, the perturbation series no longer converges in the ladder approximation. We might try to avoid this difficulty by making an analytic continuation and using Eq. (17), in spite of the lack of convergence of its power-series expansion. Such a procedure is used in the theory of Montroll and Ward,[2] but there are two reasons for not using it here. One reason is that, before we take the limit of an infinite system, $Q(K, 0)$ may take the value unity for a point on the reciprocal lattice, and the pressure will be infinite in such a case. There will therefore be violent and quite meaningless fluctuations of the pressure which depend on the size of and shape of the system. The second reason is that we know from BCS theory that interactions of particles with total spin and momentum zero contribute a finite amount to the pressure (or to the energy per particle at zero temperature), and no such effect appears from Eq. (17).

We therefore take the BCS theory as a starting point, and then apply perturbation theory. The most convenient form of the theory is that proposed by Bogoliubov, Zubarev, and Tseikovnikov,[9] which is equivalent to BCS theory. We make the canonical transformation

$$\alpha_{m0} = x_m^* \, a_{m,+} - y_m \, a^\dagger_{-m,-} \quad ,$$

$$\alpha_{m1} = x_m^* \, a_{-m,-} + y_m \, a^\dagger_{m,+} \quad , \qquad (47)$$

where

$$|x_m|^2 + |y_m|^2 = 1 ,$$

$$|x_m|^2 - |y_m|^2 = \epsilon_m/\Omega_m ,$$

$$x_m y_m = C_m/2\Omega_m ,$$

$$\Omega_m = \sqrt{(\epsilon_m^2 + |C_m|^2)} . \qquad (48)$$

The quantity C_m is determined by

$$C_m = - \sum_n V_{mn}(C_n/2\Omega_n) \tanh(\tfrac{1}{2} \beta \Omega_n) , \qquad (49)$$

where V_{mn} is the interaction between particles of opposite spin which occurs in Eq. (3). If the transformation (47) is applied to the part of the Hamiltonian which takes into account interaction between particles with opposite spin and momentum--the first sum on the right of Eq. (3) together with the $K = 0$ term of the second sum--the Hamiltonian becomes

$$(H - \mu N)_B = \sum_n [\epsilon_n - \Omega_n + (|C_n|^2/2\Omega_n) \tanh(\tfrac{1}{2} \beta \Omega_n)]$$

$$+ \sum_n \Omega_n(\alpha^\dagger_{n0} \alpha_{n0} + \alpha^\dagger_{n1} \alpha_{n1}) + \sum_m \sum_n V_{mn} B^\dagger_m B_n , \qquad (50)$$

where

$$B_m = a_{-m,-} a_{m,+} - (C_m/2\Omega_m) \tanh(\tfrac{1}{2} \beta \Omega_m) . \qquad (51)$$

The operators α_{n0} and α_{n1} are annihilation operators for "quasiparticles," and the BCS solution of the statistical mechanical problem is the solution for noninteracting quasiparticles, so that the probability of a quasiparticle states being occupied is $[\exp(\beta \Omega_n) + 1]^{-1}$. The first two sums on the right of Eq. (50) give the unperturbed Hamiltonian, while the third sum on the right of Eq. (50), together with the "residual" terms not included in Eq. (50), give the perturbation.

In this section we assume that W is zero and that V is separable, given by Eq. (12) with $g = -1$, as we assumed in the preceding section. In Sec. V we consider the effects of removing this restriction in some simple cases. Equations (49) and (48) now have a unique solution for $\beta > \beta_C$ (apart from an arbitrary phase factor), and no solution for $\beta < \beta_C$. There are two kinds of terms in the perturbation series which we must consider separately. Firstly there are those which result from interactions between particles with total momentum zero. It is shown in Appendix B that the series of ladder diagrams with total momentum zero gives an infinite sum even when the BCS energy has been subtracted. This indicates that the Hamiltonian (50) requires more careful study, but we ignore this, and neglect these terms on the usual grounds that their contribution to the pressure is of order $\mathcal{V}^{-1}$. Secondly there are those terms which result from interactions between particles with nonzero total momentum, and it is from these that we expect results not contained in BCS.

We use a particle representation rather than a quasi-particle representation. The propagators, being diagonal in the quasi-particle representation, are nondiagonal, since they mix a particle of one spin and momentum with a hole of the opposite spin and momentum. We wish to solve Eq. (22), which is an equation in a space of two-particle states. The intermediate state in this equation could be two particles, or a hole and a particle, or two holes. We ignore the possibility of the state's containing one hole and one particle, since the corresponding matrix element of V gives a small momentum transfer K to two particles instead of involving two particles with small total momentum 2K , and we are neglecting the long-range part of the potential. We denote the state with particles in [K + n,+] and [K - n,-] by the suffix n , and the state with holes in [-K - n,-] and [-K + n,+] by the suffix $\bar{n}$. The potential is given by

$$V_{\bar{m}\bar{n}} = V_{nm} = -v^*_n v_m ,$$
$$V_{\bar{m}n} = V_{m\bar{n}} = 0 . \tag{52}$$

Application of the transformation (47) to the propagators, which look like Eq. (4) with ϵ replaced by Ω in the quasi-particle representation, gives the two-particle propagators as

$$S_{mm}(K, \nu) = \left[\frac{\frac{1}{2}| x_{K+m} |^2 | x_{K-m} |^2}{\beta(\Omega_{K+m} + \Omega_{K-m}) - 2\pi i \nu} + \frac{\frac{1}{2} | y_{K+m} |^2 | y_{K-m} |^2}{\beta(\Omega_{K+m} + \Omega_{K-m}) + 2\pi i \nu} \right]$$
$$\times [\tanh(\tfrac{1}{2} \beta \Omega_{K+m}) + \tanh(\tfrac{1}{2} \beta \Omega_{K-m})]$$
$$+ \left[\frac{\frac{1}{2} | x_{K+m} |^2 | y_{K-m} |^2}{\beta(\Omega_{K+m} - \Omega_{K-m}) - 2\pi i \nu} + \frac{\frac{1}{2} | y_{K+m} |^2 | x_{K-m} |^2}{\beta(\Omega_{K+m} - \Omega_{K-m}) + 2\pi i \nu} \right]$$
$$\times [\tanh(\tfrac{1}{2} \beta \Omega_{K+m}) - \tanh(\tfrac{1}{2} \beta \Omega_{K-m})] ,$$

$$S_{m\bar{m}}(K, \nu) = -x_{K+m} y_{K+m} x_{K-m} y_{K-m} \left[\mathrm{Re} \frac{\tanh(\frac{1}{2} \beta \Omega_{K+m}) + \tanh(\frac{1}{2} \beta \Omega_{K-m})}{\beta(\Omega_{K+m} + \Omega_{K-m}) - 2\pi i \nu} \right.$$
$$\left. - \mathrm{Re} \frac{\tanh(\frac{1}{2} \beta \Omega_{K+m}) - \tanh(\frac{1}{2} \beta \Omega_{K-m})}{\beta(\Omega_{K+m} - \Omega_{K-m}) - 2\pi i \nu} \right] , \tag{53}$$

