MANY-ELECTRON THEORY

BY

STANLEY RAIMES
Department of Mathematics
Imperial College
University of London

1972

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM - LONDON

Library of Congress Catalog Card Number: 76-183283

North-Holland ISBN: 0 7204 0242 5

American Elsevier ISBN: 0 444 10353 8

PUBLISHERS:
NORTH-HOLLAND PUBLISHING COMPANY–AMSTERDAM
NORTH-HOLLAND PUBLISHING COMPANY LTD.–LONDON

SOLE DISTRIBUTORS FOR THE U.S.A. AND CANADA:
AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE, NEW YORK, N.Y. 10017

PRINTED IN THE NETHERLANDS

My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless, when he has climbed out through them, on them, over them. (He must so to speak throw away the ladder, after he has climbed up on it.)

He must surmount these propositions; then he sees the world rightly.

Whereof one cannot speak, thereof one must be silent.

Ludwig Wittgenstein *Tractatus Logico-Philosophicus*

PREFACE

Many-body theory is undoubtedly one of the most difficult branches of quantum mechanics, yet it is of growing importance throughout the whole of theoretical physics, and particularly in the theory of the solid state. Many students and research workers, baffled and intrigued by the rash of Feynman graphs and Green functions which in recent years has appeared on the face of physics, would like to gain some familiarity with the subject. However, they often find that not only original papers, but even books purporting to be written for beginners, present an impenetrable barrier. This applies to workers in fields such as theoretical metallurgy and quantum chemistry, as well as to specialists in solid-state theory who started their research in the days before a knowledge of quantum field theory became a *sine qua non*. It may be added that even graduate students of theoretical physics, who have had the benefit of a more sophisticated training, are often obliged to accept without proof statements whose origins are shrouded in mystery and whose truth is by no means apparent. To some of those who perceive it, this causes a psychological impediment to further progress.

This book has been written with all such people in mind. The only prerequisites to its understanding are a knowledge of elementary quantum mechanics and of the sort of mathematics now normally given in undergraduate courses to students of physics, metallurgy and chemistry. As far as the quantum mechanics is concerned, a reading of my *Wave Mechanics of Electrons in Metals* (North-Holland, 1961) would be more than adequate. Some mathematics which may not have been encountered in undergraduate courses is given in the Appendixes to the present work.

The book does not presume to cover all aspects of many-body theory. Only many-electron systems are considered, but the techniques described apply equally well to all many-fermion systems, and, with slight modifications, to many-boson systems also. Some well-developed branches of many-electron theory, such as the dielectric theory of plasma oscillations and the equation-of-motion method, are not included, since they parallel the methods described. The intention has been to force *one* path through the jungle, leaving no obstruction which requires a mental helicopter to surmount. It is not anticipated that anyone who reads the book will then be in a position to do research in the subject or even to read all original papers with ease. However, the contents may be adequate for those whose interest in the subject is secondary, and for the others the book will serve as a solid foundation for studying more comprehensive, if less comprehensible, texts.

Some problems and exercises relating to each chapter are collected at the end of the book, but are mainly confined to those which emphasize points of the text and which can be done without consulting external sources.

Although the primary purpose of the book is pedagogical, it is not a mere reshuffle of existing material. Many of the proofs and demonstrations have not, to my knowledge, appeared in print before, and few of the detailed explanations are available in any other work. For this reason, if for no other, it is probable that errors and inelegancies exist. However, no part of the book has been read by experts before publication, so all its shortcomings must be laid at my door. I extend my gratitude in advance to those who will point out to me any flagrant errors and suggest possible improvements.

Imperial College S. RAIMES

CONTENTS

CHAPTER 1

RÉSUMÉ OF THE MANY-ELECTRON PROBLEM

1.1. Introduction

Calculation of the properties of systems containing many electrons becomes almost impossibly complicated when an attempt is made to include the interaction of the electrons. Fortunately, in some cases it is possible to obtain reasonable results by neglecting this interaction altogether, or by considering only its average effect, so that the electrons are assumed to move more or less independently of one another. In other cases this *independent electron approximation* is inadequate.

The many-electron problem is the problem of taking proper account of the interaction of the electrons. It thus covers a large number of problems, but the one which will concern us most in this book is that of calculating the total energy of a many-electron system, or, more specifically, that part of it known as the *correlation energy*†. The techniques we shall describe, such as perturbation theory and the use of Green functions, are of course applicable to the calculation of other properties of the system.

