PHYSICAL REVIEW A VOLUME 28, NUMBER 4 OCTOBER 1983

New variational approach to the many-body problem and application to the high-density electron gas

Yasutami Takada*
Department of Physics, Purdue University, West Lafayette, Indiana 47907

(Received 29 July 1982; revised manuscript received 9 June 1983)

A new variational approach is proposed for the study of the many-body problem with the purpose of combining the advantages of perturbation-theoretic and variational approaches, while avoiding their disadvantages. A trial wave function is constructed so that a general expansion for the expectation value of an arbitrary physical quantity can be derived. Each term in the expansion can be calculated with the use of the techniques of Feynman diagrams, but in contrast with usual perturbation-theoretic approaches, every term in the expansion always converges. The procedure for applying the method to real systems is illustrated by a calculation of the ground-state energy of the electron gas at high densities.

I. INTRODUCTION

In most cases of interest, problems in an interacting many-particle system cannot be solved exactly. Thus two approaches have frequently been adopted to study the properties of a many-body system. One is to employ perturbation theory and the other is to make a variational calculation. Each method has its own advantages as well as disadvantages. In perturbation treatments, the most advantageous point is that the higher-order corrections can be calculated quite systematically by the application of quantum-field-theoretic methods,[1] in particluar, the techniques of Feynman diagrams.[2] However, one of the troubles in a perturbation-theoretic calculation is that the inclusion of higher-order terms does not always lead to an improvement in the results. We often encounter the situation where inclusion of the next-order terms leads to a poorer result. To make matters worse, perturbation theory itself cannot decide whether or not the next-order terms should be included. Another trouble is the frequent appearance of divergences in some terms of the perturbation series. Usually, this kind of divergent series is treated by an infinite summation, but the validity of the obtained result is very difficult to justify from the method itself. A different method of summation of such a formally divergent series generally produces a different answer.

A variational approach is superior to a perturbation-theoretic approach in several respects. First, the variational principle can decide which approximation, or which type of trial wave function should be taken. A trial wave function providing a lower energy is always considered to be better. Second, a complication such as the increase in the number of variational parameters always gives a better answer, because this makes the energy lower than before. Third, the variational principle also leads to a definite statement about the upper limit of the true energy. Such a limit cannot be obtained in perturbation theory for which the calculated energy can be either higher or lower than the true energy. Fourth, even if we take a trial function having an error of the order of δ, the error of the calculated energy is small, i.e., of order δ^2. In contrast, in a perturbation-theoretic calculation, particularly when the coupling-constant integration algorithm due to Pauli, Hellmann, and Feynman is employed,[3] the error of the energy is of order δ.

In spite of such advantages, the variational approach is less used in the many-body problem than the perturbation-theoretic approach. There are reasons for this. In a variational method, one has to choose some special form of trial function appropriate to the problem. Therefore, the first trouble is that a good choice requires deep physical insight and this is hard to prove. Another trouble is connected with the first one. When the form of trial function differs from problem to problem, we cannot give a general prescription for solving problems in a systematic way. Lack of such a general prescription makes calculations so tedious and difficult, especially with the use of a complicated trial function which many-body systems usually need, that in most cases, one has to introduce some further approximations in order to perform the calculations. However, the great advantage of the variational approach, namely, the existence of the variational upper bounds, is generally lost if the energy expectation value is not evaluated exactly. In this respect, the variational approach suggested by Jastrow[4] seems to provide a way to overcome such troubles. His trial function is constructed from two-particle correlation functions. Many attempts have already been made along this line, primarily in the study of the nuclear matter problem.[5] However, in the Jastrow-type formulation, it seems to be quite difficult to improve trial functions systematically by the inclusion of three-, four-, and higher-order-particle correlations. Besides, even if we obtain some results with extensive use of large computers, we cannot usually understand what the essential physics among many processes is (which may be described in terms of diagrams in perturbation theory).

Since the perturbation-theoretic and variational approaches are complementary to each other, we can expect to obtain a very powerful method for combining these two approaches. There are two important attempts in this direction. One is the correlated—basis-function (CBF) method[6] and the other is the coupled-cluster formalism [or the exp(S) method].[7] The CBF method is essentially a perturbation-theoretic approach: A perturbation expansion is developed with a nonorthogonal set of correlated basis functions (which can be given by a Jastrow-type variational calculation) in order to improve systematically on the Jastrow description of the system. The variational principle cannot be applied to the coupled-cluster formalism, either. The wave function of the ground state $|\Phi_0\rangle$ is assumed to be $\exp(S)|0\rangle$ in the formalism, where S is an operator to be determined by some integral equation and $|0\rangle$ is the ground state of the noninteracting system. Then the ground-state energy is calculated by $\langle 0|H|\Phi_0\rangle/\langle 0|\Phi_0\rangle$, not by $\langle\Phi_0|H|\Phi_0\rangle/\langle\Phi_0|\Phi_0\rangle$, where H is the Hamiltonian of the system.

Taking these points into account, we propose another method in the present paper which combines the advantages of perturbation-theoretic and variational approaches. The present method is based on the variational principle, in contrast with both the CBF and exp(S) methods. But this method is so formulated as to avoid the disadvantages of other variational methods. In particular, this method can include higher-order effects systematically. It can also incorporate the techniques of Feynman diagrams for calculating physical quantities like the total energy. We also mention that divergences never appear in the calculation.

As a first example, we treat a homogeneous electron gas neutralized by a uniform positive background, and consider the system at high densities. Several many-body techniques have been applied to this system[8-10] and have already given a rigorous expression for the correlation energy per electron ϵ_c. The result, first derived by Gell-Mann and Brueckner, is expressed as[11]

$$\epsilon_c = 0.0622\ln r_s - 0.0933 + O(r_s \ln r_s)\ , \tag{1}$$

where the energy is in rydbergs and r_s is the parameter related to the average interparticle distance in units of the Bohr radius. The same expression is also obtained in the exp(S) method.[12] Variational calculations with simple Jastrow-type trial function have also been done by many workers,[13-15] but it seems certain that this approach cannot reproduce Eq. (1). Gaskell[13] gave an expression for ϵ_c in the high-density limit ($r_s \ll 1$) as[16]

$$\epsilon_c = 0.057\ln r_s - 0.0836 + O(r_s \ln r_s)\ , \tag{2}$$

and Zabolitzky[15] obtained the same coefficient of the $\ln r_s$ term. When we compare Eq. (2) with Eq. (1) at $r_s = 0.1$, the error is 9.2%. The correct $\ln r_s$ term is obtained only after an improvement is done with the use of the CBF method.[17]

When we employ our new method, we can obtain ϵ_c easily as

$$\epsilon_c = 0.0622\ln r_s - 0.0910 + O(r_s \ln r_s)\ . \tag{3}$$

This is a much better result than Eq. (2). By improving the trial function systematically, we can approach Eq. (1) more closely.

We describe our method in Sec. II by presenting a trial wave function and by formulating a prescription to calculate the expectation value of an arbitrary physical operator. Section III is devoted to the calculation of the correlation energy of the electron gas in the high-density limit. This illustrates how the formalism of Sec. II is used in an actual problem. After deriving Eq. (3) by a rigorous treatment of variational calculations, we propose a simplified variational procedure from a physical consideration. The procedure treats only a finite number of terms in the expansion of the energy expectation value, but it gives a very good result for ϵ_c. Systematic inclusion of the higher order corrections is done with this procedure. In the end of Sec. III, we show how to take account of collective excitations explicitly in our trial function and also present how to include the nonlocal effect which is brought about by the retardation of the effective interaction. In Sec. IV, we conclude with an outlook for this new method.

II. FORMALISM

A. Generalized linked-cluster theorem

Consider a system of N particles which are contained in a volume Ω, interacting with one another through a two-body potential, and moving in an external potential field. The Hamiltonian for such a system is written as

$$H = H_0 + V\ , \tag{4}$$

where H_0 includes the kinetic energy of the particles and any external field, and V represents the interaction energy of the particles. Let us consider the following wave function[18]:

$$|\Phi_0\rangle = \sum_{n=0}^{\infty} \frac{1}{n!}\left[\sum_{l=1}^{\infty} \mu_l U_l(0,-\infty)\right]^n |0\rangle \tag{5a}$$

with $U_l(0,-\infty)$ defined by

$$U_l(0,-\infty) = \frac{(-i)^l}{l!}\int_{-\infty}^{0} e^{\epsilon t_1}dt_1 \cdots \int_{-\infty}^{0} e^{\epsilon t_l}dt_l\, T[\tilde{V}(t_1)\cdots\tilde{V}(t_l)]_L\ , \tag{6a}$$

where μ_l's are arbitrary parameters, ϵ is an infinitesimally small positive quantity, and $\tilde{V}(t)$ is defined as

$$\tilde{V}(t) = \exp(i\tilde{H}_0 t)\tilde{V}\exp(-i\tilde{H}_0 t)\ , \tag{7}$$

with some suitably chosen noninteracting Hamiltonian $\tilde{H}_0$ and effective interaction $\tilde{V}$. The state $|0\rangle$ is the normalized ground state of $\tilde{H}_0$; the symbol T and the subscript L represent the T product and the instruction to collect only the terms described by linked graphs in their ordinary sense.[1,19]

The expectation value of some physical quantity A with respect to the wave function (5a) is given by

$$\langle A\rangle=\langle\Phi_0|A\,|\Phi_0\rangle/\langle\Phi_0|\Phi_0\rangle\ . \tag{8}$$

We first consider the numerator which is given by

$$\langle\Phi_0|A\,|\Phi_0\rangle=\sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}\frac{1}{n!n'!}\left\langle 0\left|\left[\sum_{l=1}^{\infty}\mu_l U_l^\dagger(0,\infty)\right]^n A\left[\sum_{l'=1}^{\infty}\mu_{l'}U_{l'}(0,-\infty)\right]^{n'}\right|0\right\rangle . \tag{9}$$

In deriving Eq. (9), $|\Phi_0\rangle$ was written in a way other than Eq. (5a):

$$|\Phi_0\rangle=\sum_{n=0}^{\infty}\frac{1}{n!}\left[\sum_{l=1}^{\infty}\mu_l U_l(0,\infty)\right]^n|0\rangle\ , \tag{5b}$$

with $U(0,\infty)$ defined by

$$U_l(0,\infty)=\frac{i^l}{l!}\int_0^\infty e^{-\epsilon u_1}du_1\cdots\int_0^\infty e^{-\epsilon u_l}du_l T[\widetilde{V}(u_1)\cdots\widetilde{V}(u_l)]_L\ , \tag{6b}$$

and this equation was used for $\langle\Phi_0|$. In the further reduction of Eq. (9), we classify the contributions of the term

$$\left\langle 0\left|\left[\sum_{l=1}^{\infty}\mu_l U_l^\dagger(0,\infty)\right]^n A\left[\sum_{l'=1}^{\infty}\mu_{l'}U_{l'}(0,-\infty)\right]^{n'}\right|0\right\rangle \tag{10}$$

according to the number of $U_l^\dagger$ and $U_{l'}$ which are connected with A. The simplest contribution is the term in which A does not connect with either $U_l^\dagger$ or $U_{l'}$. Namely, we obtain the term

$$\langle 0|A\,|0\rangle\left\langle 0\left|\left[\sum_l\mu_l U_l^\dagger\right]^n\left[\sum_{l'}\mu_{l'}U_{l'}\right]^{n'}\right|0\right\rangle . \tag{10$'$}$$