(Eq. (53) Cont.)

$$S_{m\bar{m}}(K,\nu) = S^{*}_{m\bar{m}}(K,\nu) \quad ,$$
$$S_{\bar{m}\bar{m}}(K,\nu) = S^{*}_{mm}(K,\nu) \quad , \tag{53}$$

with all other matrix elements zero.

The potential defined by Eq. (52) is the sum of two separable parts, the first part being zero for all hole states, and the second part being zero for all particle states. We can define the two-by-two matrix $S^{ij}(K,\nu)$ in accordance with Eq. (24); this has components equal to $-\sum_m |v_m|^2 S_{mm}(K,\nu)$, $-\sum_m v_m^{\,2}\, S_{m\bar{m}}(K,\nu)$, $-\sum_m v^{*\,2}_m\, S_{\bar{m}m}(K,\nu)$, and $-\sum_m |v_m|^2\, S_{\bar{m}\bar{m}}(K,\nu)$. This can be evaluated by using Eq. (53), together with Eq. (48), to give

$$S^{ij}(K,\nu) = -\sum_m \frac{1}{2} \left\{ \mathrm{Re}\, \frac{|v_m|^2}{\beta(\Omega_{K+m}+\Omega_{K-m}) - 2\pi i \nu} \right.$$
$$\left. + \left(|x_{K+m}|^2\, |y_{K-m}|^2 - |y_{K+m}|^2\, |x_{K-m}|^2 \right) \mathrm{Im}\, \frac{|v_m|^2}{\beta(\Omega_{K+m}+\Omega_{K-m}) - 2\pi i \nu} \right\}$$
$$\times \left\{ \tanh\!\left(\tfrac{1}{2}\beta\,\Omega_{K+m}\right) + \tanh\!\left(\tfrac{1}{2}\beta\,\Omega_{K-m}\right) \right\} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$+ \sum_m \left\{ \mathrm{Re}\, \frac{\tanh\!\left(\tfrac{1}{2}\beta\,\Omega_{K+m}\right) + \tanh\!\left(\tfrac{1}{2}\beta\,\Omega_{K-m}\right)}{\beta(\Omega_{K+m}+\Omega_{K-m}) - 2\pi i \nu} - \mathrm{Re}\, \frac{\tanh\!\left(\tfrac{1}{2}\beta\Omega_{K+m}\right) - \tanh\!\left(\tfrac{1}{2}\beta\Omega_{K-m}\right)}{\beta(\Omega_{K+m}-\Omega_{K-m}) - 2\pi i \nu} \right\}$$
$$\begin{pmatrix} \tfrac{1}{2}\{ |x_{K+m}y_{K-m}|^2 + |y_{K+m}x_{K-m}|^2 \}\, |v_m|^2 & x_{K+m}y_{K+m}\, x_{K-m}y_{K-m}\, v_m^{\,2} \\ x^{*}_{K+m}y^{*}_{K+m}\, x^{*}_{K-m}y^{*}_{K-m}\, v^{*\,2}_m & \tfrac{1}{2}\{ |x_{K+m}y_{K-m}|^2 + |y_{K+m}x_{K-m}|^2 \}\, |v_m|^2 \end{pmatrix} \tag{59}$$

In the case $K = 0$, the second matrix on the right of Eq. (54) can be simplified, since Eq. (49) shows that C_m/v^*_m must be a constant A. If we take the factor $|v_m|^4/2\Omega_m$ out of the matrix, according to Eq. (48) we are left with a matrix independent of m, whose elements are $|A|^2$, A^2, A^{*2}, and $|A|^2$, and whose eigenvalues are zero and $2|A|^2$. The eigenvalues of $S^{ij}(0, \nu)$ are therefore

$$\lambda_1(0, \nu) = -\sum_m |v_m|^2 \{\mathrm{Re}(2\beta\Omega_m - 2\pi i \nu)^{-1} + (\epsilon_m/\Omega_m)\mathrm{Im}(2\beta\Omega_m - 2\pi i \nu)^{-1}\} \tanh(\tfrac{1}{2}\beta\Omega_m) \tag{55}$$

$$\begin{aligned}\lambda_2(0, \nu) = -\sum_m |v_m|^2 \{&(\epsilon_m{}^2/\Omega_m{}^2)\,\mathrm{Re}(2\beta\Omega_m - 2\pi i \nu)^{-1} \\ &+ (\epsilon_m/\Omega_m)\mathrm{Im}(2\beta\Omega_m - 2\pi i \nu)^{-1}\}\tanh(\tfrac{1}{2}\beta\Omega_m) \\ &- \sum_m (|C_m v_m|^2/4\Omega_m{}^2)\,\mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m)\delta_{\nu 0} \ .\end{aligned}$$

Equation (49) gives

$$\sum_n (|v_n|^2/2\Omega_n)\tanh(\tfrac{1}{2}\beta\Omega_n) = 1 \ , \tag{56}$$

and so $\lambda_1(0, 0)$ is $-\beta^{-1}$, while $\lambda_1(0, \nu)$ has modulus less than β^{-1} for $\nu \neq 0$. The modulus of $\lambda_2(0, \nu)$ is clearly less than β^{-1} for $\nu \neq 0$, and the modulus of $\lambda_2(0, 0)$ is less than β^{-1} because $(2\beta\Omega_m)^{-1}\tanh(\tfrac{1}{2}\beta\Omega_m)$ is greater than $\tfrac{1}{4}\mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m)$. Since the condition for convergence of the perturbation series is that $|\beta\lambda|$ should always be less than unity, we need be concerned only with $\lambda_1(K, 0)$, when K is close to zero.