In principle the methods to be described are suitable for any many-electron system, such as an atom, a molecule or a solid, but in practice insuperable computational difficulties, if nothing worse, may be encountered. Each system must therefore be considered on its merits. We shall attempt initially to keep the treatment as general as possible, but in

---

† See §4.4 and ch. 9. An elementary discussion of the many-electron problem is given in the author's book: Raimes, S., 1961, *The Wave Mechanics of Electrons in Metals* (North-Holland, Amsterdam). Since a number of references will be made to this work it will be convenient in future to denote it simply by WM. The correlation energy, for example, is defined and discussed in WM, §9.5.

the ultimate application we shall have in mind a metal such as sodium, which can be thought of as a gas of almost-free valence electrons moving in a lattice of positive ion-cores, and we shall give a detailed discussion of the free-electron approximation to such a system.

In the present chapter we begin by looking at the many-body problem in the light of elementary wave mechanics and deriving some results which will be needed for comparison purposes when describing the more sophisticated methods of later chapters.

1.2. The Schrödinger Equation

The system we wish to consider is a general one consisting of N electrons moving in an external electrostatic field, due to which the potential energy of an electron at position $\boldsymbol{r}$ is $V(\boldsymbol{r})$. The first approximation we make is to take into account only the electrostatic interactions of the electrons with each other and with the external field, and to ignore magnetic interactions, which are of a very much smaller order of magnitude. The quantum-mechanical Hamiltonian operator of the system is then

$$H = H_0 + H', \tag{1.1}$$

where

$$H_0 = \sum_{i=1}^{N} \left[-\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i) \right], \tag{1.2}$$

and

$$H' = \tfrac{1}{2} \sum_{i \neq j}^{N} \sum^{N} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{1.3}$$

H_0 is the Hamiltonian the system would have if the electrons did not interact, and H' is the potential energy due to the electrostatic or Coulomb interaction of the electrons.

The Schrödinger equation of the system is

$$H\Psi = E\Psi, \tag{1.4}$$

the wave function Ψ being a function of the coordinates, both space and spin, of all the electrons. The eigenvalues E of this equation are the energy levels of the system. We shall be particularly interested in the energy and wave function of the *ground state*, that is, the state of lowest energy – all other states are called *excited states*.

Were it not for the presence of H', the solution of eq. (1.4) would be a relatively simple matter, for the equation would then be separable. Thus, if the equation were

$$H_0\Psi = E\Psi, \tag{1.5}$$

we could write

$$\Psi = \psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2)\ldots\psi_N(\boldsymbol{r}_N), \tag{1.6}$$

and inserting this in the equation and dividing by Ψ would then give

$$\sum_{i=1}^{N} \frac{1}{\psi_i} \left[-\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i) \right] \psi_i = E. \tag{1.7}$$

Each term in the sum on the left-hand side depends upon the coordinates of one electron only, so that the equation can be separated into the N equations

$$-\frac{\hbar^2}{2m}\nabla_i^2\psi_i + V(\boldsymbol{r}_i)\psi_i = \epsilon_i\psi_i, \qquad i = 1, 2, \ldots N, \tag{1.8}$$

with

$$\sum_{i=1}^{N} \epsilon_i = E. \tag{1.9}$$

Although it is a solution of eq. (1.5) the function Ψ in (1.6) is not a correct wave function for two reasons. First, it should include the spin coordinates of the electrons. If spin–orbit interaction is neglected, we may include spin by writing, instead of $\psi_i(\boldsymbol{r})$,

$$\phi_i(\boldsymbol{x}) = \psi_i(\boldsymbol{r})\chi_i(\zeta), \tag{1.10}$$

where ζ is the spin coordinate, $\boldsymbol{x}$ standing for both $\boldsymbol{r}$ and ζ, and the spin function χ_i is either α or β, defined by

$$\begin{aligned} \alpha(1) &= 1, & \alpha(-1) &= 0, \\ \beta(1) &= 0, & \beta(-1) &= 1. \end{aligned}$$

The presence of the spin functions does not affect eqs. (1.8) and (1.9).

Secondly, we know (see WM, p. 115) that the wave function of any system of particles must be either symmetric or antisymmetric in the coordinates of the particles. Particles whose wave functions are symmetric are called *bosons*, and those whose wave functions are antisymmetric are called *fermions*. Photons, for example, are bosons, while electrons are fermions. Instead of a single product like (1.6) we must therefore take an antisymmetric sum of such products with the electronic coordinates permuted in all possible ways. The result, including spin, is the function

$$\Phi = \frac{1}{(N!)^{1/2}} \sum_P (-1)^p P\phi_1(\boldsymbol{x}_1)\ldots\phi_N(\boldsymbol{x}_N), \tag{1.11}$$

where P is a permutation operator upon the electronic coordinates, p is the number of interchanges in P, and the sum is over the $N!$ different permutations. This is just a compact way of writing the determinant

$$\Phi = \frac{1}{(N!)^{1/2}} \begin{vmatrix} \phi_1(\boldsymbol{x}_1) & \phi_1(\boldsymbol{x}_2) \ldots \phi_1(\boldsymbol{x}_N) \\ \phi_2(\boldsymbol{x}_1) & \phi_2(\boldsymbol{x}_2) \ldots \phi_2(\boldsymbol{x}_N) \\ \vdots & \vdots \qquad\quad \vdots \\ \vdots & \vdots \qquad\quad \vdots \\ \phi_N(\boldsymbol{x}_1) & \phi_N(\boldsymbol{x}_2) \ldots \phi_N(\boldsymbol{x}_N) \end{vmatrix}. \tag{1.12}$$