The next one is the term in which one $U_l^\dagger$ is connected with A. Since each term in $(\sum_l\mu_l U_l^\dagger)^n$ is unlinked to the others, this contribution can be written as

$$n\sum_{l=1}^{\infty}\mu_l\left\langle 0\left|U_l^\dagger(0,\infty)A\right|0\right\rangle_C\left\langle 0\left|\left[\sum_l\mu_l U_l^\dagger\right]^{n-1}\left[\sum_{l'}\mu_{l'}U_{l'}\right]^{n'}\right|0\right\rangle , \tag{10$''$}$$

where the subscript C means that only connected diagrams should be considered in the evaluation of the expectation value. Similarly, the contribution in which one $U_{l'}$ is connected with A is given by

$$n'\sum_{l'=1}^{\infty}\mu_{l'}\langle 0|AU_{l'}(0,-\infty)|0\rangle_C\left\langle 0\left|\left[\sum_l\mu_l U_l^\dagger\right]^{n}\left[\sum_{l'}\mu_{l'}U_{l'}\right]^{n'-1}\right|0\right\rangle . \tag{10$'''$}$$

When there are $mU_l^\dagger$'s and $m'U_{l'}$'s connected with A, the contribution obtained is

$$\begin{aligned}\frac{n!}{m!(n-m)!}\frac{n'!}{m'!(n'-m')!}&\sum_{l_1=1}^{\infty}\cdots\sum_{l_m=1}^{\infty}\sum_{l'_1=1}^{\infty}\cdots\sum_{l'_{m'}=1}^{\infty}\mu_{l_1}\cdots\mu_{l_m}\mu_{l'_1}\cdots\mu_{l'_{m'}}\\&\times\langle 0|U_{l_1}^\dagger(0,\infty)\cdots U_{l_m}^\dagger(0,\infty)AU_{l'_1}(0,-\infty)\cdots U_{l'_{m'}}(0,-\infty)|0\rangle_C\\&\times\left\langle 0\left|\left[\sum_l\mu_l U_l^\dagger\right]^{n-m}\left[\sum_{l'}\mu_{l'}U_{l'}\right]^{n'-m'}\right|0\right\rangle .\end{aligned} \tag{10$''''$}$$

In the evaluation of the connected terms, we should consider only such terms in which each $U^\dagger$ (U) is linked to other $U^\dagger$'s (U's) after it is connected with A and/or U's ($U^\dagger$'s). This situation is shown schematically in Fig. 1.

Substituting terms (10$'$)—(10$''''$) into Eq. (9) and arranging each term, we get

$$\begin{aligned}\langle\Phi_0|A\,|\Phi_0\rangle=&\sum_{m=0}^{\infty}\sum_{m'=0}^{\infty}\sum_{l_1=1}^{\infty}\cdots\sum_{l_m=1}^{\infty}\sum_{l'_1=1}^{\infty}\cdots\sum_{l'_{m'}=1}^{\infty}\mu_{l_1}\cdots\mu_{l_m}\mu_{l'_1}\cdots\mu_{l'_{m'}}\frac{1}{m!m'!}\langle 0|U_{l_1}^\dagger\cdots U_{l_m}^\dagger AU_{l'_1}\cdots U_{l'_{m'}}|0\rangle_C\\&\times\left[\sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}\frac{1}{n!n'!}\left\langle 0\left|\left[\sum_l\mu_l U_l^\dagger\right]^n\left[\sum_{l'}\mu_{l'}U_{l'}\right]^{n'}\right|0\right\rangle\right] .\end{aligned} \tag{11}$$

Since the second factor in Eq. (11) is nothing but $\langle\Phi_0|\Phi_0\rangle$, we obtain the result for the expectation value as the first factor in Eq. (11). This factor can be written in another way to give $\langle A\rangle$ as

$$\langle A \rangle = \left\langle 0 \left| \exp\left[\sum_{l=1}^{\infty} \mu_l U_l^{\dagger}(0,\infty) \right] A \exp\left[\sum_{l'=1}^{\infty} \mu_{l'} U_{l'}(0,-\infty) \right] \right| 0 \right\rangle_C . \tag{12}$$

Introducing the correlation function $C(A : m_1, m_2, \ldots ; m_1', m_2', \ldots)$, defined by

$$C(A : m_1, m_2, \ldots ; m_1', m_2', \ldots) = \frac{\langle 0 | U_1^{\dagger m_1} U_2^{\dagger m_2} \cdots A U_1^{m_1'} U_2^{m_2'} \cdots | 0 \rangle_C}{m_1! m_2! \cdots m_1'! m_2'! \cdots} , \tag{13}$$

we can rewrite Eq. (12) further as

$$\langle A \rangle = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \cdots \sum_{m_1'=0}^{\infty} \sum_{m_2'=0}^{\infty} \cdots C(A : m_1, m_2, \ldots ; m_1', m_2', \ldots) \mu_1^{m_1+m_1'} \mu_2^{m_2+m_2'} \cdots . \tag{14}$$

Although $U_1^{\dagger}, U_2^{\dagger}, \ldots$ and $U_1, U_2, \ldots$ are written in this order in Eq. (13), we should include the contribution of any order of $U_1^{\dagger}, U_2^{\dagger}, \ldots$ and $U_1, U_2, \ldots$ in the calculation of $C(A : m_1, m_2, \ldots ; m_1', m_2', \ldots)$.

Since the result of Eq. (12), or Eq. (14), contains only the contribution from connected diagrams, this result may be classified as a linked-cluster theorem. However, compared to Goldstone's linked-cluster theorem,[19] the present theorem is more general in the following two aspects. First, although Eq. (5) is a type of perturbation expansion, the expansion is not done with respect to the bare interaction V but to some effective interaction $\tilde{V}$. Second, Goldstone's theorem has been derived only in the case of $\mu_l = 1$, but we have succeeded here in deriving a rigorous result with any value of μ_l.

B. Variational procedure

Let us try to solve the system described by H in Eq. (4) variationally. For the trial function of the ground state, we consider a series of functions $\{ | \Phi_0^{(j)} \rangle \}$. The function $| \Phi_0^{(j)} \rangle$ is defined by Eq. (5a) with $\mu_l = 0$ for $l \geq j+1$. The parameters $\mu_1, \ldots, \mu_j$ in $| \Phi_0^{(j)} \rangle$ can be determined by the variational procedure. It should be noted that μ_l in $| \Phi_0^{(j)} \rangle$ thus obtained is generally different from that in $| \Phi_0^{(j-1)} \rangle$.

With the use of an excited state $| l \rangle$ of $\tilde{H}_0$ which satisfies

$$\tilde{H}_0 | l \rangle = \tilde{E}_l | l \rangle , \tag{15}$$

$| \Phi_0^{(j)} \rangle$ with $j = 1$, for example, can be rewritten as

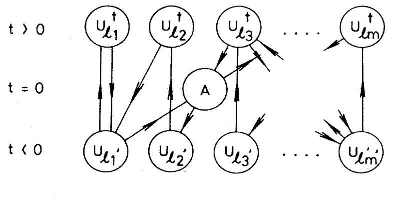

FIG. 1. Schematic representation of the structure of $\langle 0 | U_{l_1}^{\dagger} \cdots U_{l_m}^{\dagger} A U_{l_1'} \cdots U_{l_m'} | 0 \rangle_C$. Operator A is assumed to be a two-body interaction.

$$| \Phi_0^{(1)} \rangle = | 0 \rangle - \mu_1 \sum_{l \neq 0} \frac{| l \rangle \langle l | \tilde{V} | 0 \rangle}{\tilde{E}_l - \tilde{E}_0} + \tfrac{1}{2} \mu_1^2 \sum_{l \neq 0} \sum_{l' \neq 0} \frac{| l, l' \rangle \langle l | \tilde{V} | 0 \rangle \langle l' | \tilde{V} | 0 \rangle}{(\tilde{E}_l - \tilde{E}_0)(\tilde{E}_{l'} - \tilde{E}_0)} - \cdots . \tag{16}$$

When we take μ_1 to be unity and do not consider the terms higher than the first order, $| \Phi_0^{(1)} \rangle$ is reduced to the trial wave function suggested by Macke.[20] Thus, our method may be regarded as a direct extension of Macke's idea. The important consequence of this extension is that we can achieve a progressive and systematic improvement by changing the trial function from $| \Phi_0^{(j)} \rangle$ to $| \Phi_0^{(j+1)} \rangle$.

When we consider the normal state of metals and take the effective interaction $\tilde{V}$ as a two-body potential, $| l \rangle$ in Eq. (16) is the state in which two electrons below the Fermi surface are excited above it and leave two holes. We will denote such a state by (2*e*-2*h*). Then, the *n*th order term in Eq. (16) has *n* uncorrelated (2*e*-2*h*) states. For example, $| l, l' \rangle$ is the state constructed by two uncorrelated (2*e*-2*h*) states l and l'. Thus $| \Phi_0^{(1)} \rangle$ contains only two-particle correlations. In general, $| \Phi_0^{(j)} \rangle$ can include the higher-order-particle correlations up to the $(j+1)$th order. If we assume further in Eq. (16) that $\mu_1 \langle l | \tilde{V} | 0 \rangle / (\tilde{E}_l - \tilde{E}_0)$ depends only on the transferred momentum $\vec{q}$, i.e.,

$$\mu_1 \langle l | \tilde{V} | 0 \rangle / (\tilde{E}_l - \tilde{E}_0) = f_{\vec{q}} , \tag{17}$$

our trial function $| \Phi_0^{(1)} \rangle$ is just the same as that proposed by Jastrow.[4] The correlation factor F in his wave function is related to $f_{\vec{q}}$ through

$$F = \prod_{\substack{i,j \\ i<j}} \exp\left[- \sum_{\vec{q}} f_{\vec{q}} \exp(i \vec{q} \cdot \vec{r}_{ij}) \right] . \tag{18}$$

Thus the Jastrow-type approach is included in our formalism as a special case.

The choice of $\tilde{H}_0$ and $\tilde{V}$ is at our disposal. In principle, the most suitable $\tilde{H}_0$ and $\tilde{V}$ are determined by the variational procedure. But in order to obtain a good answer even with the use of $| \Phi_0^{(1)} \rangle$ and with only a few terms in Eq. (14), rather deep physical insight is necessary. Generally speaking, when we treat normal Fermi-liquid systems, $\tilde{H}_0$ may be taken as the kinetic energy of the quasiparticles and $\tilde{V}$ as the effective interaction between such quasiparticles. When we study a condensed phase like a charge-density-wave or superconducting state, we should include such information in $\tilde{H}_0$ and start from the condensed state $|0\rangle$, e.g., a BCS state in superconductivity.[21] In the construction of appropriate forms for $\tilde{H}_0$ and $\tilde{V}$, we need to include parameters like the effective mass and strength of the effective interaction. These parameters will be variational parameters in addition to μ_l.