Considering just this eigenvalue $\lambda_1(K, 0)$, with K close to zero, we shall certainly not decrease the modulus of the eigenvalue if we replace the nondiagonal elements of the second matrix of Eq. (54) by their moduli; for $K = 0$ this does not change the eigenvalue. The eigenvalue is

$$\lambda_1(K,\ 0) = -\sum_m |v_m|^2 \left[2\beta(\Omega_{K+m} + \Omega_{K-m})\right]^{-1} \left[\tanh(\tfrac{1}{2}\beta\Omega_{K+m}) + \tanh(\tfrac{1}{2}\beta\Omega_{K-m})\right]$$

$$+ \sum_m \frac{|v_m|^2(\Omega_{K+m}\,\Omega_{K-m} - \epsilon_{K+m}\,\epsilon_{K-m} - |C_{K+m}\,C_{K-m}|)}{4\,\Omega_{K+m}\,\Omega_{K-m}}$$

$$\times \left[\frac{\tanh(\frac{1}{2}\beta\Omega_{K+m}) + \tanh(\frac{1}{2}\beta\Omega_{K-m})}{\beta(\Omega_{K+m} + \Omega_{K-m})} - \frac{\tanh(\frac{1}{2}\beta\Omega_{K+m}) - \tanh(\frac{1}{2}\beta\Omega_{K-m})}{\beta(\Omega_{K+m} - \Omega_{K-m})}\right] . \tag{57}$$

This equation is also valid if v_m is everywhere real. Equation (48) shows

$$\Omega_{K+m}\,\Omega_{K-m} - \epsilon_{K+m}\epsilon_{K-m} - |C_{K+m}\,C_{K-m}|$$

$$= (\epsilon_{K+m}\,|C_{K-m}| - |C_{K+m}|\,\epsilon_{K-m})^2(\Omega_{K+m}\,\Omega_{K-m} + \epsilon_{K+m}\,\epsilon_{K-m} + |C_{K+m}\,C_{K-m}|)^{-1}, \tag{58}$$

which is always positive. The second sum on the right of Eq. (57) is therefore always positive, but whether the first sum on the right is greater or less than $-\beta^{-1}$ depends on the form of $|v_m|^2$. For $|v_m|^2$ roughly constant in the neighborhood of the Fermi surface, we expect $\lambda_1(K,\ 0)$ to have a minimum at $K = 0$.

We now make an expansion of Eq. (57), keeping only the second order in K. We neglect the variation of C_m with m, so that we have

$$\Omega_{K+m} \doteq \Omega_m + \epsilon_m(m\cdot K)/M\Omega_m + \epsilon_m K^2/2M\Omega_m + (|C_m|^2/2M^2\Omega_m{}^3)(m\cdot K)^2 \quad . \tag{59}$$

In this approximation, Eq. (57) becomes

$$1 + \beta\lambda_1(K,\ 0) = \sum_m \left[\frac{|v_m|^2\ \epsilon_m\ K^2}{4M\ \Omega_m^{\ 3}} + \frac{3|v_m\ C_m|^2\ (m\cdot K)^2}{4M^2\ \Omega_m^{\ 5}} \right]$$

$$\times \left[\tanh(\tfrac{1}{2}\beta\ \Omega_m) - \tfrac{1}{2}\beta\Omega_m\ \mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m) \right]$$

$$+ \sum_m \frac{|v_m|\ \beta^2\ \epsilon_m^{\ 2}\ (m\cdot K)^2}{8M^2\ \Omega_m^{\ 3}}\ \mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m)\ \tanh(\tfrac{1}{2}\beta\Omega_m)\ . \tag{60}$$

The assumption of very large k_F, with $|v_m|^2$ an even function of $m - k_F$, which was made in Sec. III and in BCS, can be used to simplify this considerably. If we do the angular integration, and then cancel out terms in the sums which have an odd dependence on ϵ_m, we are left with

$$1 + \beta\lambda_1(K,\ 0) = \sum_m (|v_m|^2\ k_F^{\ 2}\ K^2/12M^2\Omega_m^{\ 5})[3|C_m|^2\ \tanh(\tfrac{1}{2}\beta\Omega_m)$$

$$- \tfrac{3}{2}\ |C_m|^2\ \beta\Omega_m\ \mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m) + \tfrac{1}{2}\ \epsilon_m^{\ 2}\ \beta^2\ \Omega_m^{\ 2}\ \mathrm{sech}^2(\tfrac{1}{2}\beta\Omega_m)\ \tanh(\tfrac{1}{2}\beta\Omega_m)\ , \tag{61}$$

which is certainly positive.

Specific Heat Anomaly

Just below the critical temperature, as $|C_m|$ tends to zero, Eq. (61) tends to the limit

$$1 + \beta\lambda_1(K, 0) = \sum_m (|v_m|^2 \beta^2 k_F^2 K^2/24M^2 \epsilon_m) \operatorname{sech}^2(\tfrac{1}{2}\beta \epsilon_m) \tanh(\tfrac{1}{2}\beta \epsilon_m)$$

$$\doteq (J \beta^2 k_F^3 K^2/48 \pi^2 M) \int_{-\infty}^{\infty} x^{-1} \operatorname{sech}^2 x \tanh x \, dx$$

$$= 7 \zeta(3) J \beta_C^2 k_F^3 K^2/24 \pi^4 M , \qquad (62)$$

which agrees with Eq. (35) in the same limit. Although $\lambda_1(K, 0)$ gives no anomalous effects, $\lambda_2(K, 0)$ differs from $\lambda_1(K, 0)$ by approximately

$$\lambda_2(K, 0) - \lambda_1(K, 0) \doteq \sum_m (|C_m v_m|^2/2\beta\Omega_m^3)[\tanh(\tfrac{1}{2}\beta\Omega_m) - \tfrac{1}{2}\beta\Omega_m \operatorname{sech}^2(\tfrac{1}{2}\beta\Omega_m)] . \qquad (63)$$

Expanding Eq. (56) in powers of $|C_n|^2$ and in powers of $\theta = (T - T_C)/T_C$, we get, close to the critical temperature,

$$\sum_m (|C_m v_m|^2/4 \epsilon_m^3)[\tanh(\tfrac{1}{2}\beta \epsilon_m) - \tfrac{1}{2}\beta \epsilon_m \operatorname{sech}^2(\tfrac{1}{2}\beta \epsilon_m)]$$

$$= -\beta \theta \sum_m \tfrac{1}{4} |v_m|^2 \operatorname{sech}^2(\tfrac{1}{2}\beta \epsilon_m) \qquad (64)$$

$$= -J M k_F \theta/2\pi^2 .$$

Close to the critical temperature, we get from Eqs. (62), (63) and (64)

$$1 + \beta_C \lambda_2(K, 0) = (J M k_F/2\pi^2)[- 2 \theta + 7 \zeta(3) \beta_C^2 k_F^2 K^2/12 \pi^2 M^2] . \qquad (65)$$

This produces a specific-heat anomaly very similar to the one predicted for the normal state. The contribution to the thermodynamic potential is $\frac{1}{2} \beta^{-1} \sum_K \log[1 + \beta \lambda_2(K, 0)]$, since we have counted graphs in such a way that we have gone round each ladder once in each direction, so that each graph has been counted twice. Comparison of Eqs. (65) and (37) shows that the anomalous specific heat just below the critical temperature is $\sqrt{2}$ times Eq. (39), with θ replaced by $-\theta$, and so it also is far too small to be detected.