The factor $1/(N!)^{1/2}$ normalizes Φ, provided the one-electron functions are normalized, so that†

$$\int |\phi_i(\boldsymbol{x})|^2 \,\mathrm{d}\boldsymbol{x} = \int |\psi_i(\boldsymbol{r})|^2 \,\mathrm{d}\boldsymbol{r} = 1. \tag{1.13}$$

The functions ψ_i are, of course, orthogonal, since they are the eigenfunctions of eq. (1.8).

For the non-interacting system we have been considering, whose Hamiltonian is H_0, the determinantal wave function Φ gives the same energy E, eq. (1.9), as the single product function (1.6). The wave function of the ground state is just a determinant whose elements are the one-electron functions corresponding to the lowest energy levels ϵ_i (two functions with opposite spin factors to each orbital state).

When the interaction term H' is included in the Hamiltonian, so that we have the Schrödinger eq. (1.4), this equation can no longer be separated, and the wave function cannot be expressed as a single determinant of one-electron functions. However, it would seem possible to use the determinantal solutions of eq. (1.5) as the basis of a perturbation treatment of H' – in other words, to take the non-interacting system as the unperturbed system, with H' as a perturbation. We shall show in later chapters how this may be done, but it is by no means easy. The difficulty is that second-order perturbation theory will not work, owing to the fact that the second-order energy correction diverges. In order to obtain a finite result it is necessary to consider perturbation theory of infinite order.

Although it is possible to use the one-electron functions ϕ_i, whose orbital factors are solutions of eq. (1.8), in constructing the solutions of eq. (1.4), these may not be the best choice – it may, for example, be better to use the one-electron functions obtained by the Hartree or Hartree–Fock methods. Determinants formed from the latter functions would in general no longer be solutions of eq. (1.5), but solutions of (1.4) could still be expressed in terms of them. In the following sections we shall consider determinantal functions in general, without specifying the one-electron functions which are their elements, except to say that they form a complete, orthonormal set.

1.3. Determinantal Wave Functions

Let us then consider a complete, orthonormal set of one-electron functions $\phi_i(\boldsymbol{x})$, including spin. The orthonormality is expressed by

$$\int \phi_i^*(\boldsymbol{x})\phi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \delta_{ij}, \tag{1.14}$$

and the completeness by†

$$\sum_i \phi_i^*(\boldsymbol{x}')\phi_i(\boldsymbol{x}) = \delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{1.15}$$

---

† Integration with respect to $\boldsymbol{x}$ implies a sum over the two values ±1 of the spin variable ζ; thus

$$\int |\phi_i(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x} \equiv \sum_{\zeta=\pm1} \int |\psi_i(\boldsymbol{r})|^2 |\chi_i(\zeta)|^2\,\mathrm{d}\boldsymbol{r} = \int |\psi_i(\boldsymbol{r})|^2\,\mathrm{d}\boldsymbol{r}.$$

† $\delta(\boldsymbol{x}-\boldsymbol{x}')$ stands for $\delta(\boldsymbol{r}-\boldsymbol{r}')\delta_{\zeta\zeta'}$. The important property of a complete, orthogonal set of functions is that it is possible, within the interval of orthogonality, to expand almost any function in terms of them (see WM, Appendix 2). That this property leads to eq. (1.15) can easily be seen by expanding $\delta(\boldsymbol{x}-\boldsymbol{x}')$ in terms of the functions $\phi_i(\boldsymbol{x})$. Thus we write

$$\delta(\boldsymbol{x}-\boldsymbol{x}') = \sum_i B_i\phi_i(\boldsymbol{x}),$$

where the B_i are constants. Multiplying both sides by $\phi_j^*(\boldsymbol{x})$ and integrating over all $\boldsymbol{x}$, we find

$$\int \delta(\boldsymbol{x}-\boldsymbol{x}')\phi_j^*(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_i B_i \int \phi_j^*(\boldsymbol{x})\phi_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = B_j,$$

from eq. (1.14). The left-hand side is just $\phi_j^*(\boldsymbol{x}')$, so that

$$B_i = \phi_i^*(\boldsymbol{x}'),$$

and eq. (1.15) follows. This is also sometimes called the *closure relation* for the functions ϕ_i.