C. Calculation of the correlation function

The correlation function $C(A:m_1,m_2,\ldots;m'_1,m'_2,\ldots)$ can be evaluated with the use of quantum-field-theoretic techniques, because the structure of this function is very similar to that appearing in the ordinary perturbation-theoretic approach. Let us first define a causal Green's function $G_l(A:\omega_1,\ldots,\omega_l)$ as

$$G_l(A:\omega_1,\ldots,\omega_l)=\frac{(-i)^l}{l!}\int_{-\infty}^{\infty}\exp(i\omega_1 t_1-\epsilon|t_1|)dt_1\cdots\int_{-\infty}^{\infty}\exp(i\omega_l t_l-\epsilon|t_l|)dt_l \times\langle 0|T[\tilde{V}(t_1)\cdots\tilde{V}(t_l)A]|0\rangle_C\;, \tag{19}$$

where the operator A should be regarded as having a time of zero, namely, $A=A(t=0)$. This Green's function can be calculated with the techniques of Feynman diagrams. Since the rules for application of Feynman diagrams can be found in any standard textbook on the many-body problem,[1] we will omit the details of the method to calculate $G_l(A:\omega_1,\ldots,\omega_l)$.

In Eq. (19), the parameters $\omega_1,\ldots,\omega_l$ are external frequencies and if they are taken to be zero, we obtain the following relation by comparing the definition of Eq. (13) with that of Eq. (19):

$$G_l(A:0,0,\ldots,0)=\sum_{m_1}\sum_{m_2}\cdots\sum_{m'_1}\sum_{m'_2}\cdots C(A:m_1,m_2,\ldots;m'_1,m'_2,\ldots)\;, \tag{20}$$

where the summation is done under the condition

$$m_1+2m_2+\cdots+m'_1+2m'_2+\cdots=l\;. \tag{21}$$

Taking $\mu_l=1$ and $\tilde{V}=V$ and substituting Eq. (20) into Eq. (14), we find easily that the expansion (14) reduces to that in the ordinary perturbation-theoretic approach. However, it is very important to extract $C(A:m_1,m_2,\ldots;m'_1,m'_2,\ldots)$ from G_l in our formalism. For that purpose, we need to calculate G_l as a function of ω_n and then to investigate the analytic behavior of this Green's function in the ω_n plane. For example, $G_2(A:\omega_1,\omega_2)$ is composed of eight parts, $G_2^{(+,+)}(A:\omega_1,\omega_1+\omega_2)$, $G_2^{(+,+)}(A:\omega_1+\omega_2,\omega_2)$, $G_2^{(+)(+)}(A:\omega_1,\omega_2)$, $G_2^{(+)(-)}(A:\omega_1,\omega_2)$, $G_2^{(-)(+)}(A:\omega_1,\omega_2)$, $G_2^{(-,-)}(A:\omega_1,\omega_1+\omega_2)$, $G_2^{(-,-)}(A:\omega_1+\omega_2,\omega_2)$, and $G_2^{(-)(-)}(A:\omega_1,\omega_2)$. The part $G_2^{(+,+)}(A:\omega_1,\omega_1+\omega_2)$ is analytic in both upper ω_1 and $(\omega_1+\omega_2)$ planes, while $G_2^{(+)(+)}(A:\omega_1,\omega_2)$ is analytic in both upper ω_1 and ω_2 planes. The part $G_2^{(+)(-)}(A:\omega_1,\omega_2)$ is analytic in upper ω_1 and lower ω_2 planes. Other parts can be defined in a similar way. From this analytic behavior, we obtain

$$\begin{aligned}
&C(A:0,1,0,\ldots;0,0,0,\ldots)\\
&\qquad=G_2^{(+,+)}(A:0,0+0)+G_2^{(+,+)}(A:0+0,0)\;,\\
&C(A:2,0,0,\ldots;0,0,0,\ldots)=G_2^{(+)(+)}(A:0,0)\;,\\
&C(A:1,0,0,\ldots;1,0,0,\ldots)\\
&\qquad=G_2^{(+)(-)}(A:0,0)+G_2^{(-)(+)}(A:0,0)\;, \qquad (22)\\
&C(A:0,0,0,\ldots;0,1,0,\ldots)\\
&\qquad=G_2^{(-,-)}(A:0,0+0)+G_2^{(-,-)}(A:0+0,0)\;,\\
&C(A:0,0,0,\ldots;2,0,0,\ldots)=G_2^{(-)(-)}(A:0,0)\;.
\end{aligned}$$

It might seem to be confusing to distinguish $G_2^{(+,+)}(A:\omega_1,\omega_1+\omega_2)$, or $G_2^{(+,+)}(A:\omega_1+\omega_2,\omega_2)$, from $G_2^{(+)(+)}(A:\omega_1,\omega_2)$, but in actual calculations, it is easy to classify the contributions of $G_l(A:\omega_1,\ldots,\omega_l)$. This process is probably best illustrated by examples and some of them will be shown in Sec. III.

III. ELECTRON GAS AT HIGH DENSITIES

A. Preliminaries

For the case of an electron gas in a uniform positive background, H_0 and V are given by

$$H_0=\sum_{\vec{k},\sigma}\epsilon_{\vec{k}}\mathscr{A}^{\dagger}_{\vec{k},\sigma}\mathscr{A}_{\vec{k},\sigma}\;, \tag{23}$$

and

$$V=\tfrac{1}{2}\sum_{\vec{q}\neq\vec{0}}\;\sum_{\vec{k}_1,\vec{k}_2}\;\sum_{\sigma_1,\sigma_2}V(\vec{q})\mathscr{A}^{\dagger}_{\vec{k}_1+\vec{q},\sigma_1}\mathscr{A}^{\dagger}_{\vec{k}_2-\vec{q},\sigma_2}\times\mathscr{A}_{\vec{k}_2,\sigma_2}\mathscr{A}_{\vec{k}_1,\sigma_1}\;, \tag{24}$$

in second quantization, where $\epsilon_{\vec{k}}=\vec{k}^2/2m$, $V(\vec{q})=4\pi e^2/\Omega q^2$, $\mathscr{A}_{\vec{k},\sigma}$ is the destruction operator of an electron with momentum $\vec{k}$ and spin σ, and the omission of the term $\vec{q}=\vec{0}$ is a consequence of charge neutrality of the system.

Now, let us choose $\tilde{H}_0$ and $\tilde{V}$. In the normal state, each electron moves with an effective mass m^* instead of a bare mass m. In the high-density limit, however, m^* approaches m. Thus we will take H_0 itself for $\tilde{H}_0$. The normalized ground state $|0\rangle$ is, therefore, a state described by a plane-wave Slater determinant, which is written in second quantization as

$$|0\rangle=\prod_{|\vec{k}|\le k_F}\prod_{\sigma}\mathscr{A}^{\dagger}_{\vec{k},\sigma}|\,\text{vacuum}\rangle\ , \tag{25}$$

where k_F is the Fermi momentum.

Unlike $\tilde{H}_0$, we cannot choose V for $\tilde{V}$. Although there are two types of excitation in an electron gas, i.e., the individual and collective, only the former is included explicitly in our trial function at this stage. For example, the wave function (16) is composed of uncorrelated $(2e\text{-}2h)$ states. (They are individual excitations.) However, we cannot neglect the effect of the collective excitation, or plasmon, in the choice of $\tilde{V}$. When an electron is excited, a charge fluctuation is created which induces a simultaneous collective response of other electrons. Thus we have to introduce a screened potential which produces a $(2e\text{-}2h)$ state less frequently than a bare potential. From such a consideration, we will take $\tilde{V}$ as

$$\tilde{V}=\frac{1}{2}\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}_1,\vec{k}_2}\sum_{\sigma_1,\sigma_2}\tilde{V}(\vec{q})\mathscr{A}^{\dagger}_{\vec{k}_1+\vec{q},\sigma_1}\mathscr{A}^{\dagger}_{\vec{k}_2-\vec{q},\sigma_2}\times\mathscr{A}_{\vec{k}_2,\sigma_2}\mathscr{A}_{\vec{k}_1,\sigma_1}\ , \tag{26}$$

with

$$\tilde{V}(\vec{q})=\frac{4\pi e^2}{\Omega(q^2+\lambda q_{\mathrm{TF}}^2)}\ . \tag{27}$$

λ is a variational parameter which adjusts the strength of the coupling between electrons and plasmons; q_{TF} is the Thomas-Fermi screening constant, given by

$$q_{\mathrm{TF}}^2/k_F^2=\frac{4}{\pi}\alpha r_s\ , \tag{28}$$

with $\alpha=(4/9\pi)^{1/3}=0.521$, and $r_s=me^2/\alpha k_F$.

When the state (25) is used for the evaluation of $\langle 0|H|0\rangle$, we obtain easily

$$\langle 0|H|0\rangle\Big/\left[N\frac{me^4}{2}\right]=\frac{2.21}{r_s^2}-\frac{0.916}{r_s}\ , \tag{29}$$

in which the first term is the contribution of the kinetic energy in rydbergs and the second is that of the exchange energy. The correlation energy is defined as the difference between the true energy and $\langle 0|H|0\rangle$. Since $\langle 0|H|0\rangle$ is the first term in the expansion of Eq. (14), the correlation energy in our formalism is given by Eq. (14) with the omission of the first term, namely,

$$\epsilon_c(\lambda,\mu_1,\mu_2,\ldots)=\sum_{\substack{m_1\ m_2\\ m_1+m_2+\cdots\ m_1'+m_2'+\cdots\neq 0}}\sum\cdots\sum_{m_1'}\sum_{m_2'}\cdots\ C(H:m_1,m_2,\ldots;m_1',m_2',\ldots)\mu_1^{m_1+m_1'}\mu_2^{m_2+m_2'}\cdots\Big/\left[N\frac{me^4}{2}\right]. \tag{30}$$

In the following calculations, we consider only those terms in ϵ_c which survive as r_s approaches zero, because we aim to obtain the result to the order of the constant term at high densities.

B. Calculation with trial function having two-particle correlations

We first calculate ϵ_c with the simplest trial function (16). In this case, the correlation functions we have to calculate are $C(H_0{:}m,0,0,\ldots;m',0,0,\ldots)$ and $C(V{:}m,0,0,\ldots,;m',0,0,\ldots)$. These two will be denoted by $C_{mm'}(H_0)$ and $C_{mm'}(V)$, respectively.

