Reformulation of the Theory of Pairing Correlations

MICHEL BARANGER*†
Centre d'Etudes Nucléaires de Saclay, Gif-sur-Yvette, France
(Received 26 December 1962)

A variational approach to the ground state of a many-fermion system is developed with the trial function of condensed pairs proposed by Blatt. The calculation of matrix elements is done by a diagrammatic method which is much simpler than Blatt's. The method is susceptible of generalization to higher correlations. It is also shown that the theory is identical with Hartree-Bogolyubov theory. The Hartree-Bogolyubov equations are derived from the variational point of view.

1. INTRODUCTION

THE motivation for the present work originates in nuclear physics. The problem is to find a formulation of the theory of pairing correlations which is immediately generalizable to α-particle correlations. This could be useful in the theory of light nuclei, for which neutron-proton correlations are just as important as correlations between nucleons of the same kind. It is only for heavy nuclei that one can restrict oneself to correlations of the pairing type. There, the lack of overlap between neutron and proton wave functions near the Fermi surface decreases the strength of neutron-proton correlations. As a consequence, recent successes based on the BCS theory[1] are limited to heavy nuclei.

The necessary reformulation has been found. The generalization to α particles, and to many other types of correlations, is straightforward in principle. However, whereas pairing correlations lead to a formalism which is fairly simple, the application to α-particle correlations is beset by calculational difficulties which have not yet been resolved. Since the new formulation of the theory of pairing correlations is quite interesting in its own right, it is made the only subject of the present paper.

The theory of pairing correlations, as it is usually developed, possesses two distinct stages. First, there is the BCS theory,[1] in which the quantum numbers of the two particles that are being paired are given *a priori*, which is possible because these quantum numbers are precisely those that are conserved by the interaction. The second stage is Hartree-Bogolyubov theory[2] (to be abbreviated in the following as HB theory). It does not use a conservation law; the manner of pairing is unknown at the start and left completely arbitrary, to be determined by minimization of the energy. The theory we want to reformulate is HB, not BCS, because if we are trying to associate four particles instead of two, even a conservation law does not help much in fixing the four sets of quantum numbers. They are still to a large extent arbitrary, as they are in HB theory. Attempts to generalize BCS to α-particle correlations have been made and have failed.

Besides BCS and HB theory there exists a different point of view on the theory of pairing correlations which is that of Blatt and collaborators.[3] It turns out that the Blatt approach is precisely the reformulation that we need in order to make the generalization to α particles. Therefore, the present work is partly devoted to rederiving the results of Blatt in a way quite different from his and which we feel is much simpler. We certainly do not wish to claim that the present paper covers all aspects of the work of Blatt and collaborators. In the work of Blatt and Matsubara,[4] in particular, the emphasis is on statistical mechanics, evaluation of the partition function with various assumptions concerning correlations, etc. We do not touch this topic and limit ourselves to the study of the ground state with pairing correlations of the HB type.

Let us first recall the analogous reformulation of BCS theory,[5] which is much simpler than the HB case. Start from the BCS wave function

$$\prod_k (u_k + v_k c_k^\dagger c_{\bar{k}}^\dagger)\,|0\rangle. \tag{1}$$

The product is over half the total number of states, the index k constitutes a complete label for a state, and $\bar{k}$ represents the state which is paired with k (e.g., opposite momentum and spin). With a change of normalization, this can also be written

$$\prod_k [1 + (v_k/u_k) c_k^\dagger c_{\bar{k}}^\dagger]\,|0\rangle. \tag{2}$$

* National Science Foundation Senior Postdoctoral Fellow, 1961–62.

† Present address: Carnegie Institute of Technology, Pittsburgh, Pennsylvania.

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957); N. N. Bogolyubov, Nuovo Cimento **7**, 794 (1958); J. G. Valatin, *ibid.* **7**, 843 (1958).

[2] Sometimes also called generalized Hartree-Fock or generalized BCS, Hartree-Bogolyubov theory originates in the paper of N. N. Bogolyubov, Uspekhi Fiz. Nauk **67**, 549 (1959) [translation: Soviet Phys.—Usp. **67** (*2*), 236 (1959)]. For later developments, see Y. Nambu, Phys. Rev. **117**, 648 (1960); M. Baranger, Phys. Rev. **122**, 992 (1961); J. G. Valatin, *ibid.* **122**, 1012 (1961); C. Bloch, course given at the Centre d'Etudes Nucléaires de Saclay, 1961–62 (unpublished). See also M. Baranger in *Cargèse Summer School Lectures, 1962*, edited by M. Lévy (W. A. Benjamin, New York, 1963).

[3] J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **24**, 851 (1960). Earlier references are given there. We thank Dr. Blatt for sending reprints of his papers.

[4] J. M. Blatt and T. Matsubara, Progr. Theoret. Phys. (Kyoto) **20**, 533 (1958); T. Matsubara and J. M. Blatt, *ibid.* **23**, 451 (1960).

[5] B. Bayman, Nucl. Phys. **15**, 33 (1960); J. M. Blatt, Progr. Theoret. Phys. (Kyoto) **23**, 447 (1960); F. J. Dyson (unpublished); A. Katz, Nucl. Phys. **18**, 177 (1960); J. R. Schrieffer has also informed us that wave function (4) was very much on his mind in the early days of BCS theory, though a different point of view was adopted in the BCS paper, reference 1.

It is then easily seen that the projection of wave function (2) on the subspace of $2N$ particles or N pairs is exactly

$$(N!)^{-1}[\sum_k (v_k/u_k)c_k^\dagger c_{\bar{k}}^\dagger]^N|0\rangle. \quad (3)$$

Therefore, instead of starting from wave function (1), one may start with

$$(\sum_k \varphi_k c_k^\dagger c_{\bar{k}}^\dagger)^N|0\rangle. \quad (4)$$

Bayman[5] has rederived the main results of BCS theory with trial function (4). The two starting points are equivalent when N is large, provided one sets

$$\varphi_k = v_k/u_k. \quad (5)$$

But there is a slight difference in spirit. Namely, in wave function (1) one mixes up numbers of particles from the start, while in Bayman's approach the number of particles is kept fixed as long as possible, and it is only as the result of mathematical approximations that things look at the end as though one had mixed numbers of particles.

The treatment of the HB case, in which no simple pairing is known *a priori*, is obtained by analogy with the above. The trial function, Blatt's wave function, is taken to be

$$|\Phi_N\rangle = (\sum_{\alpha\beta} \varphi_{\alpha\beta}\, c_\alpha^\dagger c_\beta^\dagger)^N|0\rangle, \quad (6)$$

where α, β, $\cdots$ is an arbitrary representation and $\varphi_{\alpha\beta}$ is antisymmetric in its two indices. In other words, Blatt takes an arbitrary pair wave function and condenses all particles into pairs in this same state. He goes on to study the properties of the superconducting ground state represented by this wave function.[3] In particular, he gives a very lucid discussion of gauge invariance and of the Meissner effect. Unfortunately, the mathematics which Blatt has to go through, and which are published in a separate paper,[6] are extremely complicated. It also seems very difficult to extend them to α-particle correlations. Both defects are remedied in the present work. Whereas Blatt uses operator algebra, we have used diagrammatic techniques, which are quite well suited for the purpose.