Each function is the product of an orbital factor and a spin factor, as in eq. (1.10).

We now construct N-electron determinantal functions from these one-electron functions, of the type

$$\Phi_a(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N) = \frac{1}{(N!)^{1/2}} \sum_P (-1)^p P\phi_{a_1}(\boldsymbol{x}_1)\phi_{a_2}(\boldsymbol{x}_2)\ldots\phi_{a_N}(\boldsymbol{x}_N), \qquad (1.16)$$

the subscripts a_1, a_2, etc., denoting N different integers – that is, N different functions chosen from the infinite set of the $\phi_i(\boldsymbol{x})$. We shall write another such determinantal function as

$$\Phi_b(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N) = \frac{1}{(N!)^{1/2}} \sum_P (-1)^p P\phi_{b_1}(\boldsymbol{x}_1)\phi_{b_2}(\boldsymbol{x}_2)\ldots\phi_{b_N}(\boldsymbol{x}_N), \qquad (1.17)$$

where again b_1, b_2, etc., denote different integers.

We shall first prove the following theorem†:

If F is a symmetric operator and Φ_a and Φ_b are normalized determinantal functions, then

$$\int \Phi_b^* F \Phi_a \, \mathrm{d}\tau' = (N!)^{1/2} \int \Phi_b^* F \phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'. \qquad (1.18)$$

A symmetric operator is one which is symmetric in the electronic coordinates, so that $PF = F$. We have

$$\begin{aligned}
\int \Phi_b^* F \Phi_a \, \mathrm{d}\tau' &= \frac{1}{(N!)^{1/2}} \int \Phi_b^* F \sum_P (-1)^p P\phi_{a_1}(\boldsymbol{x}_1)\phi_{a_2}(\boldsymbol{x}_2)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau' \\
&= \frac{1}{(N!)^{1/2}} \sum_P (-1)^p P \int (P^{-1}\Phi_b^*) F \Phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau' \\
&= \frac{1}{(N!)^{1/2}} \sum_P P \int \Phi_b^* F \phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau' \\
&= (N!)^{1/2} \int \Phi_b^* F \phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'. \qquad (1.19)
\end{aligned}$$

---

† This is an extension of the theorem proved on p. 119 of WM. Note that the prime on the volume element $\mathrm{d}\tau'$ indicates that the multiple integral includes a sum over the two values of each of the spin variables.

In the second line we have transferred the summation over P, and the operator P, to the left of the integral sign, and this necessitates annulling the effect of P upon Φ_b^* by introducing the operator P^{-1} (which is to operate upon Φ_b^* only). The fact that P now operates upon F is of no consequence, since $PF = F$. The number of interchanges in P^{-1} is just p, the same as in P, since the two operators affect the same interchanges in reverse order. Hence

$$P^{-1}\Phi_b^* = (-1)^p \Phi_b^*,$$

for P^{-1} merely produces p interchanges of pairs of columns of the determinant. Since there is a factor $(-1)^p$ already present, and $(-1)^{2p} = 1$, we obtain the third line of eq. (1.19). Finally, we note that the value of a definite integral is independent of the symbols used for the variables of integration. Hence the integral in line three has the same value for every permutation P, and there are $N!$ permutations in all. So we arrive at the required expression.

It may be verified immediately that the determinantal functions are normalized, for

$$\begin{aligned}
\int \Phi_a^* \Phi_a \, \mathrm{d}\tau' &= (N!)^{1/2} \int \Phi_a^* \phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau' \\
&= \int \Big[\sum_P (-1)^p P\phi_{a_1}^*(\boldsymbol{x}_1)\ldots\phi_{a_N}^*(\boldsymbol{x}_N)\Big]\phi_{a_1}(\boldsymbol{x}_1)\ldots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau' \\
&= \int |\phi_{a_1}(\boldsymbol{x}_1)|^2 |\phi_{a_2}(\boldsymbol{x}_2)|^2 \ldots |\phi_{a_N}(\boldsymbol{x}_N)|^2 \,\mathrm{d}\tau' \\
&= \int |\phi_{a_1}(\boldsymbol{x}_1)|^2 \,\mathrm{d}\boldsymbol{x}_1 \ldots \int |\phi_{a_N}(\boldsymbol{x}_N)|^2 \,\mathrm{d}\boldsymbol{x}_N = 1, \qquad (1.20)
\end{aligned}$$

all the other terms of Φ_a^* giving zero contribution owing to the orthogonality of the ϕ_i.