Let us calculate the first few terms in the expansion (30) by evaluating the Green's functions G_1 and G_2. The Green's function $G_1(H_0{:}\omega)$ is easily found to vanish from its definition. Thus $C_{10}(H_0)=C_{01}(H_0)=0$. The diagrams which contribute to $G_1(V{:}\omega)$ are shown in Figs. 2(a) and 2(b). Figure 2(a) is usually called the direct term and is

$$G_{1(a)}(V{:}\omega)=2\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}_1,\vec{k}_2}V(\vec{q})\tilde{V}(\vec{q})\left[\frac{1}{\omega+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}-\frac{1}{\omega-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\right], \tag{31}$$

where

$$\Delta(\vec{q};\vec{k}_1,\vec{k}_2)=\epsilon_{\vec{k}_1+\vec{q}}+\epsilon_{\vec{k}_2+\vec{q}}-\epsilon_{\vec{k}_1}-\epsilon_{\vec{k}_2}\ . \tag{32}$$

In Eq. (31) the integral should be done under the conditions $|\vec{k}_n|<k_F$ and $|\vec{k}_n+\vec{q}|>k_F$. In the following calculations, we do not write these conditions on $\vec{k}_n$ explicitly, but it should be understood that they always apply. The exchange term of Fig. 2(b) can be calculated similarly and the result is

$$G_{1(b)}(V{:}\omega)=\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}_1,\vec{k}_2}V(\vec{q})\tilde{V}(\vec{k}_1+\vec{k}_2+\vec{q})\left[-\frac{1}{\omega+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}+\frac{1}{\omega-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\right]. \tag{33}$$

In order to obtain $C_{10}(V)$ and $C_{01}(V)$ from $G_1(V{:}\omega)$, we examine the analytic behavior of the terms in Eqs. (31) and (33). The first term in each equation is analytic in the upper ω plane, while the second one is analytic in the lower ω plane. Thus $C_{10}(V)$ and $C_{01}(V)$ are, respectively, given by the first and second terms in each equation with $\omega=0$.

A little more complicated calculation is necessary for $G_2(H_0{:}\omega_1,\omega_2)$, but the diagrams to represent this function are almost the same as those for $G_1(V{:}\omega)$. For example, the direct term for $G_2(H_0{:}\omega_1,\omega_2)$ is given in Fig. 2(c) and is calculated by

$$G_{2(a)}(H_0{:}\omega_1,\omega_2)=\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}_1,\vec{k}_2}\tilde{V}(\vec{q})^2\Delta(\vec{q};\vec{k}_1\vec{k}_2)\left[-\frac{1}{\omega_1+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\frac{1}{\omega_2-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\right.$$
$$\left.-\frac{1}{\omega_1-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\frac{1}{\omega_2+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\right]. \tag{34}$$

The analytic behavior of this function shows that both $C_{20}(H_0)$ and $C_{02}(H_0)$ vanish, while $C_{11}(H_0)$ is given by Eq. (34) with $\omega_1=\omega_2=0$. We also obtain that $C(H_0{:}0,1,0,\ldots;0,0,0,\ldots)=C(H_0{:}0,0,0,\ldots;0,1,0,\ldots)=0$.

Since both $G_{1(a)}(V{:}\omega)$ and $G_{2(a)}(H_0{:}\omega_1,\omega_2)$ give the second-order terms with respect to the interaction, they are gathered together to yield the second-order, direct terms in the correlation energy

$$\epsilon_a^{(2)}=-\frac{3}{8\pi^5}\int d^3q\int d^3k_1\int d^3k_2\frac{1}{q^2+\vec{q}\cdot(\vec{k}_1+\vec{k}_2)}\left[2\mu_1\frac{1}{q^2\left[q^2+\lambda\frac{4}{\pi}\alpha r_s\right]}-\mu_1^2\frac{1}{\left[q^2+\lambda\frac{4}{\pi}\alpha r_s\right]^2}\right]. \tag{35}$$

Momenta are measured in units of k_F. In an analogous way, calculating $G_{2(b)}(H_0{:}\omega_1,\omega_2)$, given by Fig. 2(d), and upon combining this result with $G_{1(b)}(V{:}\omega)$, we obtain the second-order exchange energy:

$$\epsilon_b^{(2)}=\frac{3}{16\pi^5}\int d^3q\int d^3k_1\int d^3k_2\frac{1}{q^2+\vec{q}\cdot(\vec{k}_1+\vec{k}_2)}\left[2\mu_1\frac{1}{q^2+\lambda\frac{4}{\pi}\alpha r_s}\frac{1}{(\vec{q}+\vec{k}_1+\vec{k}_2)^2}\right.$$
$$\left.-\mu_1^2\frac{1}{(\vec{q}+\vec{k}_1+\vec{k}_2)^2+\lambda\frac{4}{\pi}\alpha r_s}\frac{1}{q^2+\lambda\frac{4}{\pi}\alpha r_s}\right]. \tag{36}$$

When we use the state $|l\rangle$ in Eq. (16) and take $\mu_1=1$, these second-order terms can be rewritten as

$$\epsilon_a^{(2)}+\epsilon_b^{(2)}=\sum_{l\neq0}-\frac{2\langle0|V|l\rangle\langle l|\tilde{V}|0\rangle-|\langle l|\tilde{V}|0\rangle|^2}{\tilde{E}_l-\tilde{E}_0}\Big/\left[N\frac{me^4}{2}\right]. \tag{37}$$

If we take V for $\tilde{V}$, Eq. (37) reduces to the result given by an ordinary perturbation approach.[8] However, it is very important to distinguish $\tilde{V}$ from V in order to remove the divergence of the integral which ordinary perturbation approaches suffer from. In the present formalism, the problem of the divergences is solved by the use of $\tilde{V}$ instead of V in constructing the trial function. Although we have shown this here only in second order, divergence does not appear in higher order either.

As r_s tends to zero, $\epsilon_a^{(2)}$ and $\epsilon_b^{(2)}$ have the following forms:

$$\epsilon_a^{(2)}=2\mu_1\left[A\ln\left[\lambda\frac{4}{\pi}\alpha r_s\right]+\delta\right]-\mu_1^2\left[A\ln\left[\lambda\frac{4}{\pi}\alpha r_s\right]+A+\delta\right], \tag{38}$$

and

$$\epsilon_b^{(2)}=(2\mu_1-\mu_1^2)E_{2b}, \tag{39}$$

where $A=2(1-\ln2)/\pi^2=0.0622$. δ was defined by Eq. (23) in Ref. 8 and has the value of -0.0508. E_{2b} was calculated analytically by Onsager *et al.*[11] and found to be 0.0484.

The only contribution of $G_2(V{:}\omega_1,\omega_2)$ which does not vanish at $r_s=0$ is given by the diagram shown in Fig. 2(e), usually called the ring diagram. This diagram can be calculated rather easily and the result is given by

$$
\begin{aligned}
G_2(V:\omega_1,\omega_2)=4\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}_1,\vec{k}_2,\vec{k}_2}\tilde{V}(q)^2V(q)\Bigg[&\frac{1}{\omega_1+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\,\frac{1}{\omega_1+\omega_2+i\epsilon-\Delta(\vec{q};\vec{k}_2,\vec{k}_3)}\\
&+\frac{1}{\omega_2+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_3)}\,\frac{1}{\omega_1+\omega_2+i\epsilon-\Delta(\vec{q};\vec{k}_2,\vec{k}_3)}\\
&+\frac{1}{\omega_1-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\,\frac{1}{\omega_1+\omega_2-i\epsilon+\Delta(\vec{q};\vec{k}_2,\vec{k}_3)}\\
&+\frac{1}{\omega_2-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_3)}\,\frac{1}{\omega_1+\omega_2-i\epsilon+\Delta(\vec{q};\vec{k}_2,\vec{k}_3)}\\
&-\frac{1}{\omega_1+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\,\frac{1}{\omega_2-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_3)}\\
&-\frac{1}{\omega_1-i\epsilon+\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}\,\frac{1}{\omega_2+i\epsilon-\Delta(\vec{q};\vec{k}_1,\vec{k}_3)}\Bigg]\ . \qquad (40)
\end{aligned}
$$

The analytic behavior of Eq. (40) shows that both $C_{20}(V)$ and $C_{02}(V)$ vanish, while $C_{11}(V)$ is given by the last two terms with $\omega_1=\omega_2=0$. Further, $C(V{:}0,1,0,\ldots;0,0,0,\ldots)$ and $C(V{:}0,0,0,\ldots;0,1,0,\ldots)$ can be obtained by the first two and the next two terms, respectively.

Similarly, we can calculate other correlation functions by evaluating the ring diagrams and obtain the result for $\epsilon_c(\lambda,\mu_1)$ as

$$
\epsilon_c(\lambda,\mu_1)=\epsilon_b^{(2)}+\sum_{n=1}^{\infty}(-\mu_1)^n\epsilon_{n+1}\left[\frac{1}{q^2}\right]+\frac{1}{2}\sum_{n=1}^{\infty}\mu_1^{2n}\epsilon_{2n}\left[\frac{1}{q^2+\lambda\frac{4}{\pi}\alpha r_s}\right]\ , \qquad (41)
$$

where $\epsilon_n[A(q)]$ is defined by

$$
\begin{aligned}
\epsilon_n[A(q)]=2\frac{3}{8\pi^5}\left[\frac{\alpha r_s}{\pi^2}\right]^{n-2}\int d^3q\int d^2k_1\cdots\int d^3k_nA(q)\left[q^2+\lambda\frac{4}{\pi}\alpha r_s\right]^{-n+1}&\frac{1}{q^2+\vec{q}\cdot(\vec{k}_1+\vec{k}_2)}\,\frac{1}{q^2+\vec{q}\cdot(\vec{k}_2+\vec{k}_3)}\cdots\\
\times\frac{1}{q^2+\vec{q}\cdot(\vec{k}_{n-2}+\vec{k}_{n-1})}\,&\frac{1}{q^2+\vec{q}\cdot(\vec{k}_{n-1}+\vec{k}_n)}\ . \qquad (42)
\end{aligned}
$$

The structure of $\epsilon_n(A)$ is shown in Fig. 3 by a Goldstone-type diagram. As r_s tends to zero, Eq. (41) has the following form:

$$
\begin{aligned}
\epsilon_c(\lambda,\mu_1)=&\ 2\mu_1(A\ln r_s+A\ln\lambda+C+A)-\mu_1^2(A\ln r_s+A\ln\lambda+C+2A)\\
&+\sum_{n=2}^{\infty}\frac{2}{n-1}\frac{1}{(2\lambda)^{n-1}}(-\mu_1)^nD_{n+1}+\sum_{n=2}^{\infty}\frac{1}{(2n-1)(2n-2)}\frac{1}{(2\lambda)^{2n-2}}\mu_1^{2n}D_{2n}\ . \qquad (43)
\end{aligned}
$$

The constant C is given by

$$
C=E_{2b}+\delta+A\left[\ln\frac{4}{\pi}\alpha-1\right]=-0.0901\ , \qquad (44)
$$

and D_n is defined by

$$
D_n=\frac{3}{\pi^2}\int_0^1x_1dx_1\int_0^1x_2dx_2\cdots\int_0^1x_ndx_n\frac{1}{x_1+x_2}\,\frac{1}{x_2+x_3}\cdots\frac{1}{x_{n-2}+x_{n-1}}\,\frac{1}{x_{n-1}+x_n}\ . \qquad (45)
$$

The value D_2 is just the same as A and D_3 is calculated $3(1-\pi^2/12)/(2\pi^2)=0.0270$. We have computed other D_n and found that when D_{n+1}/D_n is plotted as a function of n, it is almost constant, as shown in Fig. 4. In particular, when we calculate ϵ_c with an accuracy up to 10^{-4}, we can consider that for n larger than three, D_n is given by

$$
D_n=\beta^{n-3}D_3\ , \qquad (46)
$$

with $\beta=0.437$.