Low-Temperature Paramagnetism

It has been suggested by Heine and Pippard[11] that, if proper account were taken of interactions between particles with momenta not exactly opposite, a finite spin paramagnetism might be obtained. The BCS theory gives a paramagnetism which falls off exponentially as β goes to infinity,[23] and we shall examine the perturbation-theory corrections to this result. Suppose that, when we switch on a magnetic field of strength $\mathcal{H}$, all the single-particle states with spin up lose energy $\mathcal{H} \mu_0$, and all the states with spin down gain the same amount. The same will be true of the quasi-particle states defined by Eq. (47), and they will lose or gain energy according to whether they have the label zero or one. We can do all the calculations with these altered quasi-particle energies, and then find the susceptibility by calculating $\partial^2\Omega/\partial\mathcal{H}^2$ for $\mathcal{H} = 0$. Equation (56) is altered by the magnetic field, and this gives the BCS expression for the spin susceptibility. The quantities of Eq. (48) are altered by only a small amount which falls off exponentially at low temperatures. In Eq. (54), $\mathcal{H} \mu_0$ must be subtracted from Ω_{K+m} and added to Ω_{K-m} everywhere they occur. For large β, however, this makes only exponentially small alterations, and the same is true of Eq. (60), and so there is only an exponentially small contribution to the magnetic susceptibility.

V. MORE GENERAL FORMS OF THE INTERACTION

A Criterion for Convergence

In the form of the theory developed by Bogoliubov and his collaborators,[9] the canonical transformation (47), with coefficients given by Eqs. (48) and (49), is made. The new unperturbed Hamiltonian is given by the first two sums on the right of Eq. (50), and we wish to find the condition that the sum of ladder diagrams should now converge. We know from Appendix B that the sum of ladders with momentum zero is infinite, but we require that the sum of ladders with any other momentum should converge. Since Eq. (49) in general has several sets of solutions, this requirement may enable us to determine which solutions are acceptable.

The condition for convergence is that the solution of Eq. (22) can be expanded as a power series in the coupling constant. An equivalent condition is that the equations

$$\lambda\, d_m = -\beta \sum_n (V_{mn} S_{nn} d_n + V_{m\bar{n}} S_{n\bar{n}} d_{\bar{n}}) \;,$$
$$\lambda\, d_{\bar{m}} = -\beta \sum_n (V_{\bar{m}\bar{n}} S_{\bar{n}n} d_n + V_{\bar{m}\bar{n}} S_{\bar{n}\bar{n}} d_{\bar{n}}) \qquad (66)$$

should have no eigenvalues with $|\lambda| > 1$, since the thermodynamic potential is $\frac{1}{2}\beta^{-1}$ times a sum of the possible $\log(1 - \lambda)$. The terms in Eq. (66) are understood to be functions of the total momentum $2K$ of the ladder, and of the Fourier component ν. As in Eqs. (52) and (53), $\bar{m}$ denotes a state consisting of two holes, and we neglect states consisting of a hole and a particle, on the grounds that they involve the long-range part of the interaction. The elements of the matrix S are given by Eq. (53), and $V_{\bar{m}\bar{n}} = V^*_{mn}$. For the case of total momentum close to zero and $\nu = 0$, Eq. (66) becomes

$$\lambda\, d_m = \sum_n V_{mn}(-A_n d_n + |X_n| d_n + X_n d_{\bar{n}}) \;,$$

$$\lambda\, d_{\bar{m}} = \sum_n V^*_{mn}(X^*_n d_n - A_n d_{\bar{n}} + |X_n| d_{\bar{n}}) \;, \qquad (67)$$

where

$$A_n = (2\Omega_n)^{-1} \tanh(\tfrac{1}{2}\beta\Omega_n) \;,$$

$$X_n = (C_n^2/4\Omega_n^3)[\tanh(\tfrac{1}{2}\beta\Omega_n) - \tfrac{1}{2}\beta\Omega_n \operatorname{sech}^2(\tfrac{1}{2}\beta\Omega_n)] \;. \qquad (68)$$

One solution of Eq. (67) is certainly given by $\lambda = 1$, $d_m = C_m$, $d_{\bar{m}} = -C^*_m$, since it then reduces to Eq. (49) and its complex conjugate. We require that there be no eigenvalues greater than this.

The advantage of using Eq. (67) to distinguish between different solutions of the BCS problem is that, unlike the original Eq. (49), it is a linear equation. We use it here to study the problem of more general interactions than the one we have so far studied, which is the separable S-state interaction between two particles of opposite spin.

<u>Interaction between Particles with Parallel Spin</u>

The first problem to which we apply Eq. (67) is a very artificial one. We assume that, in addition to the interaction (12) between two particles with opposite spins, there is a separable interaction

$$W_{mn} = -w^*_m w_n \qquad (69)$$

between two particles with parallel spins. This interaction is an S-state interaction, and therefore has the wrong symmetry, but the simplification produced by assuming the potential to be separable is considerable.