We may also verify that the determinantal functions are orthogonal. Suppose that Φ_a and Φ_b are *different* determinantal functions, which means that the sets of subscripts a_i and b_i must differ in at least one integer. Let us assume that

$$\left.\begin{aligned} a_i &= b_i, \qquad i \neq j, \\ a_j &\neq b_j. \end{aligned}\right\} \qquad (1.21)$$

Then

$$\begin{aligned}\int \Phi_a^* \Phi_b \, \mathrm{d}\tau' &\\ &= \int \Big[\sum_P (-1)^p P \phi_{a_1}^*(\boldsymbol{x}_1) \ldots \phi_{a_N}^*(\boldsymbol{x}_N) \Big] \phi_{b_1}(\boldsymbol{x}_1) \ldots \phi_{b_N}(\boldsymbol{x}_N) \, \mathrm{d}\tau' \\ &= \int |\phi_{a_1}(\boldsymbol{x}_1)|^2 \, \mathrm{d}\boldsymbol{x}_1 \ldots \int \phi_{a_j}^*(\boldsymbol{x}_j) \phi_{b_j}(\boldsymbol{x}_j) \, \mathrm{d}\boldsymbol{x}_j \ldots \int |\phi_{a_N}(\boldsymbol{x}_N)|^2 \, \mathrm{d}\boldsymbol{x}_N \\ &= 0, \end{aligned} \tag{1.22}$$

owing to the orthogonality of the functions ϕ_{a_j} and ϕ_{b_j}. If Φ_a and Φ_b differ in more than one function, that is, if the sets a_i and b_i differ in more than one integer, clearly the same result will follow.

The determinantal functions thus form an orthonormal set, and we shall assume that this set is complete, so that any N-electron wave function Ψ can be expanded in an infinite series of the type

$$\Psi = \sum_a B_a \Phi_a, \tag{1.23}$$

the coefficients B_a being suitably chosen constants. This function is correctly antisymmetric, because interchange of the coordinates of two electrons changes the sign of every determinant.

1.4. Matrix Elements

The problem is to find the eigenfunctions and eigenvalues of the Schrödinger equation

$$H\Psi = E\Psi. \tag{1.24}$$

Let us assume an expansion of the form (1.23) for Ψ and substitute in this equation. We obtain

$$\sum_a B_a H \Phi_a = E \sum_a B_a \Phi_a. \tag{1.25}$$

Multiplication by Φ_b^* and integration over the configuration space of the system now give

$$\sum_a B_a \int \Phi_b^* H \Phi_a \, \mathrm{d}\tau' = E \sum_a B_a \int \Phi_b^* \Phi_a \, \mathrm{d}\tau'. \tag{1.26}$$

If we write

$$H_{ba} = \int \Phi_b^* H \Phi_a \, \mathrm{d}\tau', \tag{1.27}$$

which is called the *matrix element* of H between the functions Φ_b and Φ_a, and remember that

$$\int \Phi_b^* \Phi_a \, \mathrm{d}\tau' = \delta_{ba}, \tag{1.28}$$

owing to the orthonormality of the Φ_a, eq. (1.26) becomes

$$\sum_a B_a (H_{ba} - E\delta_{ba}) = 0. \tag{1.29}$$

There is an infinite number of such equations, corresponding to all the functions Φ_b, and each equation has an infinite number of terms. We thus have an infinite set of simultaneous linear equations in the B_a, and for consistency the determinant of the coefficients of the B_a must vanish, that is,

$$|H_{ba} - E\delta_{ba}| = 0. \tag{1.30}$$

The determinant is of infinite order, E appearing only in the elements of the leading diagonal. The matrix elements H_{ab} are, of course, simply numbers, and, once these have been calculated, the determinantal equation can in principle be solved to give an infinite number of values of E, which are the energy levels of the system.

If the functions Φ_a are the eigenfunctions of a Hamiltonian which differs from H by only a small perturbing term, then the approximate evaluation of the roots of eq. (1.30) is very much simpler, as shown in WM, §3.8. However, it is important to notice that we have made no such assumption here – eq. (1.30) is correct, and in principle will yield the correct energy levels for *any* complete, orthonormal set of N-electron determinantal functions Φ_a. In other words, the energy levels of the system are completely specified by the matrix elements of H with respect to any such set of functions. These matrix elements thus constitute a complete representation of H, and any operator which has the same matrix elements as H is entirely equivalent to H.

We shall need to use this fact in the following chapter, but meanwhile it is necessary to obtain more definite information about the matrix elements of H.