Substituting Eq. (46) into Eq. (43), we finally obtain

$$\epsilon_c(\lambda,\mu_1)=\mu_1(2-\mu_1)A\ \ln r_s+2\mu_1(1-\mu_1)(A+C)+C+\mu_1^2D_3/\beta+\mu_1(2-\mu_1)\left[A-\frac{D_3}{\beta}\right]\ln\lambda$$
$$+\frac{\mu_1D_3}{\beta}\left[\frac{\lambda}{\beta}-\frac{\mu_1}{2}\right]\ln\left[\lambda-\frac{\beta\mu_1}{2}\right]+\frac{\mu_1D_3}{\beta}\left[2-\frac{\mu_1}{2}-\frac{\lambda}{\beta}\right]\ln\left[\lambda+\frac{\beta\mu_1}{2}\right]\ . \qquad (47)$$

This expression is valid only when

$$\lambda\geq\tfrac{1}{2}\beta\mu_1=0.214\mu_1\ . \qquad (48)$$

If λ is smaller than $\beta\mu_1/2$, Eq. (43) diverges.

In order to minimize the leading term, i.e., the $\ln r_s$ term, in Eq. (47), we should take μ_1 to be unity first. Then we optimize λ to get the minimum of ϵ_c as

$$\min[\epsilon_c(\lambda,\mu_1)]=0.0622\ln r_s-0.0910\ , \qquad (49)$$

which occurs for $\lambda=0.329$. Compared with Eq. (1), Eq. (49) has the same $\ln r_s$ term. The error in the constant term is only 2.5%. At $r_s=0.1$, the difference between Eqs. (1) and (49) is less than 1%. We have investigated other possibilities for $\tilde{V}(\vec{q})$, but the choice of Eq. (27) for $\tilde{V}(\vec{q})$ seems to be the best, as long as $\tilde{V}$ is static and local, i.e., depends only on $\vec{q}$.

When $\tilde{V}$ is a nonlocal potential, written as $\tilde{V}(\vec{k}_1,\vec{k}_2;\vec{k}_1+\vec{q},\vec{k}_2-\vec{q})$ in Eq. (26) instead of $\tilde{V}(\vec{q})$, Eq. (41) is changed into the following form:

$$\epsilon_c(\lambda,\mu_1)=\epsilon_b^{(2)}+\sum_{n=1}^{\infty}(-\mu_1)^nE_{n+1}(V)\Big/\left[N\frac{me^4}{2}\right]+\frac{1}{2}\sum_{n=1}^{\infty}\mu_1^{2n}E_{2n}(\tilde{V})\Big/\left[N\frac{me^4}{2}\right]\ , \qquad (50)$$

where

$$E_n(V)=2^n\sum_{\vec{q}}\sum_{\vec{k}_1}\cdot\cdot\sum_{\vec{k}_n}\frac{V(\vec{q})\tilde{V}(\vec{k}_1,\vec{k}_2+\vec{q};\vec{k}_1+\vec{q},\vec{k}_2)\tilde{V}(\vec{k}_2,\vec{k}_3+\vec{q};\vec{k}_2+\vec{q},\vec{k}_3)\cdots\tilde{V}(\vec{k}_{n-1},\vec{k}_n+\vec{q};\vec{k}_{n-1}+\vec{q},\vec{k}_n)}{\Delta(\vec{q};\vec{k}_1,\vec{k}_2)\Delta(\vec{q};\vec{k}_2,\vec{k}_3)\cdots\Delta(\vec{q};\vec{k}_{n-1},\vec{k}_n)}\ , \qquad (51)$$

and

$$E_n(\tilde{V})=2^n\sum_{\vec{q}}\sum_{\vec{k}_1}\cdots\sum_{\vec{k}_n}\tilde{V}(\vec{k}_1,\vec{k}_2+\vec{q};\vec{k}_1+\vec{q},\vec{k}_2)\tilde{V}(\vec{k}_2,\vec{k}_3+\vec{q};\vec{k}_2+\vec{q},\vec{k}_3)\cdots\tilde{V}(\vec{k}_{n-1},\vec{k}_n+\vec{q};\vec{k}_{n-1}+\vec{q},\vec{k}_n)$$
$$\times\tilde{V}(\vec{k}_n,\vec{k}_1+\vec{q};\vec{k}_n+\vec{q},\vec{k}_1)/\Delta(\vec{q};\vec{k}_1,\vec{k}_2)\Delta(\vec{q};\vec{k}_2,\vec{k}_3)\cdots\Delta(\vec{q};\vec{k}_{n-1},\vec{k}_n)\ . \qquad (52)$$

We should also modify $\epsilon_b^{(2)}$ in an appropriate way, but $\epsilon_b^{(2)}$ always tends to E_{2b} as r_s goes to zero. If we include the $\vec{k}_1$ and $\vec{k}_2$ dependence of the energy denominator $\Delta(\vec{q};\vec{k}_1,\vec{k}_2)$ in the choice of $\tilde{V}$ to give a Jastrow-type trial function [Eqs. (17) and (18)], i.e., if we take

$$\tilde{V}(\vec{k}_1,\vec{k}_2+\vec{q};\vec{k}_1+\vec{q},\vec{k}_2)=\frac{4\pi e^2}{\Omega(q^2+\lambda q_{\mathrm{TF}}^2)}\frac{m\Delta(\vec{q};\vec{k}_1,\vec{k}_2)}{k_Fq}\ , \qquad (53)$$

the coefficient of the $\ln r_s$ term is given by

$$\frac{3}{2\pi^2}\mu_1-\frac{1}{\pi^2}\mu_1^2\ , \qquad (54)$$

which has the maximum value 0.057 at $\mu_1=\frac{3}{4}$. This is the result in Eq. (2). Thus, the failure of the Jastrow-type approach to reproduce the $\ln r_s$ term is due to the inappropriate treatment of the energy denominator $\Delta(\vec{q};\vec{k}_1,\vec{k}_2)$ which compels us to use an unsuitable nonlocal potential like Eq. (53). It is difficult to give a suitable nonlocal potential at this stage. This problem will be investigated in Sec. III E.

C. Variational calculation with a finite number of terms

Since the minimum value of ϵ_c is found in the convergent region of λ in Eq. (43), we can also obtain the result of Eq. (49) by the following procedure. First, we truncate the series in Eq. (43), keeping terms up to the order of μ_1^{2L}, namely,

$$\epsilon_{c,L}(\lambda,\mu_1)=2\mu_1(A\ \ln r_s+A\ \ln\lambda+C+A)-\mu_1^2(A\ \ln r_s+A\ \ln\lambda+C+2A)$$
$$+\sum_{n=2}^{2L}\frac{2}{n-1}\frac{1}{(2\lambda)^{n-1}}(-\mu_1)^nD_{n+1}+\sum_{n=2}^{L}\frac{1}{(2n-1)(2n-2)}\frac{1}{(2\lambda)^{2n-2}}\mu_1^{2n}D_{2n}\ . \qquad (55)$$

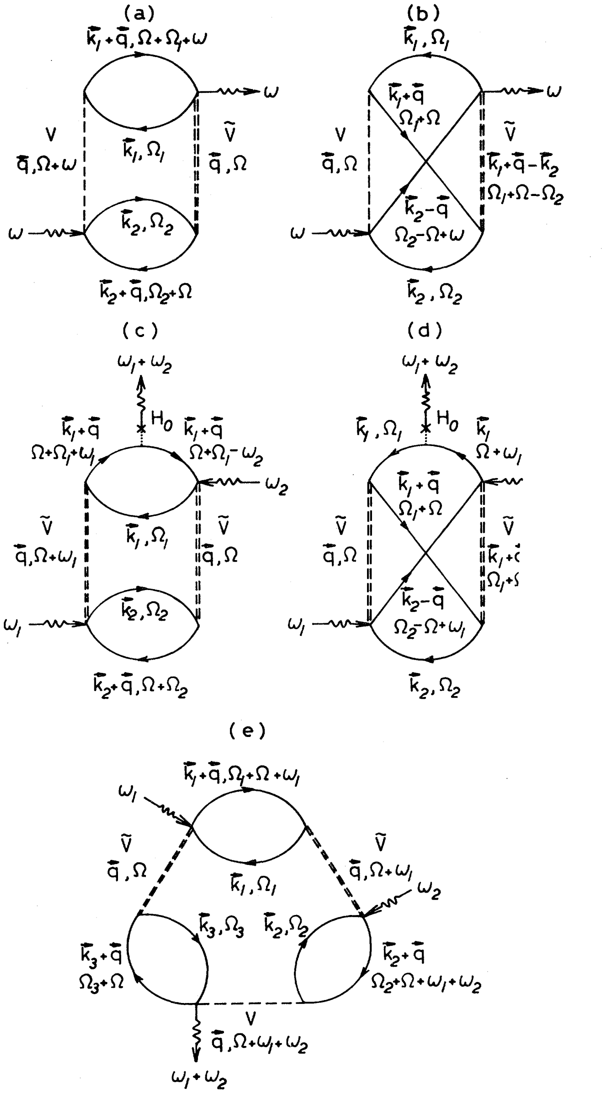

FIG. 2. Diagrams for the Green's functions $G_1(V:\omega)$, $G_2(H_0:\omega_1,\omega_2)$, and $G_2(V:\omega_1,\omega_2)$. (a) and (b) are, respectively, the direct and exchange terms of $G_1(V:\omega)$, while (c) and (d) are the corresponding ones of $G_2(H_0:\omega_1,\omega_2)$. (e) provides the only term for $G_2(V:\omega_1,\omega_2)$ which does not vanish at high densities.

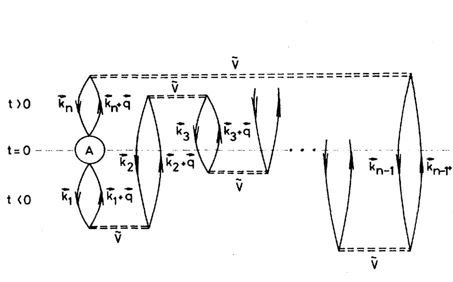

FIG. 3. Structure of $\epsilon_n(A)$ defined in Eq. (42) shown in a Goldstone-type diagram. Integer n is assumed to be odd here. Case of even n is given by a similar structure. Only change is that both $(\vec{k}_1,\vec{k}_1+\vec{q})$ and $(\vec{k}_n,\vec{k}_n+\vec{q})$ pairs are in the same region of t.