The net result of our work is that Blatt's theory and HB theory are identical. This should have been obvious from the beginning for the following reasons. Bloch and Messiah[7] have proved a theorem saying that given any antisymmetric two-body state $\varphi_{\alpha\beta}$, there exists a change of representation of the one-body labels, i.e., a one-body unitary transformation, that puts the two-body function in the specially simple form shown in Fig. 1. In this new representation, the one-body states are paired in such a way that $\varphi_{\alpha\beta}$ exists only if α and β are paired states. We may designate two paired states by k and $\bar{k}$. Then, in the new representation, wave

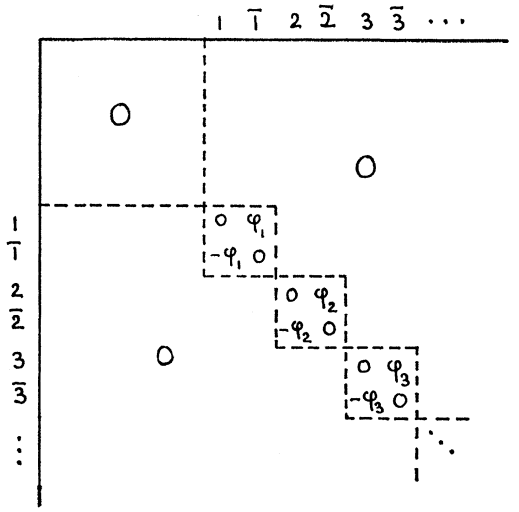

FIG. 1. Canonical form of a two-fermion wave function according to reference 7.

function (6) becomes identical with (4). In the same paper Bloch and Messiah show that the ground state of HB theory can always be written in form (1), again with suitable choice of the one-body representation. Therefore, one sees in this representation, as one did in the BCS case, that Blatt's wave function is the projection of the HB wave function on the subspace of $2N$ particles, and the two must give identical results in the limit of large N. This equivalence between Blatt's wave function and wave function (1) was not recognized by Blatt in reference 6. To understand how this confusion may have arisen, it is good to realize that HB theory is not the same as BCS theory plus a simple change of representation, even though the ground state of HB theory may be made to look like the BCS ground state.[8] What is missing in HB theory is the conservation law which simplifies BCS considerably.

In spite of the equivalence just noted, it is highly interesting to develop the complete theory with Blatt's point of view. First, it provides an attractive alternative to the usual HB procedure, an alternative which is neither harder nor longer. Second, it is susceptible of generalization to α-particle correlations. The usual point of view, with its emphasis on quasiparticle operators that are linear combinations of the original field operators, is powerless there. The trial function for the α-particle theory should obviously be

$$(\sum_{\alpha\beta\gamma\delta} \varphi_{\alpha\beta\gamma\delta}\, c_\alpha^\dagger c_\beta^\dagger c_\gamma^\dagger c_\delta^\dagger)^N|0\rangle. \quad (7)$$

Everything that we shall prove in Secs. 3 and 4 about the simpler trial function (6) can be immediately generalized to this one. The results are not simple because, unlike (6), wave function (7) leads to an

[6] J. M. Blatt, J. Australian Math. Soc. **1**, 465 (1960).
[7] C. Bloch and A. Messiah, Nucl. Phys. **39**, 95 (1962). The same theorem has also been proved by B. Zumino, J. Math. Phys. **3**, 1055 (1962).

[8] The differences between the two theories have been especially emphasized by Valatin in the paper cited under reference 2.

infinite number of irreducible diagrams. But, at least, the formalism exists.

In Sec. 2 we summarize briefly the results of conventional HB theory. In Sec. 3 we show how to calculate expectation values of operators with Blatt's trial function and find that the outcome is the same as that of Sec. 2. The argument depends critically on an assertion concerning the constancy or near-constancy of the function $F(2N)$ defined by Eq. (21), provided φ is suitably normalized. This assertion is examined critically and proved in Sec. 4. After this point, HB theory and Blatt's theory can be considered identical. However, they are not yet in a form suitable for practical applications. For the latter, it is useful to derive a set of equations analogous to the Hartree-Fock equations, namely, the Hartree-Bogolyubov equations. Although this belongs properly with conventional HB theory, the derivation of these equations from the variational point of view does not seem to exist in the literature: We give it in Sec. 5.

2. HARTREE-BOGOLYUBOV THEORY

We consider a general Hamiltonian with one-body and two-body parts,

$$H=\sum_{\alpha\gamma} T_{\alpha\gamma}\, c_\alpha^\dagger c_\gamma+\tfrac{1}{4}\sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta}\, c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma. \quad (8)$$

The c's are Fermion field operators. The coefficients T and V have the following symmetry properties

$$T_{\alpha\gamma}=T_{\gamma\alpha}^*, \quad (9)$$

$$V_{\alpha\beta\gamma\delta}=V_{\gamma\delta\alpha\beta}^*, \quad (10)$$

$$V_{\alpha\beta\gamma\delta}=-V_{\beta\alpha\gamma\delta}=-V_{\alpha\beta\delta\gamma}=V_{\beta\alpha\delta\gamma}. \quad (11)$$

These antisymmetry relations mean that the exchange term is already included together with the direct term in the interaction.

The reader is referred to reference 2 for details of Hartree-Bogolyubov theory. The idea is to make the HB ground state $|\Psi\rangle$ similar to an independent particle state, in the sense that it is defined by

$$a_i|\Psi\rangle=0, \quad (12)$$

the a_i's being a complete set of Fermion operators called quasiparticle operators. These a operators are arbitrary linear combinations of the original c operators, with the possibility of mixing creation and annihilation operators, such as for instance

$$a_i^\dagger=\sum_\alpha (A_{\alpha i} c_\alpha^\dagger+B_{\alpha i} c_\alpha). \quad (13)$$

Because of this mixing, the quasiparticles are linear combinations of particles and holes and the state $|\Psi\rangle$ defined by (12) does not contain a definite number of particles. The following expectation values are therefore all nonzero:

$$\langle\Psi|c_\alpha^\dagger c_\gamma|\Psi\rangle=\rho_{\gamma\alpha}=\rho_{\alpha\gamma}^*, \quad (14a)$$

$$\langle\Psi|c_\beta c_\alpha|\Psi\rangle=\kappa_{\beta\alpha}=-\kappa_{\alpha\beta}, \quad (14b)$$

$$\langle\Psi|c_\alpha^\dagger c_\beta^\dagger|\Psi\rangle=\kappa_{\beta\alpha}^*=-\kappa_{\alpha\beta}^*. \quad (14c)$$

Since $|\Psi\rangle$ is of the independent-particle type, all other expectation values can be computed in terms of the two basic ones ρ and κ by application of Wick's theorem. This is true in particular of $\langle\Psi|c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma|\Psi\rangle$ which appears in the expectation value of H, and one finds

$$H_0=\langle\Psi|H|\Psi\rangle=\sum_{\alpha\gamma} T_{\alpha\gamma}\rho_{\gamma\alpha}+\tfrac{1}{2}\sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta}\rho_{\gamma\alpha}\rho_{\delta\beta} +\tfrac{1}{4}\sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta}\kappa_{\beta\alpha}^*\kappa_{\delta\gamma}. \quad (15)$$