For the N-electron system we have been considering H is the sum of H_0, eq. (1.2), and H', eq. (1.3), both of which are symmetric operators. H_0 is the sum of terms, each of which depends upon the coordinates of one electron only, and H' is the sum of terms, each of which depends upon the coordinates of two electrons only. Let us for convenience write†

$$H_0 = \sum_{i=1}^{N} f(\boldsymbol{x}_i), \tag{1.31}$$

where

$$f(\boldsymbol{x}_i) = -\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i), \tag{1.32}$$

and

$$H' = \tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N} v(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{1.33}$$

where

$$v(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{1.34}$$

Also, to obviate a confusion of subscripts, we shall throughout the remainder of the book consistently use the following shorthand notation for matrix elements. For the matrix element of H, or any other many-electron operator, between two many-electron functions we shall write‡, for example,

$$\langle \Phi_b | H | \Phi_a \rangle = \int \Phi_b^* H \Phi_a \, \mathrm{d}\tau', \tag{1.35}$$

which is what we have previously denoted by H_{ba}. If the operator is unity, we shall simplify the notation further and write

$$\langle \Phi_b | \Phi_a \rangle = \int \Phi_b^* \Phi_a \, \mathrm{d}\tau'. \tag{1.36}$$

---

† The results we shall obtain here and in the following chapter will apply quite generally to any system of N fermions whose Hamiltonian consists of the sum of two symmetric operators like (1.31) and (1.33), even when f and v have not the special forms, independent of spin, given in (1.32) and (1.34). This is why we have used $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ rather than $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$.

‡ Although the use of angular brackets originated with Dirac, it must not be supposed that we are going to make use of his very abstract formulation of quantum mechanics. In fact, for the benefit of the unsophisticated reader, and at the risk of a little clumsiness, we shall present a wave-mechanical treatment throughout, so that the bracketed symbols may be thought of as merely a convenient notation for the integrals which they represent.

On the other hand, for the matrix elements of f and v with respect to a chosen set of *one-electron* functions $\phi_i(\boldsymbol{x})$ we shall consistently use the notation

$$\langle i | f | j \rangle = \int \phi_i^*(\boldsymbol{x}) f(\boldsymbol{x}) \phi_j(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \tag{1.37}$$

and

$$\langle ij | v | kl \rangle = \iint \phi_i^*(\boldsymbol{x}_1)\phi_j^*(\boldsymbol{x}_2) v(\boldsymbol{x}_1, \boldsymbol{x}_2)\phi_k(\boldsymbol{x}_1)\phi_l(\boldsymbol{x}_2) \, \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}\boldsymbol{x}_2. \tag{1.38}$$

1.4.1. *Matrix elements of H_0*

We have, in general,

$$\begin{aligned}\langle \Phi_b | H_0 | \Phi_a \rangle &= \int \Phi_b^* H_0 \Phi_a \, \mathrm{d}\tau' \\ &= (N!)^{1/2} \int \Phi_b^* H_0 \phi_{a_1}(\boldsymbol{x}_1)\phi_{a_2}(\boldsymbol{x}_2) \ldots \phi_{a_N}(\boldsymbol{x}_N) \, \mathrm{d}\tau' \\ &= (N!)^{1/2} \sum_{i=1}^{N} \int \Phi_b^* f(\boldsymbol{x}_i) \phi_{a_1}(\boldsymbol{x}_1) \ldots \phi_{a_N}(\boldsymbol{x}_N) \, \mathrm{d}\tau'. \end{aligned} \tag{1.39}$$

Now let us apply this to the following particular cases:

(1) Suppose $\Phi_a = \Phi_b$, that is, $b_j = a_j$ for all j. Then

$$\begin{aligned}(N!)^{1/2} &\int \Phi_b^* f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1) \ldots \phi_{a_N}(\boldsymbol{x}_N) \, \mathrm{d}\tau' \\ &= \int \Big[\sum_P (-1)^p P \phi_{a_1}^*(\boldsymbol{x}_1) \ldots \phi_{a_N}^*(\boldsymbol{x}_N)\Big] f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1) \ldots \phi_{a_N}(\boldsymbol{x}_N) \, \mathrm{d}\tau' \\ &= \int \phi_{a_i}^*(\boldsymbol{x}_i) f(\boldsymbol{x}_i) \phi_{a_i}(\boldsymbol{x}_i) \, \mathrm{d}\boldsymbol{x}_i, \end{aligned} \tag{1.40}$$

owing to the orthonormality of the ϕ_{a_i}.

Thus

$$\langle \Phi_a | H_0 | \Phi_a \rangle = \sum_{i=1}^{N} \langle a_i | f | a_i \rangle. \tag{1.41}$$

(2) Next, suppose that Φ_a and Φ_b differ in only one function. Let us say $a_j = b_j$ for $j \neq k$, but $a_k \neq b_k$.