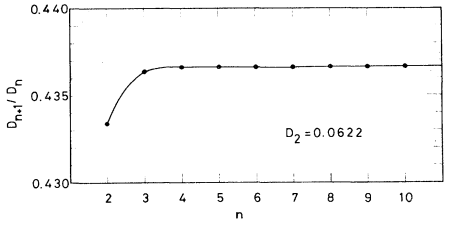

FIG. 4. Ratio D_{n+1}/D_n plotted vs n. Integral D_n is defined in Eq. (45).

Next, we find the minimum value of $\epsilon_{c,L}(\lambda,\mu_1)$. Then we can reach Eq. (49) by extrapolation:

$$\min[\epsilon_c(\lambda,\mu_1)]=\lim_{L\to\infty}\min[\epsilon_{c,L}(\lambda,\mu_1)]\ . \tag{56}$$

Physically, $\epsilon_{c,L}(\lambda,\mu_1)$ is the energy calculated for the case in which we neglect terms beyond the Lth order in the trial wave function (16). Thus, by making a variational calculation for $\epsilon_{c,L}(\lambda,\mu_1)$, we can obtain the optimized state under the condition that there are, at most, L uncorrelated ($2e$-$2h$) states.

Now, we apply this procedure to the present problem. As before, μ_1 can always be determined to be unity in order to minimize the $\ln r_s$ term. Then, for $L=1$, we obtain

$$\epsilon_{c,1}(\lambda,\mu_1=1)=A\ln r_s+A\ln\lambda+C+D_3/\lambda\ . \tag{57}$$

The optimized λ is easily found to be $D_3/A=0.434$ and the minimum value of $\epsilon_{c,1}(\lambda,\mu_1)$ is given by

$$\min[\epsilon_{c,1}(\lambda,\mu)]=0.0622\ln r_s-0.0798\ . \tag{58}$$

This is a pretty good result for ϵ_c. Compared to Eq. (1), the error is only 6% at $r_s=0.1$. Thus, this is even better than Eq. (2).

As L is increased, the minimum value of $\epsilon_{c,L}$ decreases, as shown in Fig. 5 in which the constant term in

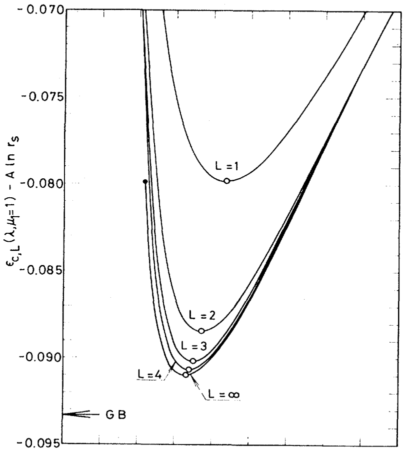

FIG. 5. Constant term in $\epsilon_{c,L}(\lambda,\mu_1=1)$ as a function of λ for several values of L. In the limit of $L=\infty$, $\epsilon_{c,L}$ diverges for λ less than 0.214. Effective potential is a static and local one, defined in Eq. (27). Point indicated by GB is the value obtained by Gell-Mann and Brueckner.

$\epsilon_{c,L}(\lambda,\mu_1=1)$ is plotted as a function of λ. The optimized λ also decreases, and the effective potential $\tilde{V}$ approaches the bare one V. In order to see the physical meaning of this result, let us go back in Fig. 3. The figure tells us that a larger number of polarization pairs $(\vec{k},\vec{k}+\vec{q})$ is taken into account in the calculation of ϵ_c with the increase of L. Therefore, we need less of the polarization effect in $\tilde{V}$, which results in the decrease of the optimized λ.

The above consideration also clarifies the reason why Eq. (57) has given such a good result as in Eq. (58) for ϵ_c. We have considered only a few terms in obtaining Eq. (57), but by choosing $\tilde{V}$ appropriately, we have been able to include most of the physically important effects which other infinite-order terms produce. The meaing of the fact is profound. It suggests that we can reduce the N-body problem to the two-body one, because when we calculate $\epsilon_{c,1}(\lambda,\mu_1)$, we treat the problem as if there were only two particles in the system. The presence of other particles is taken into account by the effective medium which produces an effective potential between the two particles. The properties of the effective interaction are determined variationally by minimizing the total energy of the two-body system $\epsilon_{c,1}(\lambda,\mu_1)$.

The obtained minimum energy of $\epsilon_{c,1}$ is larger than the exact value of ϵ_c in the present case. At present, it is not proved mathematically that $\min(\epsilon_{c,1})$ is always an upper bound of the true ground-state energy. However, we can expect the above statement to be true from the following consideration. First, the true ground-state energy is always smaller than $\lim_{L\to\infty}\min(\epsilon_{c,L})$. Second, $\min(\epsilon_{c,L})$ is expected to be a monotonically decreasing function of L, because when we increase L, we treat more particles in the system. Any increase of degrees of freedom always leads to a decrease of the minimum value of the energy.

From a physical point of view, however, it is more important that $\min(\epsilon_{c,1})$ is found to be a good approximation to the true value of ϵ_c than that $\min(\epsilon_{c,1})$ is proved to be an upper bound of ϵ_c. This is particularly useful when we apply this new method to more complicated systems. It is probably impossible to sum a series like Eq. (43), or even to obtain the form of higher-order terms in a complicated system, but we can expect to get a physically meaningful result by calculating a few terms in the expansion of the energy expectation value with a suitably chosen $\tilde{V}$.

D. Inclusion of higher-order corrections

One of the advantages of the present method is that even if we cannot choose the best form of $\tilde{V}$, we can improve the result by changing the wave function from $|\Phi_0^{(j)}\rangle$ to $|\Phi_0^{(j+1)}\rangle$. The actual calculations can be done by the extension of the procedure proposed in Sec. III C. Namely, we calculate $\epsilon_{c,L}(\lambda,\mu_1,\ldots,\mu_j)$, defined by

$$\epsilon_{c,L}(\lambda,\mu_1,\ldots,\mu_j)=\sum_{m_1}\sum_{m_2}\cdots\sum_{m_j}\sum_{m_1'}\sum_{m_2'}\cdots\sum_{m_j'}C(H:m_1,\ldots,m_j;m_1',\ldots,m_j')\mu_1^{m_1+m_1'}\cdots\mu_j^{m_j+m_j'}/(Nme^2/2)\,, \tag{59}$$

where the following restriction is imposed on the sum:

$$1\le m_1+\cdots+m_j+m_1'+\cdots+m_j'\le 2L\ . \tag{60}$$

Then we take the limit of $\min[\epsilon_{c,L}(\lambda,\mu_1,\ldots,\mu_j)]$ as $L\to\infty$.

The case of $L=1$ is shown in Fig. 6. The constant term in the energy and the optimized values of λ, μ_2, μ_3, and μ_4 are plotted as a function of j. (As before, μ_1 is always taken to be unity.) As j is increased, more polarization effects are included in such a way as shown in Fig. 7 which gives a Goldstone-type diagram for

$$C(V:0,\ldots,m_l=1,\ldots,0;0,\ldots,m_{l'}=1,\ldots,0)\ .$$

Thus, we need less of the polarization effect of $\tilde{V}$ and obtain the smaller value for the optimized λ with the in-

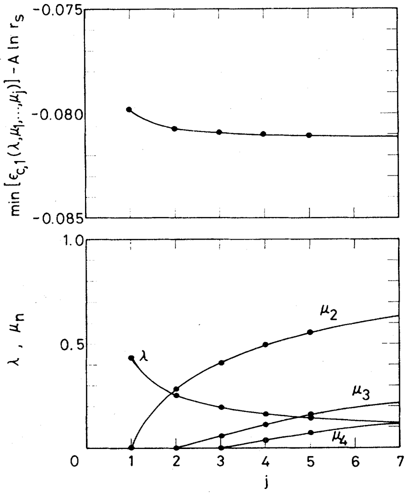

FIG. 6. Minimum energy of $\epsilon_{c,1}(\lambda,\mu_1,\ldots,\mu_j)$, optimized λ, and μ_n at each step of j. Parameter μ_1 is always unity.

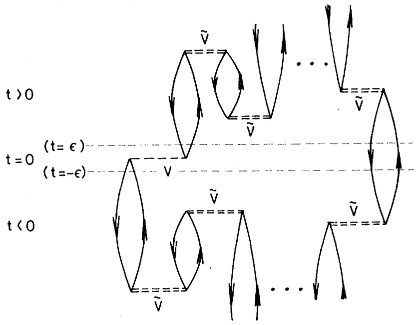

FIG. 7. Goldstone-type diagram for $C(V:0,\ldots,m_l=1,\ldots,0;0,\ldots,m_{l'}=1,\ldots,0)$ which contributes to Eq. (59).

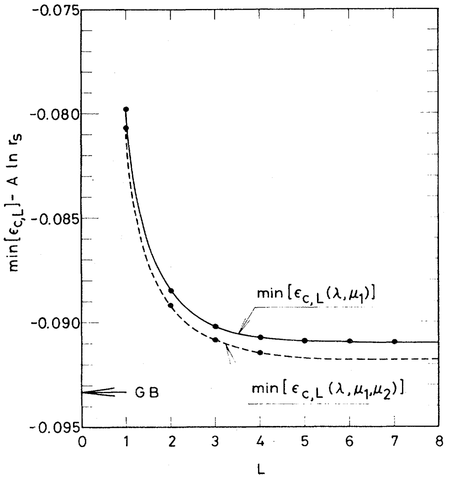

FIG. 8. Minimum energy of $\epsilon_{c,L}$ plotted vs L. Case of $|\Phi_0^{(1)}\rangle$ is shown by the solid curve, while that of $|\Phi_0^{(2)}\rangle$ is by the broken curve.

crease of j. The parameter μ_l depends how good $\tilde{V}$ is. For example, if $\tilde{V}$ takes almost all physically important effects into account, $\mu_2,\mu_3,\ldots$ will be very small. In the case of Fig. 6, $\tilde{V}$ includes less of the polarization effect with the increase of j, which results in the increase of the values of μ_2, μ_3, and μ_4. Since μ_l is correlated with λ in this manner, it is quite important not to take $\mu_l=1$ from the outset, but to take μ_l as a variational parameter.

Let us compare Fig. 7 with Fig. 3. Both diagrams treat the polarization effect, but they are a little different. There are many polarized pairs at $t=0$ in Fig. 3, but only one $(2e\text{-}2h)$ state is present at $t=\pm\epsilon$ in Fig. 7. Thus, even if we calculate with $|\Phi_0^{(2)}\rangle$, $|\Phi_0^{(3)}\rangle$, and so on, these wave functions have only two-particle correlations in the high-density limit in which only the ring diagrams are important. Therefore, the value, obtained as

$$\lim_{j\to\infty} \min[\epsilon_{c,1}(\lambda,\mu_1,\ldots,\mu_j)]=0.0622\ln r_s-0.0812\ , \tag{61}$$

corresponds to the case in which we employ a trial function with the best choice of $\tilde{V}$, but having only one $(2e\text{-}2h)$ state.