Suppose that the usual Bogoliubov transformation (47), pairing particles of opposite spin, is made. Then the sum of ladder graphs with total spin unity is given by an equation like Eq. (67), with V replaced by W . The solutions of this equation are given by d_m/w^*_m and $d_{\bar{m}}/w_m$ equal to constants, and there are two eigenvalues,

$$\lambda = \sum_n (A_n - |X_n|)|w_n|^2 \pm |\sum_n X_n w_n^2| \quad . \tag{70}$$

If we take both v_n and w_n to be real, the largest eigenvalue is

$$\lambda = \sum_n (w_n^2/2\Omega_n) \tanh(\tfrac{1}{2}\beta\Omega_n) \quad . \tag{71}$$

If v_n and w_n are proportional, then Eq. (71) gives an eigenvalue less than one for $v_n^2 > w_n^2$, and greater than one for $w_n^2 > v_n^2$. In the latter case, we could make a transformation which correlated particles with parallel spin, and we would then get a set of equations analogous to Eqs. (48), (49), and (71),

$$\begin{aligned} C_n/w_n &= B \quad , \\ \Omega'^2_n &= \epsilon_n^2 + C_n^2 \\ \sum_n (w_n^2/2\Omega'_n) \tanh(\tfrac{1}{2}\beta\Omega'_n) &= 1 \quad , \\ \sum_n (v_n^2/2\Omega'_n) \tanh(\tfrac{1}{2}\beta\Omega'_n) &\leq 1 \quad . \end{aligned} \tag{72}$$

If v_n and w_n are proportional, one and only one of these two sets of equations is consistent. If they are not proportional, it is possible for neither to be consistent, but it is not possible for both to be consistent. The relations which would have to be satisfied can be written as

$$
\begin{aligned}
A^2 &= \sum_n p_n \, f_n(p_n) \quad , \\
B^2 &\geqslant \sum_n q_n \, f_n(p_n) \quad , \\
B^2 &= \sum_n q_n \, f_n(q_n) \quad , \\
A^2 &\geqslant \sum_n p_n \, f_n(p_n) \quad ,
\end{aligned}
\tag{73}
$$

where we have written $\Omega_n{}^2 = \epsilon_n{}^2 + A^2 v_n{}^2$, $p_n = A^2 v_n{}^2$, $q_n = B^2 w_n{}^2$, and $f_n(x) = \frac{1}{2}(\epsilon_n{}^2 + x^2)^{-1/2} \tanh[\, \frac{1}{2} \beta(\epsilon_n{}^2 + x^2)^{1/2} \,]$. These four relations imply

$$
\sum_n (p_n - q_n)[f_n(p_n) - f_n(q_n)] \geqslant 0 \quad , \tag{74}
$$

which is impossible, since $f_n(x)$ is a monotone decreasing function of x for positive x . We have therefore proved the statement that both of the possible solutions cannot give a convergent perturbation series.

At zero temperature, we can show how the intermediate region, in which neither coupling with spins opposite nor coupling with spins parallel gives a convergent perturbation expansion in the ladder approximation, really gives rise to a more complicated coupling, even when we use a realistic Hamiltonian like Eq. (3). We take all the matrix elements of the interaction to be real, and use the trial wave function

$$
\begin{aligned}
\psi = \prod_k \, [\alpha_k &+ \beta_k(a^\dagger{}_{k,+} \, a^\dagger{}_{-k,-} + a^\dagger{}_{-k,+} \, a^\dagger{}_{k,-}) \\
&+ \gamma_k(a^\dagger{}_{k,+} \, a^\dagger{}_{-k,+} + a^\dagger{}_{k,-} \, a^\dagger{}_{-k,-}) + \delta_k \, a^\dagger{}_{k,+} \, a^\dagger{}_{-k,-} \, a^\dagger{}_{-k,+} \, a^\dagger{}_{k,-} \,] |0\rangle \; ,
\end{aligned}
$$

$$
\alpha_k{}^2 + 2\,\beta_k{}^2 + 2\,\gamma_k{}^2 + \delta_k{}^2 = 1 \quad , \tag{75}
$$

where the product is taken over all single-particle states which have positive component of momentum in a chosen direction. The expectation value of the reduced Hamiltonian, the part of Eq. (3) which includes only the interactions between pairs of particles with total momentum zero, is

$$E = \sum_m 2\epsilon_m(\beta_m^2 + \gamma_m^2 + \delta_m^2) + \sum_m \sum_n (\alpha_m \beta_m + \beta_m \delta_m) V_{mn} (\alpha_n \beta_n + \beta_n \delta_n)$$

$$+ \sum_m \sum_n (\alpha_m \gamma_m + \gamma_m \delta_m) W_{mn} (\alpha_n \gamma_n + \gamma_n \delta_n) \quad , \tag{76}$$

where the sums go over all states in momentum space. We understand that α, β, and δ for states of opposite momentum are the same, but that γ has opposite signs in the two states. One set of coefficients which gives a stationary value of Eq. (76) is $\alpha_m = x_m^2$, $\beta_m = x_m y_m$, $\gamma_m = 0$, $\delta_m = y_m^2$, where x and y are given by Eqs. (48) and (49), since Eq. (75) is then the BCS trial wave function. We make a small variation of the coefficients away from this solution of the variational problem, by making γ_m proportional to $d_m/2\Omega_m$, where d_m satisfies the equation

$$d_m = -\lambda \sum_n (V_{mn}/2\Omega_n) d_n \quad . \tag{77}$$

We have to make changes of order y^2 in α, β, and δ, but Eq. (76) is stationary with respect to changes which satisfy $\alpha^2 + 2\beta^2 + \delta^2 = 1$, and so the change in Eq. (76) depends only on the values of the γ_m. Equations (48), (49), and (77) show that the change in E is simply $(1 - \lambda) \sum_m 2 \, |\gamma_m|^2 \, \Omega_m$. The existence of an eigenvalue of Eq. (77), which must also be an eigenvalue of Eq. (67), greater than unity is a sufficient condition for a trial wave function like Eq. (75) to give a lower energy than the BCS wave function. The best solution can therefore be more complicated than the solutions considered by BCS, and the failure of convergence of the perturbation expansion in the ladder approximation seems to be a sign of this possibility.

Angle-Dependent Interactions

It is usually assumed that the solution of Eq. (49) is spherically symmetric, so that only the S-state part of V is effective. It is possible for there to be solutions which have a dependence on angle, either because of a strong interaction in some higher angular momentum state, or because the effective mass depends on angle.[24] There are many possible forms of the dependence of the solution on angle, since Eq. (49) is nonlinear, and Eq.(67) may tell us which forms are acceptable.