Then

$$(N!)^{1/2}\int \Phi_b^* f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$\begin{aligned}&=\int\Big[\sum_P(-1)^p P\phi_{a_1}^*(\boldsymbol{x}_1)\dots\phi_{b_k}^*(\boldsymbol{x}_k)\dots\phi_{a_N}^*(\boldsymbol{x}_N)\Big]\\&\qquad\times f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_k}(\boldsymbol{x}_k)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'\\&=\begin{cases}0, & \text{if}\quad i\neq k,\\ \displaystyle\int\phi_{b_k}^*(\boldsymbol{x}_k)f(\boldsymbol{x}_k)\phi_{a_k}(\boldsymbol{x}_k)\,\mathrm{d}\boldsymbol{x}_k, & \text{if}\quad i=k.\end{cases}\end{aligned}\tag{1.42}$$

Thus

$$\langle\Phi_b|H_0|\Phi_a\rangle=\langle b_k|f|a_k\rangle.\tag{1.43}$$

(3) Finally, suppose that Φ_a and Φ_b differ in two functions. Let us say $a_j=b_j$ for $j\neq k,l$, but $a_k\neq b_k$ and $a_l\neq b_l$ (also $a_k\neq b_l$ and $a_l\neq b_k$). Then

$$(N!)^{1/2}\int \Phi_b^* f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$\begin{aligned}&=\int\Big[\sum_P(-1)^p P\phi_{a_1}^*(\boldsymbol{x}_1)\dots\phi_{b_k}^*(\boldsymbol{x}_k)\dots\phi_{b_l}^*(\boldsymbol{x}_l)\dots\phi_{a_N}^*(\boldsymbol{x}_N)\Big]\\&\qquad\times f(\boldsymbol{x}_i)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_k}(\boldsymbol{x}_k)\dots\phi_{a_l}(\boldsymbol{x}_l)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'\\&=0,\quad\text{for all } i.\end{aligned}\tag{1.44}$$

Thus

$$\langle\Phi_b|H_0|\Phi_a\rangle=0,\tag{1.45}$$

in this case. Clearly, the same result follows if Φ_b and Φ_a differ in more than two functions.

To summarize: H_0 can have non-zero matrix elements only between two determinantal functions which are either the same or differ in a single one-electron function only, and these matrix elements are given by eqs. (1.41) and (1.43), respectively†.

---

† In the latter case the matrix element also vanishes unless the two different one-electron functions have the same spin factor, since

$$\begin{aligned}\langle b_k|f|a_k\rangle&=\int\phi_{b_k}^*(\boldsymbol{x})f(\boldsymbol{x})\phi_{a_k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\&=\sum_{\zeta=\pm1}\chi_{b_k}(\zeta)\chi_{a_k}(\zeta)\int\psi_{b_k}^*(\boldsymbol{r})f(\boldsymbol{x})\psi_{a_k}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\\&=0\qquad\text{if}\qquad\chi_{b_k}\neq\chi_{a_k}.\end{aligned}$$

1.4.2. *Matrix elements of H′*

We have, in general,

$$\begin{aligned}\langle\Phi_b|H'|\Phi_a\rangle&=\int\Phi_b^*H'\Phi_a\,\mathrm{d}\tau'\\&=(N!)^{1/2}\int\Phi_b^*H'\phi_{a_1}(\boldsymbol{x}_1)\phi_{a_2}(\boldsymbol{x}_2)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'\\&=\tfrac{1}{2}(N!)^{1/2}\sum_{i\neq j}^{N}\sum^{N}\int\Phi_b^*v(\boldsymbol{x}_i,\boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau',\end{aligned}\tag{1.46}$$

and we shall again apply this to the following cases:

(1) Suppose $\Phi_a=\Phi_b$. Then

$$(N!)^{1/2}\int\Phi_b^*v(\boldsymbol{x}_i,\boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$\begin{aligned}&=\int\Big[\sum_P(-1)^pP\phi_{a_1}^*(\boldsymbol{x}_1)\dots\phi_{a_N}^*(\boldsymbol{x}_N)\Big]\\&\qquad\times v(\boldsymbol{x}_i,\boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'\\&=\iint[\phi_{a_i}^*(\boldsymbol{x}_i)\phi_{a_j}^*(\boldsymbol{x}_j)-\phi_{a_j}^*(\boldsymbol{x}_i)\phi_{a_i}^*(\boldsymbol{x}_j)]\\&\qquad\times v(\boldsymbol{x}_i,\boldsymbol{x}_j)\phi_{a_i}(\boldsymbol{x}_i)\phi_{a_j}(\boldsymbol{x}_j)\,\mathrm{d}\boldsymbol{x}_i\,\mathrm{d}\boldsymbol{x}_j\\&=\langle a_ia_j|v|a_ia_j\rangle-\langle a_ja_i|v|a_ia_j\rangle.\end{aligned}\tag{1.47}$$

The second term comes from that permutation which interchanges a_j and a_i only. For a single interchange, $p=1$, which accounts for the minus sign.