Now, one seeks to determine $|\Psi\rangle$ by minimizing H_0. It is important to realize that determining $|\Psi\rangle$ does not mean determining the coefficients A and B of linear transformation (13). Indeed, definition (12) shows that one could perform a unitary transformation on the annihilation operators a_i without changing $|\Psi\rangle$. This is why A and B themselves do not appear in (15), but only some special combinations of them, ρ and κ, which are invariant under said unitary transformation. However, because of the way they have been obtained, ρ and κ are not completely arbitrary but obey the following matrix relations (the asterisk denotes the complex-conjugate matrix):

$$\rho^2-\kappa\kappa^*=\rho, \quad (16a)$$

$$\rho\kappa-\kappa\rho^*=0. \quad (16b)$$

In addition, since $|\Psi\rangle$ mixes numbers of particles, one must make sure in the minimization that the average number of particles stays equal to what one wants, e.g., $2N$, hence,

$$\mathrm{Tr}\rho=2N. \quad (17)$$

Finally, the HB problem is that of minimizing H_0 [Eq. (15)] in which the parameters ρ and κ are restricted by supplementary conditions (16) and (17).

3. CALCULATION OF EXPECTATION VALUES WITH BLATT'S TRIAL FUNCTION

Since wave function (6) is not normalized, it is necessary to calculate $\langle\Phi_N|\Phi_N\rangle$, $\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle$, and $\langle\Phi_N|c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma|\Phi_N\rangle$ in order to find the average energy. Let each pair $(\alpha\beta)$ be included only once in the sum of Eq. (6). The order of the two members of the pair makes no difference since $\varphi_{\beta\alpha}=-\varphi_{\alpha\beta}$.

First consider

$$\langle\Phi_N|\Phi_N\rangle=\langle 0|(\textstyle\sum_{\alpha\beta}\varphi_{\alpha\beta}^* c_\beta c_\alpha)^N(\sum_{\gamma\delta}\varphi_{\gamma\delta} c_\gamma^\dagger c_\delta^\dagger)^N|0\rangle. \quad (18)$$

This is calculated by Wick's theorem in the well-known manner, i.e., it is the sum of all possible sets of contractions. As usual, each contraction will be represented by a line labeled by the state in question. Figure 2 shows a typical term of (18). Each term consists of a

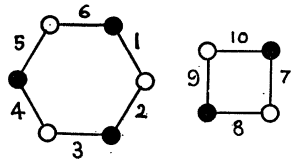

FIG. 2. A particular set of closed chains arising on the right-hand side of Eq. (18).

number of closed chains. Each chain has an even number of lines. White vertices represent the wave function φ; black vertices stand for φ^*. The total number of wave functions of either type is $2N$.

To make things clearer, we show for instance how the six-sided chain of Fig. 2 arises. In one of the N factors $\sum_{\gamma\delta}\varphi_{\gamma\delta}c_\gamma^\dagger c_\delta^\dagger$ occurring at the right in (18), we pick a term $\varphi_{12}c_1^\dagger c_2^\dagger$. Operator $c_2^\dagger$ gets contracted with a c_2 occurring in one of the N factors at the left. This c_2 is part of a term $\varphi_{23}^* c_3 c_2$. Operator c_3 gets contracted with a $c_3^\dagger$ at the right. This $c_3^\dagger$ is itself associated with a $c_4^\dagger$, which contracts with a c_4 at the left, which c_4 is associated with a c_5, etc., until we happen upon an operator c_6 which is associated with c_1. Then, we have a chance of closing the chain by contracting this c_1 with the $c_1^\dagger$ of the original term. To find the resulting sign, we write all operators in proper order, namely,

$$(\varphi_{61}^* c_1 c_6)(\varphi_{45}^* c_5 c_4)(\varphi_{23}^* c_3 c_2)(\varphi_{12} c_1^\dagger c_2^\dagger) \times(\varphi_{34} c_3^\dagger c_4^\dagger)(\varphi_{56} c_5^\dagger c_6^\dagger). \quad (19)$$

We move c_1 to the immediate left of $c_1^\dagger$, which produces a minus sign. Then, all contractions can be made without additional minus signs, so that the contribution of this particular chain is

$$-\varphi_{12}\varphi_{23}^*\varphi_{34}\varphi_{45}^*\varphi_{56}\varphi_{61}^*. \quad (20)$$

In other words, one goes around a chain in either direction, writing down the wave functions and the states in the order in which they occur, and one puts a minus sign in front of the whole.

The next problem is that of determining how many times a given term of (18), i.e., a given set of chains, occurs. If all states of the set are different, the number of times is obviously $(N!)^2$ because each elementary constituent such as $\varphi_{12}c_1^\dagger c_2^\dagger$ is chosen out of a product of N identical factors. One may, if one wants, include only sets whose states are all different, since the product of two identical Fermi operators is zero. However, doing so would make calculations very complicated and would defeat the purpose of the diagrammatic method. It is much easier to allow states to be repeated, i.e., to ignore the Pauli principle, since terms with repeated states will cancel each other mutually in the end. Naturally, adjacent states are always different. Once repeating of states has been allowed, it becomes necessary to re-examine a little more carefully the frequency of occurrence of a given set in the total.

To do this, imagine that all the φ's occurring in a given set of chains have been numbered from 1 to N, and similarly that the φ^*'s have been numbered from 1 to N. The number associated with each wave function is supposed to indicate the factor from which it is coming, among the N identical factors in (18). The number of distinct ways to do the numbering is the number of times the set occurs. If the set has no two positions alike, that is, if the relationship of each wave function to the whole is unique, then the number of numberings is $(N!)^2$. This can happen even though some states are repeated, either in the same chain or in different chains. But if the set of chains possesses a symmetry group, then the number of distinct numberings is decreased. There are two possible kinds of symmetries. First, a given chain c may remain identical to itself through a group of p_c circular permutations. This decreases the number of numberings by a factor p_c. We shall incorporate this factor in the definition of the contribution R_c of chain c to the set of contractions, i.e., R_c is calculated as in (20) with an extra division by p_c. The other kind of symmetry is that a given chain, say c, may occur m_c times in the set. This decreases the number of numberings by a factor $m_c!$, which we shall write explicitly in the contribution of the set. We also define n_c, the order of chain c, as being the total number of φ's plus φ^*'s in it. Then,

$$\langle\Phi_N|\Phi_N\rangle/(N!)^2=\sum\prod_c[(R_c)^{m_c}/m_c!]=F(2N), \quad (21)$$

this being the definition of the function F. In (21), the sum is over all possible sets of closed chains such that

$$\sum_c m_c n_c=2N, \quad (22)$$

the product is over all the *different* chains occurring in the set, and so is the sum in (22).