Suppose we have an interaction

$$
\begin{aligned}
V_{mn} &= - \sum_{\ell=0}^{\infty} (2\ell + 1) v^{*}_{m,\ell} \, v_{n,\ell} \, P_{\ell}(\cos \theta_{mn}) \\
&= -4\pi \sum_{\ell=0}^{\infty} \sum_{\mu=-\ell}^{\ell} v^{*}_{m,\ell} \, Y^{*}_{\ell\mu}(\theta_m, \phi_m) v_{n,\ell} \, Y_{\ell\mu}(\theta_n, \phi_n) \quad ,
\end{aligned}
\tag{78}
$$

where $v_{m,\ell} = \sqrt{J_\ell}$ for m within a certain distance of the Fermi surface. Equation (21) for the critical temperature is a linear equation, and it separates into angular-momentum components in such a way that C_n must be a pure spherical harmonic. If the largest of the numbers J_ℓ is J_0 , then the S-state potential alone will determine the critical temperature. Below the critical temperature, we have the usual spherically symmetric solution, and Eq. (67) is also separable. The larger eigenvalue corresponding to a solution with the angular dependence of an ℓth-order spherical harmonic is

$$
\lambda = \sum_m (v_{m,\ell}^2 / 2\Omega_m) \tanh(\tfrac{1}{2} \beta \, \Omega_m) \quad , \tag{79}
$$

and this will be less than unity if the S-state interaction is the strongest interaction.

If the S-state interaction does not dominate, and solutions of Eq. (49) other than the spherically symmetric one are involved, the situation becomes much more complicated, because of the angular dependence of Ω_m . We take, for example, a potential which acts only in the state with angular momentum ℓ , and, at zero temperature, Eq. (49) becomes

$$C_m = \sum_\mu v^*_m \Gamma_\mu Y^*_{\ell\mu}(\theta_m, \phi_m) \quad , \tag{80}$$

where

$$\Gamma_\mu = 4\pi \sum_m v_m C_m Y_{\ell\mu}(\theta_m, \phi_m) \Big/ 2(\epsilon_m^2 + |C_m|^2)^{1/2} \quad . \tag{81}$$

The dependence of $|C_m|^2$ on angle can therefore be quite complicated, although it does not depend on the magnitude of m if $v_m^2 = J$, and the integration over energies can be done by the usual methods. We assume that Ω_m is axially symmetric, since this is the simplest case. Only one of the numbers Γ_μ is nonzero, and the equation that must be satisfied if Eqs. (80) and (81) hold is

$$4\pi \sum_m |v_m Y_{\ell\mu}(\theta_m, \phi_m)|^2 \Big/ 2(\epsilon_m^2 + |C_m|^2)^{1/2} = 1 \quad . \tag{82}$$

With this solution for C_m , we look for a solution of Eq. (67) of the form

$$\begin{aligned} d_m &= v^*_m \Delta_{\mu'} Y^*_{\ell\mu'}(\theta_m, \phi_m) \quad , \\ d_{\bar{m}} &= v_m E_{\mu'} Y_{\ell,2\mu-\mu'}(\theta_m, \phi_m) \quad , \end{aligned} \tag{83}$$

and this satisfies Eq. (67) if

$$\lambda \Delta_{\mu'} = 4\pi \sum_n \Big\{ [\, |v_n Y_{\ell\mu'}(\theta_n, \phi_n)|^2 / 2\Omega_n - |v_n Y_{\ell\mu'}(\theta_n, \phi_n) C_n|^2 / 4\Omega_n^3] \Delta_{\mu'}$$

$$- [v_n^2 C_n^2 Y_{\ell\mu'}(\theta_n, \phi_n) Y_{\ell,2\mu-\mu'}(\theta_n, \phi_n) / 4\Omega_n^3] E_{\mu'} \Big\}$$

$$\lambda E_{\mu'} = 4\pi \sum_n \Big\{ [\, |v_n Y_{\ell,2\mu-\mu'}(\theta_n, \phi_n)|^2 / 2\Omega_n - |v_n Y_{\ell,2\mu-\mu'}(\theta_n, \phi_n) C_n|^2 / 4\Omega_n^3] E_{\mu'}$$

$$- [v_n^{*2} C_n^{*2} Y^*_{\ell,2\mu-\mu'}(\theta_n, \phi_n) Y^*_{\ell,\mu'}(\theta_n, \phi_n) / 4\Omega_n^3] \Delta_{\mu'} \Big\} . \qquad (84)$$

There are also solutions of Eq. (67) which have a different value of ℓ, but we do not consider those here. If we take $v_m = \sqrt{J}$ for $k_F - w < m < k_F + w$, we can perform the integrations over energy which occur in Eqs. (83) and (84). The integrals we need are

$$\int_{-wk_F/M}^{wk_F/M} \frac{d\epsilon}{2(\epsilon^2 + |C|^2)^{1/2}} \doteq \log(2wk_F/M|C|)$$

$$(85)$$

$$\int_{-\infty}^{\infty} \frac{|C|^2 \, d\epsilon}{4(\epsilon^2 + |C|^2)^{3/2}} \doteq \frac{1}{2} .$$

The eigenvalues of Eq. (84) can now be determined by comparing the angular integrals which occur in Eqs. (82) and (84).

We take the simplest example, which is a P-state interaction. For $\mu = 0$, there is a solution of Eq. (84) with $\Delta_1 = E_1$, and the corresponding eigenvalue is

$$\lambda = \frac{\frac{3}{4}\int_{-1}^{1} \sin^2\theta \log(A/|\cos\theta|)\, d(\cos\theta)}{\frac{3}{2}\int_{-1}^{1} \cos^2\theta \log(A/|\cos\theta|)\, d(\cos\theta)} \tag{86}$$

$$= (\log A + \tfrac{4}{3})/(\log A + \tfrac{1}{3}) ,$$

which is greater than unity, since log A is positive. The solution with $\mu = 0$ is not allowed, so we try $\mu = 1$. There are then solutions of Eq. (84) with $E_0 = 0$ or with $E_{-1} = 0$, and both have eigenvalue unity. The eigenvalue for $\mu' = 0$ is

$$\lambda = \frac{\frac{3}{2}\int_{-1}^{1} \cos^2\theta\, [\log(A/\sin\theta) - \frac{1}{2}]\, d(\cos\theta)}{\frac{3}{4}\int_{-1}^{1} \sin^2\theta \log(A/\sin\theta)\, d(\cos\theta)} \tag{87}$$

$$= (\log A + \tfrac{5}{6} - \log 2)\,(\log A + \tfrac{5}{6} - \log 2) = 1 .$$

In this way, we have shown which of the axially symmetric solutions of the BCS equations is possible with a pure P-state interaction.