The fact that there is a difference between Eqs. (1) and (61) indicates that many $(2e\text{-}2h)$ states are necessary in order to reproduce Eq. (1). Physically, this can be explained as follows. In the high-density limit, there are many particles in the range of the interaction, estimated by $1/q_{\rm TF}$, because the ratio of the interparticle spacing to $1/q_{\rm TF}$ is proportional to $\sqrt{r_s}$ and very small for $r_s \ll 1$. Therefore, the polarization of many particles will play an important role in the calculation of ϵ_c.

As L is increased, we can include larger number of $(2e\text{-}2h)$ states and obtain an improvement of the result for ϵ_c. The case of $j=2$ is shown in Fig. 8. The minimum values of $\epsilon_{c,L}(\lambda,\mu_1,\mu_2)$ are plotted as a function of L by a broken curve. (A solid curve corresponds to the case of $j=1$.) By extrapolation, we can estimate that

$$\min[\epsilon_c(\lambda,\mu_1,\mu_2)]=0.0622\ln r_s-0.0918\ . \tag{62}$$

Thus the result of Eq. (49) is improved. The minimum values of $\epsilon_{c,L}(\lambda,\mu_1,\ldots,\mu_j)$ are plotted in Fig. 9 as a function of the optimized λ for each j and L. The result of Eq. (1) seems to be obtained at $\lambda=0$, $j=\infty$, and $L=\infty$.

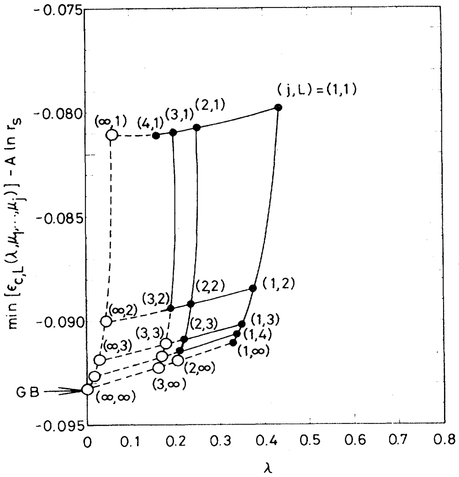

FIG. 9. Minimum energy $\epsilon_{c,L}(\lambda,\mu_1,\ldots,\mu_j)$ vs optimized λ for several values of j and L. Solid points are given by calculations, while the points indicated by open circles are estimated by extrapolation.

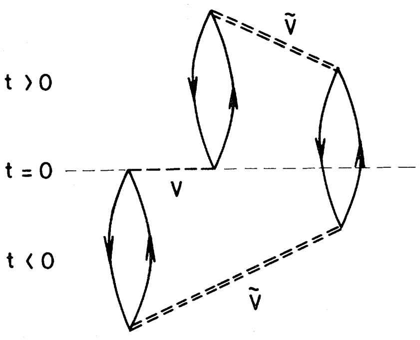

FIG. 10. Schematic representation of the process in Fig. 7 by introducing a time-dependent $\tilde{V}$.

E. Inclusion of dynamic effects in the effective interaction

The analysis in Sec. III D suggests that Eq. (1) may be reproduced with the simplest wave function (16), if we can find a better form of $\tilde{V}$ than Eq. (27). Let us return to Fig. 7 in order to see what kind of improvement should be made. At $t=-\epsilon(t=\epsilon)$, there are two electron-hole pairs, but these two are created (annihilated) at different times. Thus, if we wish to simulate the process in Fig. 7 by using one effective interaction in each region of t, we have to employ a time-dependent $\tilde{V}$ as shown by a Goldstone-type diagram in Fig. 10, though, strictly speaking, this diagram is not a Goldstone-type diagram, because the interaction line is not vertical to the time axis.

The advantage in formulation of the whole problem with the use of Feynman diagrams is that each term in Eq. (14) can be calculated easily even if $\tilde{V}$ depends on t. Of course, once $\tilde{V}$ has a dynamic effect, $\tilde{V}$ cannot be written in the form of Eq. (26). However, this equation is not necessary for actual calculations, because $\tilde{V}$ appears only in the calculation of the Green's functions G_l which is expressed by Feynman diagrams. We need only the ω-dependent form of $\tilde{V}$. The only complexity that occurs is that we have to introduce an external frequency at each end of $\tilde{V}$ in the diagrams, as shown in Fig. 11 for $G_1(H_0)$.

Since both ends of each effective interaction lie in the same region of t in Fig. 10, we will consider only the terms in which both end points of one effective interaction are characterized by the same analytic behavior. Then, we obtain $\epsilon_c(\lambda,\mu_1)$ in just the same form as in Eq. (50). The only difference is that in Eqs. (51) and (52) $\tilde{V}(\vec{k},\vec{k}'+\vec{q};\vec{k}+\vec{q},\vec{k}')$ should be replaced by

$$\tilde{V}^*(\vec{k},\vec{k}'+\vec{q};\vec{k}+\vec{q},\vec{k}')=\Delta(\vec{q};\vec{k},\vec{k})\int_0^\infty \frac{d\omega}{\pi}\tilde{V}(\vec{k},\vec{k}'+\vec{q};\vec{k}+\vec{q},\vec{k}':i\omega)\frac{\omega^2+\Delta(\vec{q};\vec{k})\Delta(\vec{q};\vec{k}')}{[\omega^2+\Delta(\vec{q};\vec{k})^2][\omega^2+\Delta(\vec{q};\vec{k}')^2]}\ , \tag{63}$$

with

$$\Delta(\vec{q};\vec{k})=\epsilon_{\vec{k}+\vec{q}}-\epsilon_{\vec{k}}\ . \tag{64}$$

Therefore, the introduction of dynamic effects does not produce any new terms in $\epsilon_c(\lambda,\mu_1)$. However, Eq. (63) is very important and provides a good prescription to give a proper nonlocal potential. The nonlocal behavior appearing in Eq. (63) is due to the retarded effects of the interaction.

We have investigated several forms for $\tilde{V}$ and found that a very good, if not the best, choice for $\tilde{V}$ is given by

$$\tilde{V}(\vec{k},\vec{k}'+\vec{q};\vec{k}+\vec{q},\vec{k}':i\omega)=\frac{4\pi e^2}{\Omega[q^2+\lambda q_{\mathrm{TF}}^2 f(q/2k_F,im\omega/k_Fq)]}\times\frac{2\Delta(\vec{q};\vec{k})\Delta(\vec{q};\vec{k}')}{\omega^2+\Delta(\vec{q};\vec{k})\Delta(\vec{q};\vec{k}')}\ , \tag{65}$$

with the random-phase approximation function $f(z,iu)$, given by[22]

$$f(z,iu)=\tfrac{1}{2}\left[1+\frac{1-z^2+u^2}{4z}\ln\frac{(1+z)^2+u^2}{(1-z)^2+u^2}-u\tan^{-1}\frac{2u}{u^2+z^2-1}\right]\ , \tag{66}$$

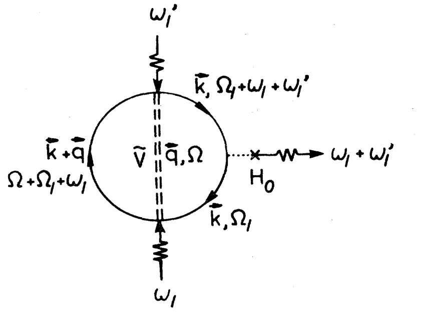

FIG. 11. Diagram for the Green's function $G_1(H_0:\omega_1,\omega_1')$. Effective potential $\tilde{V}$ depends on Ω as well as $\vec{q}$.

where the branch of $\tan^{-1}z$ is between 0 and π. The process to minimize $\epsilon_c(\lambda,\mu_1)$ is the same as shown in Sec. III C. The constant term in $\epsilon_{c,L}(\lambda,\mu_1=1)$ is plotted as a function of λ in Fig. 12 for several values of L. By taking the limit of $L=\infty$, we can estimate

$$\min[\epsilon_c(\lambda,\mu_1)]=0.0622\ln r_s-0.0929\ , \tag{67}$$

which occurs at $\lambda=0.585$. The dynamic interaction like Eq. (65) gives a weaker screening effect than the static one [Eq. (27)] for the processes with high ω. In order to compensate such a situation, a stronger coupling constant λ was necessary in Eq. (67) than that in Eq. (49).

There is still a difference between Eqs. (1) and (67), but we believe that if we can find the form for $\tilde{V}$ for which $\min[\epsilon_{c,1}(\lambda,\mu_1)]$ reproduces Eq. (61), we can reach Eq. (1). In the case of $\tilde{V}$ in Eq. (65), we have obtained

$$\min[\epsilon_{c,1}(\lambda,\mu_1)]=0.0622\ln r_s-0.0810\ , \tag{68}$$

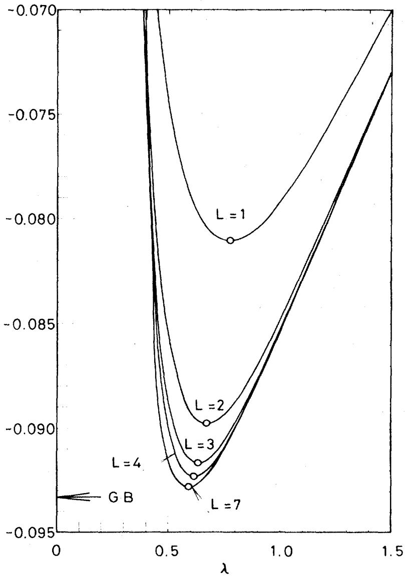

FIG. 12. Constant term in $\epsilon_{c,L}(\lambda,\mu_1=1)$ as a function of λ for several value of L. Effective potential is a nonlocal one, given by Eq. (65) with Eq. (66).

and there is a small difference between Eqs. (61) and (68).

In the rest of this section, we will show that our trial functions can encompass the states of collective excitations as well as those of individual ones, when dynamic effects are included in $\widetilde{V}$. In the derivation of Eq. (67), we have neglected the terms having at least one effective interaction whose end points are characterized by the different analytic behavior. In order to see the physical meaning of those neglected terms, $G_1(H_0)$, given in Fig. 11, is calculated as

$$G_1(H_0:\omega_1,\omega_1')=2\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}}\int_0^\infty \frac{d\omega}{\pi}\mathrm{Im}\widetilde{V}^R(\vec{k},\vec{k}+\vec{q};\vec{k}+\vec{q},\vec{k}:\omega)\left[\frac{\epsilon_{\vec{k}+\vec{q}}}{\omega_1-i\epsilon+\omega+\Delta(\vec{q};\vec{k})}\frac{1}{\omega_1'+i\epsilon-\omega-\Delta(\vec{q};\vec{k})}\right.$$
$$\left.-\frac{\epsilon_{\vec{k}}}{\omega_1+i\epsilon-\omega-\Delta(\vec{q};\vec{k})}\frac{1}{\omega_1'-i\epsilon+\omega+\Delta(\vec{q};\vec{k})}\right], \qquad (69)$$

where the superscripts R refers to the retarded form. The analytic behavior of Eq. (69) indicates that $G_1(H_0)$ does not vanish for the process shown in Fig. 13. At $t=0$ in this diagram, there are one electron-hole pair and the state introduced by $\mathrm{Im}\widetilde{V}^R$. Since the main contribution of $\mathrm{Im}\widetilde{V}^R$ comes from plasmons in the electron gas, it is anticipated that collective excitations might be included explicitly in $|\Phi_0^{(1)}\rangle$.