Now consider the matrix element

$$\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle =\langle 0|(\sum_{\gamma\delta}\varphi_{\gamma\delta}^* c_\delta c_\gamma)^N c_\alpha^\dagger c_\beta(\sum_{\epsilon\zeta}\varphi_{\epsilon\zeta}c_\epsilon^\dagger c_\zeta^\dagger)^N|0\rangle. \quad (23)$$

Again, this is calculated as the sum of all possible sets of contractions, but now each set contains one open chain having α and β as end states. For instance, the combination of operators

$$(\varphi_{5\alpha}^* c_\alpha c_5)(\varphi_{34}^* c_4 c_3)(\varphi_{12}^* c_2 c_1)c_\alpha^\dagger c_\beta(\varphi_{\beta 1}c_\beta^\dagger c_1^\dagger) \times(\varphi_{23}c_2^\dagger c_3^\dagger)(\varphi_{45}c_4^\dagger c_5^\dagger), \quad (24)$$

when contracted, gives rise to the chain of Fig. 3, whose contribution is

$$-\varphi_{\beta 1}\varphi_{12}^*\varphi_{23}\varphi_{34}^*\varphi_{45}\varphi_{5\alpha}^*. \quad (25)$$

Call i such an open chain, n_i its order, $R_{\beta\alpha}{}^i$ its contribution as in (25). Obviously, it occurs only once in a set and does not have any invariance group of circular permutations. Then,

$$\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle/(N!)^2=\sum_i R_{\beta\alpha}{}^i F(2N-n_i), \quad (26)$$

where the sum is over all possible open chains starting in β with a φ and ending in α with a φ^*.

Finally, take $\langle\Phi_N|c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma|\Phi_N\rangle$. This will contain two open chains, but there are three possible ways of forming them. It may be two even chains like that of Fig. 3, namely, $(\gamma\cdots\alpha)(\delta\cdots\beta)$ or $(\delta\cdots\alpha)(\gamma\cdots\beta)$, the

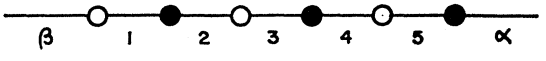

FIG. 3. An open chain such as this arises, together with a set of closed chains, in each term of the right-hand side of Eq. (23).

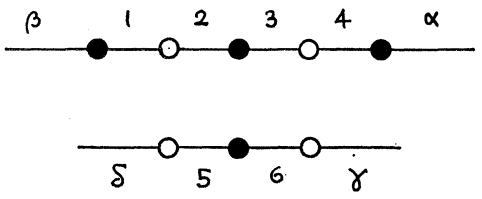

FIG. 4. Two odd open chains, together with a set of closed chains, form one of three possibilities for a term in the expansion of a two-body matrix element.

second possibility being the exchange term. But it may also be two odd chains as in Fig. 4. The chains of Fig. 4 come from the combinations of operators

$$(\varphi_{4\alpha}{}^{*}c_{\alpha}c_{4})(\varphi_{23}{}^{*}c_{3}c_{2})(\varphi_{\beta 1}{}^{*}c_{1}c_{\beta})c_{\alpha}{}^{\dagger}c_{\beta}{}^{\dagger} \times(\varphi_{12}c_{1}{}^{\dagger}c_{2}{}^{\dagger})(\varphi_{34}c_{3}{}^{\dagger}c_{4}{}^{\dagger}), \quad (27)$$

and

$$(\varphi_{56}{}^{*}c_{6}c_{5})c_{\delta}c_{\gamma}(\varphi_{\delta 5}c_{\delta}{}^{\dagger}c_{5}{}^{\dagger})(\varphi_{6\gamma}c_{6}{}^{\dagger}c_{\gamma}{}^{\dagger}), \quad (28)$$

respectively. Their contracted contributions are

$$-\varphi_{\beta 1}{}^{*}\varphi_{12}\varphi_{23}{}^{*}\varphi_{34}\varphi_{4\alpha}{}^{*}, \quad (29)$$

and

$$-\varphi_{\delta 5}\varphi_{56}{}^{*}\varphi_{6\gamma}. \quad (30)$$

We shall call $K_{\delta\gamma}{}^{i}$ the contribution of an odd chain i starting and ending with a φ, such as (30). For a chain starting and ending with φ^{*}, the contribution is the complex conjugate. One sees that

$$\langle\Phi_N|c_{\alpha}{}^{\dagger}c_{\beta}{}^{\dagger}c_{\delta}c_{\gamma}|\Phi_N\rangle/(N!)^2=\sum_{ij}R_{\gamma\alpha}{}^{i}R_{\delta\beta}{}^{j}F(2N-n_i-n_j) -\sum_{ij}R_{\delta\alpha}{}^{i}R_{\gamma\beta}{}^{j}F(2N-n_i-n_j) +\sum_{ij}K_{\beta\alpha}{}^{i*}K_{\delta\gamma}{}^{j}F(2N-n_i-n_j). \quad (31)$$

The sums in the last term are both over all odd chains starting and ending with a φ.

Now we are coming to the main point in the argument, which is that the normalization of φ can be chosen in such a way that $F(2N)$ is approximately independent of N for large N. If this is true, and provided all series converge, all the F's in Eqs. (21), (26), and (31) can be considered equal. Calling

$$\rho_{\beta\alpha}=\sum_{i}R_{\beta\alpha}{}^{i}, \quad (32a)$$

$$\kappa_{\beta\alpha}=\sum_{i}K_{\beta\alpha}{}^{i}, \quad (32b)$$

one sees that the expectation values of one- and two-body operators are given by

$$\langle\Phi_N|c_{\alpha}{}^{\dagger}c_{\beta}|\Phi_N\rangle/\langle\Phi_N|\Phi_N\rangle=\rho_{\beta\alpha}, \quad (33)$$

$$\langle\Phi_N|c_{\alpha}{}^{\dagger}c_{\beta}{}^{\dagger}c_{\delta}c_{\gamma}|\Phi_N\rangle/\langle\Phi_N|\Phi_N\rangle =\rho_{\gamma\alpha}\rho_{\delta\beta}-\rho_{\delta\alpha}\rho_{\gamma\beta}+\kappa_{\beta\alpha}{}^{*}\kappa_{\delta\gamma}. \quad (34)$$

This is formally identical to what one gets in HB theory. The expression for the average energy is identical to (15). There remains only to show that the supplementary conditions obeyed by ρ and κ are also the same as in HB theory, namely, Eqs. (16) and (17). The proof of Eq. (17) will be given in the next section, together with the proof of the main point. Now we prove Eqs. (16).