VI. CONCLUSIONS

We have shown that there is a close connection between the BCS theory of superconductivity and the ladder diagrams of perturbation theory, although we have not been able to show which graphs of perturbation theory the BCS theory takes into account. It seems likely that, if a wider class of diagrams than the ladder diagrams were taken into account in the normal state, the critical temperature would be different. The Bogoliubov transformation almost solves the convergence problem below the critical temperature, but not quite, since perturbation theory applied to the reduced, transformed Hamiltonian does not give a convergent result if only the terms independent of the extent of the system are included. Convergence of the other ladder terms in the perturbation series, which involve the residual terms of the Hamiltonian, seems to be a useful criterion for the best BCS solution, although nothing general has been proved about this. It is a simple criterion, because it involves only linear equations. This work reveals more that is new about perturbation theory than about superconductivity theory, but may provide a useful additional tool for the latter study.

I should like to thank Dr. A. E. Glassgold and Dr. B. J. Mottelson for some useful discussions.

APPENDIX A

The ring diagrams look very like the ladder diagrams, but half the lines are reversed, so that there is one particle line and one hole line joining two successive vertices, instead of two particle lines or two hole lines. We can define a propagator $R_{nm}(q; t - t')$ as the sum of graphs like those shown in Fig. 3. They start with the state $m + q$, $m - q$ at t, and end with the states $n + q$, $n - q$ at t'. Since one line is reversed in direction, it is now the relative momentum rather than the total momentum which is conserved by the interaction. The propagator now satisfies an equation analogous to Eq. (6), which is

$$R_{nm}(q; t - t') = S(n + q, t - t') \, S(n - q, t' - t)\delta_{nm}$$
$$+ \sum_p \int_0^\beta S(n + q, t'' - t') \, S(n - q, t' - t'')V(q)R_{pm}(q; t - t'')dt'' \, , \tag{A1}$$

where we have assumed the matrix element of the interaction to depend only on the momentum transfer $2q$. We have written R as a function of $t - t'$, since this can be proved in the same way as it was proved for L. It satisfies the same periodicity condition

$$R_{nm}(q; t - \beta) = R_{nm}(q; t) \, . \tag{A2}$$

In the same way as in Sec. II, we write R as a Fourier series, and solve Eq. (A1) to get

$$\beta \sum_m \sum_n R_{nm}(q, \nu)V(q) = X(q, \nu)[1 + X(q, \nu)]^{-1} \, , \tag{A3}$$

where

$$X(q, \nu) = \sum_n \frac{1}{2} \beta \quad V(q)[\beta(\epsilon_{n+q} - \epsilon_{n-q}) - 2 \pi i \nu]^{-1}$$

$$\times [\tanh(\frac{1}{2} \beta \epsilon_{n+q}) - \tanh(\frac{1}{2} \beta \epsilon_{n-q})] . \tag{A4}$$

The contribution to Ω , excluding the first-order term, is

$$\beta^{-1} \sum_{\nu=-\infty}^{+\infty} \left\{ \log[1 + X(q, \nu)] - X(q, \nu) \right\} , \tag{A5}$$

which is very similar to the expression found by Montroll and Ward.[2]

APPENDIX B

We outline the proof that ladder diagrams with zero momentum still diverge after the transformation (47) has been made. It has been shown[9] that only terms in which each B_m or $B^\dagger_m$ is paired with another B_m or $B^\dagger_m$ do not vanish, and that these terms give a contribution of order unity to the thermodynamic potential. We use this fact by introducing, as well as the two-particle states m and the two-hole states $\bar{m}$, a third kind of state $\tilde{m}$ which has no particles or holes in it. Equations (50) and (51) show

$$V_{m\tilde{n}} = v^*_m v_n (C_n/2\Omega_n) \tanh(\tfrac{1}{2}\beta\Omega_n) \quad ,$$

$$V_{\bar{m}\tilde{n}} = v_m v^*_n (C^*_n/2\Omega_n) \tanh(\tfrac{1}{2}\beta\Omega_n) \quad ,$$

$$V_{\tilde{m}\tilde{n}} = -2\,\mathrm{Re}[(v^*_m v_n C^*_m C_n/4\Omega_m\Omega_n) \tanh(\tfrac{1}{2}\beta\Omega_m) \tanh(\tfrac{1}{2}\beta\Omega_n)] \quad . \tag{B1}$$

The sum of this and Eq. (52) is still the sum of two separable potentials. The propagator of the state $\tilde{m}$ is unity, and therefore the matrix $S^{ij}(0, 0)$ is given by the limit $K = 0$ of Eq. (54) together with the additional term

$$R_{ij}(0, 0) = \sum_m \tanh^2(\tfrac{1}{2}\beta\Omega_m) \begin{pmatrix} |C_m v_m|^2/4\Omega_m^2 & C_m^2 v_m^2/4\Omega_m^2 \\ C^{*2}_m v^{*2}_m/4\Omega_m^2 & |C_m v_m|^2/4\Omega_m^2 \end{pmatrix} . \tag{B2}$$

The matrix here is the same as the second matrix on the right of Eq. (54) in the limit $K = 0$, and one eigenvalue of the sum of Eqs. (54) and (B2) is still $-\beta^{-1}$.

REFERENCES

1. T. Matsubara, Progr. Theoret. Phys. (Kyoto) 14, 351 (1955).

2. E. W. Montroll and J. C. Ward, Physics of Fluids 1, 55 (1958).

3. C. Bloch and C. De Dominicis, Nuclear Phys. 7, 459 (1958).

4. J. Goldstone, Proc. Roy. Soc. (London) A239, 267 (1957).

5. Bardeen, Cooper, and Schrieffer, Phys. Rev. 108, 1175 (1957).

6. L. N. Cooper, Phys. Rev. 104, 1189 (1956).

7. P. W. Anderson, Phys. Rev. 112, 1900 (1958).

8. N. N. Bogoliubov, Zhur. Eksptl. i Teoret. Fiz. 34, 73 (1958) [translation: Soviet Physics JETP 7, 51 (1958)].

9. Bogoliubov, Zubarev, and Tserkovnikov, Doklady Akad. Nauk S.S.S.R. 117, 788 (1957) [translation: Soviet Physics (Doklady) 2, 535 (1958)].

10. D. J. Thouless, Bull. Am. Phys. Soc. 4, 225 (1959) and Phys. Rev. (to be published).

11. V. Heine and A. B. Pippard, Phil. Mag. 3, 1046 (1958).

12. R. E. Prange, Bull. Am. Phys. Soc. 4, 225 (1959), and to be published. See also A. J. Cantor and P. C. Martin, Bull. Am. Phys. Soc. 3, 202 (1958), and C. Bloch and C. De Dominicis, Nuclear Phys. 10, 509 (1959).