In order to show that this is the case, we calculate the kinetic energy of the trial function, defined by

$$|\Phi_0^{(P)}\rangle=|0\rangle-\mu_p|\text{plasmon}\rangle+\cdots$$
$$=|0\rangle-\mu_p\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k},\sigma}\frac{\lambda_{\vec{q}}}{\omega_p+\Delta(\vec{q};\vec{k})}\mathscr{A}^\dagger_{\vec{k}+\vec{q},\sigma}\mathscr{A}_{\vec{k},\sigma}\mathscr{B}^\dagger_{-\vec{q}}|0\rangle+\cdots, \qquad (70)$$

where $\lambda_{\vec{q}}$ is the plasmon-electron coupling constant, $\mathscr{B}^\dagger_{-\vec{q}}$ is the creation operator of the plasmon and is assumed to commute with both $\mathscr{A}_{\vec{k}\sigma}$ and $\mathscr{A}^\dagger_{\vec{k}\sigma'}$, and ω_P is the plasmon energy. The result is given by

$$\langle\text{plasmon}|H_0|\text{plasmon}\rangle_C=2\sum_{\vec{q}\neq\vec{0}}\sum_{\vec{k}}\frac{|\lambda_{\vec{q}}|^2\Delta(\vec{q};\vec{k})}{[\omega_p+\Delta(\vec{q};\vec{k})]^2}. \qquad (71)$$

When we employ the plasmon-pole approximation for $\mathrm{Im}\widetilde{V}^R$ as

$$-\mathrm{Im}\widetilde{V}^R(\vec{k},\vec{k}+\vec{q};\vec{k}+\vec{q},\vec{k}:\omega)=\pi|\lambda_{\vec{q}}|^2\delta(\omega-\omega_p), \qquad (72)$$

and make both ω_1 and ω_1' zero, Eq. (69) is reduced to Eq. (71).

Similar terms can be obtained in other Green's functions $G_l(H_0)$ and $G_l(V)$. Collecting these contributions and making a variational procedure, we find that the total energy of the ground state is minimized at $\mu_p=0$. Thus, it is verified variationally that we should neglect these contributions in the calculation of ϵ_c. The above result also means that if we operate with $\mathscr{B}_{\vec{q}}$ on $|\Phi_0^{(1)}\rangle$, we obtain

$$\mathscr{B}_{\vec{q}}|\Phi_0^{(1)}\rangle=0. \qquad (73)$$

This is just the condition which Bohm and Pines[9] assumed for the true ground state.

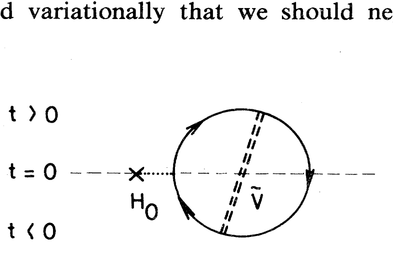

FIG. 13. Schematic representation of the term in $G_1(H_0)$ which does not vanish for t dependent $\widetilde{V}$.

IV. CONCLUSIONS AND DISCUSSION

In this paper, we have proposed a new variational method which combines two complementary approaches, namely, perturbation-theoretic and variational, to the many-body problem. The essential point of the present method is the use of a variationally determined effective potential. This is a striking difference from other methods. We have checked the usefulness of our method by applying it to the high-density electron gas.

There are many advantages in this new method, but the following two points are particularly important. First, the problem of divergences never occurs and we always attack the problem from the convergent region. Thus, even if we calculate only a finte number of terms, we can obtain a physically meaningful result. Second, improvement can be made by either a better choice of the effective potential, or the systematic change of the trial function. If a very good effective potential is found from physical intuition, we will obtain an excellent result by calculating a very small number of terms. If it is difficult to find a good effective potential, we can improve the result gradually and systematically by increasing the number of terms which we calculate. By plotting sequential results and employing extrapolation, we can find the value which a summation of an infinite number of terms produces.

It can be seen from the definition of the trial function, Eq. (5a), that the present method works well for weak coupling. The applicability of this method to strong coupling needs to be checked in the future. However, we believe this method to be useful even in the strong-coupling case. One reason is that if strong-coupling effects can be expressed by renormalization of the interaction and the single-particle states, as assumed, for example, in strong-coupling superconductivity,[23] we can take these effects into account by an appropriate choice of $\widetilde{H}_0$ and $\widetilde{V}$. Another reason is that Eq. (5a) is an expansion with respect to $\widetilde{V}$, not to the bare interaction V. Therefore, even if V becomes very large, the series in Eq. (5a) will converge rapidly, when $\widetilde{V}$ remains small. This is just the case of the electron gas. The smallness parameter in the expansion with respect to V is r_s and becomes large for r_s larger than unity. However, it is something like $r_s/(1+r_s)$ for the expansion with respect to the screened Coulomb interaction $\widetilde{V}$. This quantity remains less than unity, even if r_s is larger than unity. Consider the approach due to Jastrow[4] which has been employed in the strong-coupling situation.[5] The essential physics is to treat the two-particle correlations properly. Since two-particle correlations are also treated well in our trial function $|\Phi_0^{(1)}\rangle$, we can hope that our method will be useful in problems for which the Jastrow-type approach has been successful.

A system to which the present method can be applied in the future is the electron gas at metallic densities. One advantage of the present method is that we can include the ω dependence in the effective interaction. In the case of the Jastrow-type approach, inclusion of the ω dependence would be very difficult. In perturbation-theoretic approaches, inclusion of the ω dependence is also easy, but the problem is that there is not a guiding principle for determining the form of the ω dependence. At most there are a few sum rules to be employed. In our method, the most suitable form of $\widetilde{V}$ can be determined by the variational procedure. Another advantage is that we can investigate the effect of nonlocality, as partly shown in Sec. III E.

Another area for application of the present method is the condensed phase of the electron gas. The method can be used to study spin-density-wave and charge-density-wave states in much the same way as for the normal phase. The possibility of such phases of the electron gas has been extensively discussed.[24] However, a quantitative study of correlation in these phases remains a very interesting problem. Another area is superconductivity. The present author has already pointed out that when the usual perturbation-theoretic approach is employed in the solution of the gap equation (with the interaction treated in the lowest order), the low-density electron gas may become superconducting with the mediation of plasmons.[25] The possibility of superconductivity without mediation from phonons or excitons should be investigated in more detail.

Since the present method is related to fundamental problems in quantum mechanics, its application need not be restricted to solid-state physics. It may also be applied to nuclear and elementary-particle physics, e.g., the nuclear matter problem.

ACKNOWLEDGMENTS

The author would like to thank Professor A. W. Overhauser for suggesting this work, for stimulating discussions throughout the course of this reasearch, and for a critical reading of the manuscript. He also wishes to thank Dr. M. Huberman for useful discussions. The author is grateful to the National Science Foundation (NSF) and the NSF Materials Research Laboratories Program for support.

---

[*] Present address: Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106.

[1] See, for example, A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, 1963).

[2] R. P. Feynman, Phys. Rev. 76, 749 (1949); 76, 769 (1949).

[3] A. W. Overhauser and J. T. Tsai, Phys. Rev. B 13, 607 (1976).

[4] R. Jastrow, Phys. Rev. 98, 1479 (1955).

[5] See, for example, V. R. Pandharipande and R. B. Wiringa, Rev. Mod. Phys. 51, 821 (1979); in *The Many-Body Problem, Jastrow Correlations Versus Brueckner Theory,* edited by R. Guardiola and J. Ros (Springer, Berlin, 1980).

[6] E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).

[7] F. Coester and H. Kümmel, Nucl. Phys. 17, 477 (1960).

[8] M. Gell-Mann and K. Brueckner, Phys. Rev. 106, 364 (1957).

[9] D. Bohm and D. Pines, Phys. Rev. 82, 625 (1951); 85, 338 (1952); 92, 609 (1953); 92, 626 (1953).

[10] K. Sawada, Phys. Rev. 106, 372 (1957); J. Hubbard, Proc. R. Soc. London, Ser. A 243, 336 (1957); P. Nozieres and D. Pines, Phys. Rev. 111, 442 (1958); Nuovo Cimento 9, 470 (1958); almost all the above papers can be found in D. Pines, *The Many-Body Problem* (Benjamin, New York, 1961).

[11] In the original paper of Gell-Mann and Brueckner, the constant term in the correlation energy was -0.096 instead of -0.0933. However, L. Onsager, L. Mittag, and M. J. Stephen [Ann. Phys. (N.Y.) 18, 71 (1966)] pointed out that the second-order exchange energy could be calculated rigorously and that the value of it was larger by 0.0024 than that calculated numerically by Gell-Mann and Brueckner.

[12] R. F. Bishop and K. H. Lührmann, Phys. Rev. B 17, 3757 (1978).

[13] T. Gaskell, Proc. Phys. Soc. London 77, 1182 (1961); 80, 1091 (1962).

[14] M. S. Becker, A. A. Broyles, and T. Dunn, Phys. Rev. 175, 224 (1968); R. Mannier, Phys. Rev. A 6, 393 (1972); D. K. Lee and F. H. Ree, *ibid.* 6, 1218 (1972); G. Keiser and F. Y. Wu, *ibid.* 6, 2369 (1972); F. A. Stevens and M. A. Pokrant, *ibid.*

8, 990 (1973); J. D. Talman, *ibid.* 10, 1333 (1974); S. Chakravarty and C. W. Woo, Phys. Rev. B 13, 4815 (1976); D. Ceperley, *ibid.* 18, 3126 (1978); L. J. Lantto, *ibid.* 22, 1380 (1980).

[15] J. G. Zabolitzky, Phys. Rev. B 22, 2353 (1980).

[16] In the original result of Gaskell, the constant term was -0.132. However, this was the value without the second-order exchange energy.

[17] L. J. Lantto, E. Krotscheck, and R. A. Smith, *Recent Progress in Many-Body Theories,* Springer Lecture Notes in Physics, Vol. 142; edited by J. G. Zabolitzky, M. de Llano, M. Fortes, and J. W. Clark (Springer, Berlin, 1981), p. 287.

[18] In this paper we employ units in which $\hbar=c=k_B=1$.

[19] J. Goldstone, Proc. R. Soc. London Ser. A 239, 267 (1957).

[20] W. Macke, Z. Naturforsch. 5a, 192 (1950).

[21] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).

[22] J. Lindhard, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. 28, 8 (1954).

[23] See, for example, D. J. Scalapino, *Superconductivity,* edited by R. D. Parks (Dekker, New York, 1969), Vol. 1, p. 449.

[24] A. W. Overhauser, Phys. Rev. Lett. 4, 462 (1960); Phys. Rev. 128, 1437 (1962); 167, 691 (1968); Adv. Phys. 27, 343 (1978).

[25] Y. Takada, J. Phys. Soc. Jpn. 45, 786 (1978).