Consider the quantity $(\rho^2)_{\beta\alpha}=\sum_{\gamma}\rho_{\beta\gamma}\rho_{\gamma\alpha}$, each ρ being represented by a sum of even chains as in (32a). By thus putting two even chains end to end, one obtains a single even chain, also starting with φ and ending with φ^{*}. However, a particular single chain $R_{\beta\alpha}{}^{i}$ can be obtained in $\frac{1}{2}n_i-1$ ways, because there are that many ways to cut it into two even chains. Hence, one can write

$$(\rho^2)_{\beta\alpha}=-\sum_{i}(\tfrac{1}{2}n_i-1)R_{\beta\alpha}{}^{i}, \quad (35a)$$

in which the minus sign arises from the sign included in our definition of R. Similarly, putting two odd chains end to end also gives an even chain, but this time the number of ways to cut chain i is $\frac{1}{2}n_i$, i.e.,

$$(\kappa\kappa^{*})_{\beta\alpha}=-\sum_{i}\tfrac{1}{2}n_iR_{\beta\alpha}{}^{i}. \quad (35b)$$

Subtracting (35b) from (35a) gives (16a). Equation (16b) is proved by similar considerations leading to

$$(\rho\kappa)_{\beta\alpha}=-\sum_{i}\tfrac{1}{2}(n_i-1)K_{\beta\alpha}{}^{i}=(\kappa\rho^{*})_{\beta\alpha}. \quad (36)$$

It is also desirable to show that Eqs. (16) and (17) are the only supplementary conditions. In other words, given any ρ and κ satisfying these equations, it should be possible to find a quantity φ which, when treated by the methods of this section, reproduces the given ρ and κ. This will be shown in the next section.

4. PROOF OF THE MAIN POINT

One possible proof of the main point runs as follows. Instead of expression (21) for $F(2N)$, one can write

$$F(2N)=(2\pi i)^{-1}\oint z^{-2N-1}dz\sum\prod_{c}[(z^{n_c}R_c)^{m_c}/m_c!]. \quad (37)$$

The integral is on a contour enclosing the origin. The sum now includes all possible sets of closed chains without restriction (22). Equation (37) can also be written

$$F(2N)=(2\pi i)^{-1}\oint z^{-2N-1}dz\exp(\textstyle\sum_{c}z^{n_c}R_c)$$

$$=(2\pi i)^{-1}\oint z^{-1}dz\exp(\textstyle\sum_{c}z^{n_c}R_c-2N\ln z), \quad (38)$$

in which the sum runs over all different closed chains. For a slightly different number of particles we shall write the integral as

$$F(2N-n) =(2\pi i)^{-1}\oint z^{n-1}dz\exp(\textstyle\sum_{c}z^{n_c}R_c-2N\ln z). \quad (39)$$

This is evaluated by steepest descent. The saddle point z_0 is found by setting the derivative of the exponent

equal to zero,

$$\sum_c n_c z_0^{n_c-1} R_c - 2N/z_0 = 0. \tag{40}$$

The factor z^{n-1} is considered slowly varying compared to the exponential and is just evaluated at the saddle point. In order to make the result independent of n, we manage to make $z_0=1$, i.e., we choose the normalization of φ in a way to satisfy the condition

$$\sum_c n_c R_c = 2N. \tag{41}$$

Equation (41) is the same as supplementary condition (17). To see this, consider all possible open even chains beginning and ending with the same state, i.e., those that enter in $\mathrm{Tr}\rho$. By joining the ends of such a chain, one obtains a closed even chain. But a given closed chain c can be obtained in n_c different ways, because there are n_c places where one can break it open. Hence,

$$\sum_\alpha \sum_i R_{\alpha\alpha}{}^i = \sum_c n_c R_c. \tag{42}$$

On second thought, this might appear not quite right, since we have included in the definition of R_c the factor $1/p_c$, p_c being the number of circular permutations that leave the chain invariant. But (42) is indeed completely right, because if there is any such invariance group for chain c, the correct number of ways that it can be broken to give a different open chain is not n_c, but n_c/p_c.

As long as N is large and n is small, there seems to be little to object to in the argument at the beginning of this section. The flaw, of course, is that the summations over i and j in Eqs. (26) and (31) may not be rapidly convergent, so that values of F for numbers radically different from $2N$ will be needed. In fact, the BCS case considered a little later provides such an example of nonconvergence. There are in many-body theory many other instances of handling a nonconvergent series as though it were convergent. In the present case, however, it is possible to give a much improved derivation. We do it along the lines of Bayman's argument in reference 5.

Define

$$|\Phi(z)\rangle = \sum_N z^N |\Phi_N\rangle/N!. \tag{43}$$

One can, then, write

$$\langle\Phi(z)|\Phi(z)\rangle = \sum_N z^{2N}\langle\Phi_N|\Phi_N\rangle/(N!)^2 = \sum_N z^{2N} F(2N) = \exp\Big(\sum_c z^{n_c} R_c\Big) = F(z). \tag{44}$$

Similarly,

$$\langle\Phi(z)|c_\alpha^\dagger c_\beta|\Phi(z)\rangle = \sum_N z^{2N}\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle/(N!)^2 = F(z)R_{\beta\alpha}(z), \tag{45}$$

where

$$R_{\beta\alpha}(z) = \sum_i z^{n_i} R_{\beta\alpha}{}^i. \tag{46}$$

To go back to matrix elements with $|\Phi_N\rangle$, one uses the contour integration again, for instance

$$F(2N) = (2\pi i)^{-1} \oint z^{-2N-1} dz F(z), \tag{47}$$

which is just Eq. (38), and

$$\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle/(N!)^2 = (2\pi i)^{-1} \oint z^{-2N-1} dz \langle\Phi(z)|c_\alpha^\dagger c_\beta|\Phi(z)\rangle. \tag{48}$$

Then, the expectation value of $c_\alpha^\dagger c_\beta$ is given by

$$\langle\Phi_N|c_\alpha^\dagger c_\beta|\Phi_N\rangle/\langle\Phi_N|\Phi_N\rangle = \oint z^{-2N-1} dz\, F(z) R_{\beta\alpha}(z) \Big/ \oint z^{-2N-1} dz\, F(z). \tag{49}$$

A similar expression holds for the expectation value of $c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma$. The two integrations in (49) are done by steepest descent, but $R_{\beta\alpha}(z)$ is considered slowly varying compared to $F(z)$, so that the saddle point z_0 is the same for both integrals and expression (49) is just $R_{\beta\alpha}(z_0)$. The latter is the same as $\rho_{\beta\alpha}$, Eq. (32a), except for a change in the normalization of φ by a factor z_0. Alternatively, one may choose the normalization of φ so that $z_0=1$, i.e., so as to satisfy Eq. (41), and then $R_{\beta\alpha}(z_0)$ equals $\rho_{\beta\alpha}$.

Now, let us look at this argument in more detail. First, note that whereas in Sec. 3 all summations and multiple products contained a finite number of terms the sum over all closed chains in Eqs. (38) or (44), as well as the sum over open chains in Eq. (46), are infinite series. This is because, once the restriction to $2N$ particles has been removed, the total number of different chains (closed or open) is infinite in view of the fact that repetition of states is allowed. These series are presumably convergent if $|z|$ is small enough. The contour of integration should originally be chosen inside the circles of convergence and the extension to higher $|z|$ should be done by analytic continuation. The original divergence trouble has now been transferred to the question of whether and how the analytic continuation can be performed all the way to the saddle point.