13. J. Goldstone, private communication; P. C. Martin and C. De Dominicis, unpublished.

14. L. D. Landau and E. M. Lifshitz, Statistical Physics (John Wiley and Sons, New York, 1958), p. 453.

15. Reference 14, p. 157.

16. See, for example, reference 4.

17. For the properties of integral equations used here, see R. Courant and D. Hilbert, Methods of Mathematical Physics, vol. 1 (Interscience Publishing Company, New York, 1953), pp. 112-163.

18. M. Gell-Mann and K. A. Brueckner, Phys. Rev. 106, 364 (1957).

19. N. N. Bogoliubov, Doklady Akad. Nauk S.S.S.R. 119, 52 (1958) [translation: Soviet Physics (Doklady) 3, 279 (1958)].

20. Reference 14, p. 165.

21. J. G. Daunt, Progress in Low Temperature Physics, vol. 1 (Interscience Publishing Company, New York, 1955), p. 211.

22. Fairbank, Buckingham, and Kellers, Proceedings of the 5th International Conference on Low Temperature Physics and Chemistry (University of Wisconsin Press, Madison, 1958), p. 50.

23. K. Yosida, Phys. Rev. 110, 769 (1958).

24. L. N. Cooper, Phys. Rev. Letters 3, 17 (1959).

CAPTION FOR FIGURES

Fig. 1. Some typical ladder diagrams.

Fig. 2. Some diagrams contributing to the ladder propagator $L_{nm}(K; t', t)$.

Fig. 3. Some diagrams contributing to the ring propagator $R_{nm}(q; t - t')$.

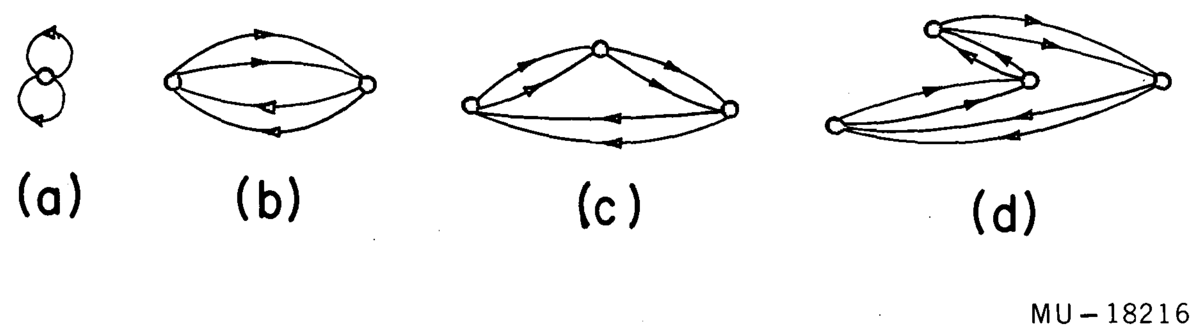

Fig. 1

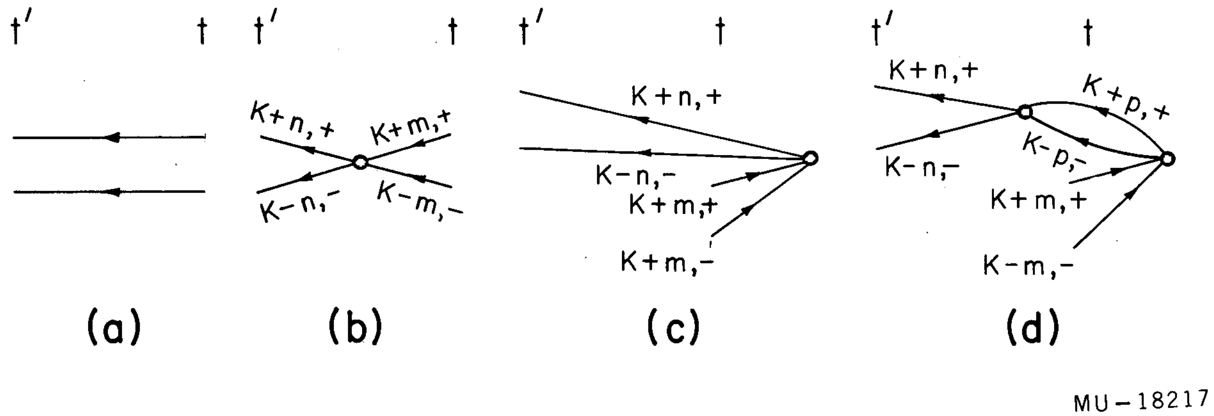

Fig. 2.

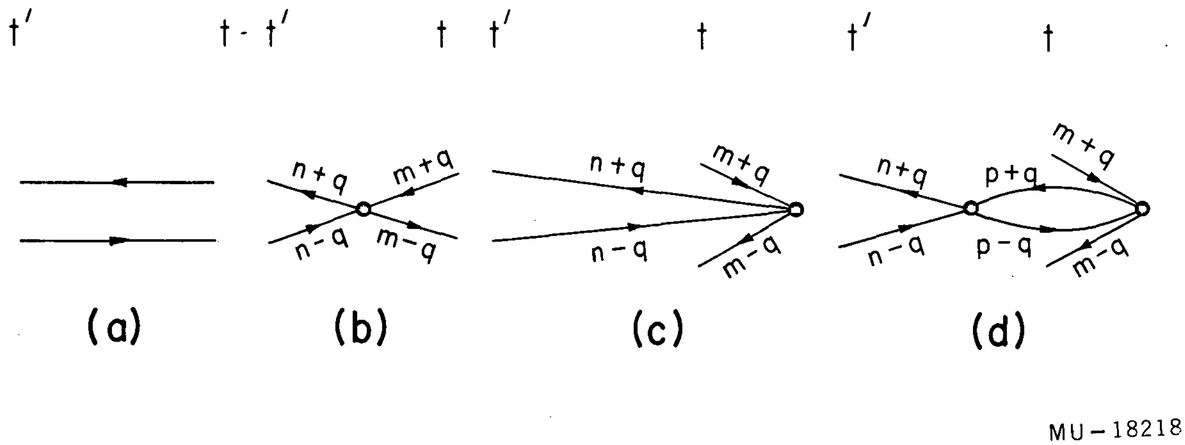

Fig. 3