Actually, the analytic continuation goes through without much difficulty. Define the quantity

$$\chi_{\alpha\beta} = -\sum_\gamma \varphi_{\alpha\gamma}\varphi_{\gamma\beta}{}^*, \tag{50}$$

which can be considered as a Hermitian matrix. Since the sum in Eq. (46) involves all possible even chains built up of all possible intermediate states, allowing repetition, we have evidently (in matrix form)

$$R(z) = z^2\chi - z^4\chi^2 + z^6\chi^3 - \cdots = z^2\chi(1+z^2\chi)^{-1}. \tag{51}$$

Continuing to the saddle point $z=z_0=1$, we get

$$\rho = \chi(1+\chi)^{-1}, \tag{52a}$$

a closed expression for ρ in terms of φ. This and similar results of this section are the same as those obtained by Blatt[6] using his algebraic method. Once ρ has been obtained in closed form, κ follows, since it is obvious

from their definitions as sums of odd and even chains, respectively, that κ is related to ρ by

$$\kappa=(\rho-1)\varphi=\varphi(\rho^*-1) \tag{53}$$

$$=-(1+\chi)^{-1}\varphi=-\varphi(1+\chi^*)^{-1}. \tag{52b}$$

In order to put $F(z)$ in closed form, note that according to Eq. (44) one can write

$$z d[\ln F(z)]/dz=\sum_c z^{n_c} n_c R_c$$
$$=\sum_\alpha \sum_i z^{n_i} R_{\alpha\alpha}{}^i=\mathrm{Tr}R(z). \tag{54}$$

The transition from the first to the second line involves the same argument that led to Eq. (42). This differential equation for $F(z)$ is easily integrated. Noting that $F(0)=1$, one obtains

$$F(z)=\exp[\tfrac{1}{2}\,\mathrm{Tr}\ln(1+z^2\chi)]. \tag{55}$$

The assumption made earlier that $R_{\beta\alpha}(z)$ varies slowly compared to $F(z)$ is reasonable since $F(z)$ involves a trace. Moreover, one could in principle calculate corrections to the saddle point approximation, since one possesses exact expressions for the expectation values, such as Eq. (49).

There is a radically different approach which consists in picking for a ground-state trial function

$$|\Phi\rangle=\sum_N |\Phi_N\rangle/N!, \tag{56}$$

i.e., abandoning conservation of the number of particles altogether. Of course, this is the way the BCS and HB theories are usually presented. In that case, one sees by setting $z=1$ in Eqs. (44) and (45) that the expectation value of $c_\alpha{}^\dagger c_\beta$ is rigorously equal to $\rho_{\beta\alpha}$. The trouble has been transferred to the wave function itself. One must introduce a supplementary condition specifying that, at least, the expectation value of the number of particles is correct, i.e., Eq. (17). One must also make sure that the spread in particle numbers in $|\Phi\rangle$, for instance the standard deviation, is small enough to make the conclusions of the calculation applicable to a system with a definite number of particles. It is the virtue of the approach through Bayman's or Blatt's trial functions that it enables one to deal most of the way with a system having a definite number of particles, and that the appearance of nonconservation of numbers is only the result of mathematical approximations. In the usual approach, the introduction of anomalous contractions such as (14b) and (14c) may leave one a little uneasy, because such matrix elements would be zero if the number of particles were conserved. In the present approach, on the other hand, $\kappa_{\beta\alpha}$ arises very naturally and on an equal footing with $\rho_{\beta\alpha}$, the first representing the contribution of odd chains, the second that of even chains.[9]

We are now in a position to show that Eqs. (16) and (17) are the only supplementary conditions (see end of Sec. 3). Start with ρ and κ satisfying these equations, but assume that ρ does not have any eigenvalue equal to unity. Such an eigenvalue corresponds to a fully-occupied single-particle state. If there are any such states, for instance if there is a filled Fermi sphere underlying trial function (6), they should be treated separately. Then Eqs. (53) can be solved to give two expressions for φ in terms of ρ and κ,

$$\varphi=(\rho-1)^{-1}\kappa=\kappa(\rho^*-1)^{-1}. \tag{57}$$

That these two expressions are the same follows from condition (16b). We proceed to use this φ to construct a ρ and a κ according to Eqs. (52) and show that they are the same as the original ones. First, calculate χ,

$$\chi=-\varphi\varphi^*=-(\rho-1)^{-1}\kappa\kappa^*(\rho-1)^{-1}. \tag{58}$$

Using supplementary condition (16a), one can write this as

$$\chi=-\rho(\rho-1)^{-1}. \tag{59}$$

Substituting this in Eq. (52a) gives an identity; substituting in (52b) gives back (53), which was our starting point, q.e.d.

Before ending this section, we consider briefly the simplifications arising in the BCS case. Then, a representation is picked *a priori*, in terms of which the trial function has the simple form (4), i.e. the only nonvanishing components of the pair wave function are of the form

$$\varphi_{k\bar{k}}=-\varphi_{\bar{k}k}=\varphi_k. \tag{60}$$

In the following, φ_k is assumed real. In Sec. 3, the only chains giving nonvanishing contributions would have successive states alternating between the two members of a $(k\bar{k})$ pair. For instance, ρ defined by Eq. (32a) would be diagonal and given by

$$\rho_k=\rho_{kk}=-\varphi_{k\bar{k}}\varphi_{\bar{k}k}-\varphi_{k\bar{k}}\varphi_{\bar{k}k}\varphi_{k\bar{k}}\varphi_{\bar{k}k}-\cdots$$
$$=\varphi_k{}^2-\varphi_k{}^4+\varphi_k{}^6-\cdots. \tag{61}$$

The possibility of divergence, unless the function $R(z)$ is introduced, is evident. Equation (50) defines a diagonal matrix χ, with

$$\chi_k=\chi_{kk}=\chi_{\bar{k}\bar{k}}=\varphi_k{}^2, \tag{62}$$

and Eqs. (52) give

$$\rho_k=\rho_{kk}=\rho_{\bar{k}\bar{k}}=\varphi_k{}^2(1+\varphi_k{}^2)^{-1}, \tag{63a}$$

$$\kappa_{\bar{k}k}=-\kappa_{k\bar{k}}=\varphi_k(1+\varphi_k{}^2)^{-1}. \tag{63b}$$

Other matrix elements of κ vanish. One can define

$$u_k=(1+\varphi_k{}^2)^{-1/2},\ v_k=\varphi_k(1+\varphi_k{}^2)^{-1/2},\ (u_k{}^2+v_k{}^2=1), \tag{64}$$

in terms of which one can write

$$\rho_k=v_k{}^2,\quad \kappa_{\bar{k}k}=u_k v_k, \tag{65}$$

and one is led back to the standard BCS formalism. The condition fixing the normalization of φ is

$$\sum_k \varphi_k{}^2(1+\varphi_k{}^2)^{-1}=\sum_k v_k{}^2=N. \tag{66}$$

[9] We thank H. J. Lipkin for this last remark.

5. HARTREE-BOGOLYUBOV EQUATIONS

In the preceding, the question of finding the ground state of a system of $2N$ interacting Fermions has been set up as a variational problem. With either Blatt's or HB's trial function, it was found that the expectation value of the energy is given by Eq. (15). At this point, it would be possible to start substituting trial values of ρ and κ, provided they satisfy the supplementary conditions, and see which gives the lowest H_0. However, suppose one wants to continue with the formal analytical solution of the problem as long as it is practicable. Then one introduces a Lagrange multiplier λ to take care of condition (17), i.e., one tries to minimize

$$H_0'=H_0-\lambda\,\mathrm{Tr}\rho. \tag{67}$$

Two approaches are possible. The first approach takes φ as the independent variable with respect to which the minimization is done. In the second approach, ρ and κ are the variables, but there are additional supplementary conditions (16). The first approach is the one adopted by Blatt.[3] He derives for φ the following equation [his Eq. (5-14)]

$$\sum_\gamma(T_{\alpha\gamma}-\lambda\delta_{\alpha\gamma}+\Gamma_{\alpha\gamma})\varphi_{\beta\gamma}-\tfrac{1}{2}\Delta_{\alpha\beta}+\tfrac{1}{2}\sum_{\gamma\delta}\Delta_{\gamma\delta}{}^*\varphi_{\alpha\gamma}\varphi_{\beta\delta}=0. \tag{68}$$

This equation is highly nonlinear. Matrices Γ and Δ are the same ones which we define later in Eqs. (89) and are themselves functions of φ. At first sight, there does not seem to be any simple systematic method for solving such an equation; none is provided by Blatt. The second approach, on the other hand, is the one that leads to the HB equations, which are a generalization of the Hartree-Fock equations and whose solution, without being easy, is more familiar: We proceed with it.

Bogolyubov[2] has shown that it is convenient to introduce matrices in which each index takes twice as many values as there are states in the original formulation. In particular, define

$$\mathfrak{R}=\begin{pmatrix}\rho & -\kappa\\ \kappa^* & 1-\rho^*\end{pmatrix}, \tag{69}$$

which can also be written

$$\begin{aligned}&\mathfrak{R}_{\alpha\beta}{}^{11}=\rho_{\alpha\beta}, \quad \mathfrak{R}_{\alpha\beta}{}^{12}=-\kappa_{\alpha\beta},\\ &\mathfrak{R}_{\alpha\beta}{}^{21}=\kappa_{\alpha\beta}{}^*, \quad \mathfrak{R}_{\alpha\beta}{}^{22}=\delta_{\alpha\beta}-\rho_{\beta\alpha}.\end{aligned} \tag{70}$$

This matrix is seen to have the properties

$$\mathfrak{R}^\dagger=\mathfrak{R}, \tag{71}$$

$$f\mathfrak{R}f=1-\mathfrak{R}^*, \tag{72}$$

f being defined by

$$f_{\alpha\beta}{}^{11}=f_{\alpha\beta}{}^{22}=0, \quad f_{\alpha\beta}{}^{12}=f_{\alpha\beta}{}^{21}=\delta_{\alpha\beta}. \tag{73}$$

Expressed in terms of $\mathfrak{R}$, relations (16) become

$$\mathfrak{R}^2=\mathfrak{R}, \tag{74}$$

which says that all eigenvalues of $\mathfrak{R}$ are 0 or 1; then (72) shows that the number of each kind is the same. In a similar notation, define a matrix $\mathfrak{V}$ by

$$\begin{aligned}&\mathfrak{V}_{\alpha\beta\gamma\delta}{}^{1111}=V_{\alpha\beta\gamma\delta}, \quad \mathfrak{V}_{\alpha\beta\gamma\delta}{}^{2222}=V_{\gamma\delta\alpha\beta},\\ &\mathfrak{V}_{\alpha\beta\gamma\delta}{}^{1212}=-V_{\alpha\delta\gamma\beta}, \quad \mathfrak{V}_{\alpha\beta\gamma\delta}{}^{2121}=-V_{\gamma\beta\alpha\delta},\\ &\mathfrak{V}_{\alpha\beta\gamma\delta}{}^{1221}=-V_{\alpha\gamma\beta\delta}, \quad \mathfrak{V}_{\alpha\beta\gamma\delta}{}^{2112}=-V_{\beta\delta\alpha\gamma}.\end{aligned} \tag{75}$$

Submatrices corresponding to the ten other possible combinations of superscripts are all taken to vanish. Let the complete matrix be called $\mathfrak{V}_{abcd}$, each Roman subscript assuming twice as many values as a Greek subscript. It has properties of Hermiticity and antisymmetry similar to (10) and (11). Define further

$$\mathcal{T}=\mathcal{T}^\dagger=\begin{pmatrix}T-\lambda+\frac{1}{2}U & 0\\ 0 & -T^*+\lambda-\frac{1}{2}U^*\end{pmatrix}, \tag{76}$$

with

$$U_{\alpha\gamma}=U_{\gamma\alpha}{}^*=\sum_\beta V_{\alpha\beta\gamma\beta}. \tag{77}$$

Then, it will be found that H_0' can be written as

$$H_0'=\tfrac{1}{2}\sum_{ac}\mathcal{T}_{ac}\mathfrak{R}_{ca}+\tfrac{1}{8}\sum_{abcd}\mathfrak{V}_{abcd}\mathfrak{R}_{ca}\mathfrak{R}_{db}+\tfrac{1}{2}\sum_\alpha(T_{\alpha\alpha}-\lambda)+\tfrac{1}{8}\sum_{\alpha\beta}V_{\alpha\beta\alpha\beta}. \tag{78}$$

The new notation, besides simplifying the argument that follows, makes the HB problem formally similar to the Hartree-Fock one.

The problem is to minimize (78), the variable $\mathfrak{R}$ being restricted by supplementary conditions (71), (72), and (74). Condition (74) is taken care of with a Lagrange multiplier, i.e., one adds to H_0'

$$-\tfrac{1}{2}\sum_{ac}\mu_{ac}(\mathfrak{R}_{ca}-\sum_b\mathfrak{R}_{cb}\mathfrak{R}_{ba}). \tag{79}$$

When $\mathfrak{R}$ is varied by a small amount $\delta\mathfrak{R}$, the first-order change in the variational expression is

$$\tfrac{1}{2}\sum_{ac}(\mathfrak{W}_{ac}-\mu_{ac}+\sum_b\mu_{ab}\mathfrak{R}_{bc}+\sum_b\mathfrak{R}_{ab}\mu_{bc})\delta\mathfrak{R}_{ca}, \tag{80}$$

with the definition

$$\mathfrak{W}_{ac}=\mathcal{T}_{ac}+\tfrac{1}{2}\sum_{bd}\mathfrak{V}_{abcd}\mathfrak{R}_{db}. \tag{81}$$

Since $\mathfrak{R}$ is still restricted by Eqs. (71) and (72), expression (80) must vanish for any $\delta\mathfrak{R}$ satisfying

$$\delta\mathfrak{R}^\dagger=\delta\mathfrak{R}, \tag{82}$$

$$f\delta\mathfrak{R}f=-\delta\mathfrak{R}^*. \tag{83}$$

Equation (82) says that $\delta\mathfrak{R}$ must be Hermitian. But Hermitian matrices form a complete set, in the sense that an arbitrary matrix can be written as a linear combination of Hermitian matrices. Hence, if (80) vanishes for any Hermitian $\delta\mathfrak{R}$, it will also vanish for arbitrary $\delta\mathfrak{R}$. In a similar fashion, matrices satisfying (83) also form a complete set. Therefore, one can conclude that the parenthesis itself must vanish in (80) and that no Lagrange multipliers are necessary for supplementary conditions (71) and (72). The equation determining $\mathfrak{R}$ is then

$$\mathfrak{W}-\mu+\mu\mathfrak{R}+\mathfrak{R}\mu=0, \tag{84}$$

from which μ must be eliminated through use of condition (74). This is done by writing the vanishing of the commutator,

$$[\mathcal{R}, \mathcal{W}-\mu+\mu\mathcal{R}+\mathcal{R}\mu]=0 \tag{85}$$

which, upon use of (74), becomes simply

$$[\mathcal{R},\mathcal{W}]=0. \tag{86}$$

In conclusion, $\mathcal{R}$ must be constructed in such a way as to commute with $\mathcal{W}$. To achieve this, one finds a representation diagonalizing $\mathcal{W}$ and one picks $\mathcal{R}$ diagonal in the same representation. This would be easy if $\mathcal{W}$ was not itself a function of $\mathcal{R}$. Since it is, the $\mathcal{R}$ that comes out of the diagonalization process must be the same as the $\mathcal{R}$ used to calculate $\mathcal{W}$ in the first place. This self-consistency requirement makes all the difficulty of the problem. In general, the only way to arrive at an exact solution, short of minimizing H_0 with all possible $\mathcal{R}$'s, is to repeat indefinitely the diagonalization procedure, $\mathcal{W}$ being recalculated at each step with $\mathcal{R}$ found in previous steps, and to hope that the sequence converges. As for the supplementary conditions on $\mathcal{R}$, they can all be satisfied. First, $\mathcal{W}$ and $\mathcal{R}$ are Hermitian together. Second, if $\mathcal{R}$ satisfies (72), then $\mathcal{W}$ satisfies

$$f\mathcal{W}f=-\mathcal{W}^*. \tag{87}$$

It follows that, if $\mathcal{W}$ has eigenvector $\mathcal{A}_i$ for eigenvalue E_i, it also has eigenvector $f\mathcal{A}_i^*$ for eigenvalue $-E_i$. The new $\mathcal{R}$ is given eigenvalue 1 for one of these eigenvectors and 0 for the other, which makes it obey (72) again.

The eigenvalue equations for $\mathcal{W}$ are the HB equations. We want to write them now in the original notation with the Greek subscripts. One finds

$$\mathcal{W}=\begin{pmatrix} T-\lambda+\Gamma & \Delta \\ -\Delta^* & -T^*+\lambda-\Gamma^* \end{pmatrix}, \tag{88}$$

where the Hartree potential Γ is given by

$$\Gamma_{\alpha\gamma}=\sum_{\beta\delta} V_{\alpha\beta\gamma\delta}\,\rho_{\delta\beta}=\Gamma_{\gamma\alpha}^* \tag{89a}$$

and the pairing potential Δ is given by

$$\Delta_{\alpha\beta}=\tfrac{1}{2}\sum_{\gamma\delta} V_{\alpha\beta\gamma\delta}\,\kappa_{\delta\gamma}=-\Delta_{\beta\alpha}. \tag{89b}$$

For the eigenvector of $\mathcal{W}$, use the notation

$$(\mathcal{A}_i)_\alpha{}^1=A_{\alpha i}, \quad (\mathcal{A}_i)_\alpha{}^2=B_{\alpha i}. \tag{90}$$

The HB equations are

$$\begin{aligned} E_iA_{\alpha i}&=\sum_\gamma(T_{\alpha\gamma}-\lambda\delta_{\alpha\gamma}+\Gamma_{\alpha\gamma})A_{\gamma i}+\sum_\beta \Delta_{\alpha\beta}B_{\beta i}, \\ -E_iB_{\alpha i}&=\sum_\gamma(T_{\alpha\gamma}^*-\lambda\delta_{\alpha\gamma}+\Gamma_{\alpha\gamma}^*)B_{\gamma i} \\ &\qquad +\sum_\beta \Delta_{\alpha\beta}^*A_{\beta i}. \end{aligned} \tag{91}$$

As mentioned earlier, solutions occur in pairs, $\mathcal{A}_i$ and $f\mathcal{A}_i^*$. We label with the index i only one of the members of the pair, so that i takes on as many values as there are states labeled by Greek indices. The orthonormality condition for the eigenvectors becomes

$$\begin{aligned} \textstyle\sum_\alpha(A_{\alpha i}^*A_{\alpha j}+B_{\alpha i}^*B_{\alpha j})&=\delta_{ij}, \\ \textstyle\sum_\alpha(A_{\alpha i}B_{\alpha j}+B_{\alpha i}A_{\alpha j})&=0, \end{aligned} \tag{92}$$

and the completeness relations are

$$\begin{aligned} \textstyle\sum_i(A_{\alpha i}A_{\beta i}^*+B_{\alpha i}^*B_{\beta i})&=\delta_{\alpha\beta}, \\ \textstyle\sum_i(A_{\alpha i}B_{\beta i}^*+B_{\alpha i}^*A_{\beta i})&=0. \end{aligned} \tag{93}$$

If one constructs $\mathcal{R}$ in the manner outlined earlier, giving it eigenvalue 0 for each $\mathcal{A}_i$ eigenvector of $\mathcal{W}$ and eigenvalue 1 for each $f\mathcal{A}_i^*$ eigenvector, one gets

$$\rho_{\alpha\gamma}=\textstyle\sum_i B_{\alpha i}^*B_{\gamma i}, \tag{94a}$$

$$\kappa_{\alpha\beta}=\textstyle\sum_i A_{\alpha i}B_{\beta i}^*. \tag{94b}$$

To see that κ is antisymmetric and to check conditions (16), it is necessary to use the completeness relations. Since Γ and Δ depend on ρ and κ, the HB equations are nonlinear.

In the development we have given, the eigenvalue E_i does not seem to have any particular significance. This is because we have concentrated exclusively on the ground state, both with Blatt's point of view and the HB point of view. In other formulations of HB theory,[2] in which more attention is paid to excited states, E_i is seen to be the energy of an elementary excitation or quasiparticle. For this reason, a solution cannot be stable unless all E_i's are positive. One can also show that $A_{\alpha i}$ and $B_{\alpha i}$ that solve the HB equations are precisely the coefficients to be used in Eq. (13) defining the quasiparticle operators.

The HB equations are useful even in a variational approach in which no attempt is made to solve the HB problem exactly. The question then is to pick trial matrices ρ and κ satisfying supplementary conditions (16). This can be achieved by picking arbitrary potentials Γ and Δ and solving the HB equations as linear equations. Out of the solutions, one builds ρ and κ by Eqs. (94) and they satisfy the supplementary conditions automatically.

ACKNOWLEDGMENTS

Most of this work was performed while the author was on a year's visit at the Centre d'Etudes Nucléaires de Saclay. He wishes to thank Professor Claude Bloch for the hospitality and stimulation he found in the Service de Physique Mathématique, and the National Science Foundation for financial support.