Thus

$$\langle\Phi_a|H'|\Phi_a\rangle=\tfrac{1}{2}\sum_i^N\sum_j^N[\langle a_ia_j|v|a_ia_j\rangle-\langle a_ja_i|v|a_ia_j\rangle].\tag{1.48}$$

Here it is not necessary to specify $i\neq j$, as in eq. (1.46), because the expression in square brackets vanishes when $i=j$.

(2) Suppose that Φ_a and Φ_b differ in only one function. Let us say $a_g=b_g$ for $g\neq k$, but $a_k\neq b_k$. Then

$$(N!)^{1/2}\int\Phi_b^*v(\boldsymbol{x}_i,\boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\dots\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$=\int\Big[\sum_P(-1)^pP\phi_{a_1}^*(\boldsymbol{x}_1)\dots\phi_{b_k}^*(\boldsymbol{x}_k)\dots\phi_{a_N}^*(\boldsymbol{x}_N)\Big]\times$$

$$\times\, v(\boldsymbol{x}_i, \boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\,\ldots\,\phi_{a_k}(\boldsymbol{x}_k)\,\ldots\,\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$= \begin{cases} \langle b_k a_j|v|a_k a_j\rangle - \langle a_j b_k|v|a_k a_j\rangle, & \text{if} \quad i=k, \\ \langle a_i b_k|v|a_i a_k\rangle - \langle b_k a_i|v|a_i a_k\rangle, & \text{if} \quad j=k, \\ 0, \quad \text{otherwise.} \end{cases} \tag{1.49}$$

Thus, from eq. (1.46),

$$\langle\Phi_b|H'|\Phi_a\rangle = \tfrac{1}{2}\sum_j^N \left[\langle b_k a_j|v|a_k a_j\rangle - \langle a_j b_k|v|a_k a_j\rangle\right]$$

$$+\tfrac{1}{2}\sum_i^N \left[\langle a_i b_k|v|a_i a_k\rangle - \langle b_k a_i|v|a_i a_k\rangle\right]$$

$$= \sum_i^N \left[\langle a_i b_k|v|a_i a_k\rangle - \langle b_k a_i|v|a_i a_k\rangle\right]. \tag{1.50}$$

Here we have used the fact that

$$\begin{aligned}\langle b_k a_j|v|a_k a_j\rangle &= \iint \phi_{b_k}^*(\boldsymbol{x}_1)\phi_{a_j}^*(\boldsymbol{x}_2)v(\boldsymbol{x}_1, \boldsymbol{x}_2)\phi_{a_k}(\boldsymbol{x}_1)\phi_{a_j}(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &= \iint \phi_{b_k}^*(\boldsymbol{x}_2)\phi_{a_j}^*(\boldsymbol{x}_1)v(\boldsymbol{x}_2, \boldsymbol{x}_1)\phi_{a_k}(\boldsymbol{x}_2)\phi_{a_j}(\boldsymbol{x}_1)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &= \langle a_j b_k|v|a_j a_k\rangle,\end{aligned} \tag{1.51}$$

since

$$v(\boldsymbol{x}_1, \boldsymbol{x}_2) = v(\boldsymbol{x}_2, \boldsymbol{x}_1), \tag{1.52}$$

so that the sums over j and i are the same. Again, it is unnecessary to specify $i \neq k$ because the term with $i = k$ vanishes.

(3) Suppose that Φ_a and Φ_b differ in two functions. Let us say $a_g = b_g$ for $g \neq k,l$, but $a_k \neq b_k$ and $a_l \neq b_l$ (also $a_k \neq b_l$ and $a_l \neq b_k$). Then

$$(N!)^{1/2}\int \Phi_b^* v(\boldsymbol{x}_i, \boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\,\ldots\,\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$= \int\left[\sum_P (-1)^p P\phi_{a_1}^*(\boldsymbol{x}_1)\,\ldots\,\phi_{b_k}^*(\boldsymbol{x}_k)\,\ldots\,\phi_{b_l}^*(\boldsymbol{x}_l)\,\ldots\,\phi_{a_N}^*(\boldsymbol{x}_N)\right]$$

$$\times\, v(\boldsymbol{x}_i, \boldsymbol{x}_j)\phi_{a_1}(\boldsymbol{x}_1)\,\ldots\,\phi_{a_k}(\boldsymbol{x}_k)\,\ldots\,\phi_{a_l}(\boldsymbol{x}_l)\,\ldots\,\phi_{a_N}(\boldsymbol{x}_N)\,\mathrm{d}\tau'$$

$$= \begin{cases} \langle b_k b_l|v|a_k a_l\rangle - \langle b_l b_k|v|a_k a_l\rangle, & \begin{array}{l}\text{if } i=k,\, j=l \\ \text{or } i=l,\, j=k,\end{array} \\ 0, \quad \text{otherwise.} \end{cases} \tag{1.53}$$

Thus

$$\langle\Phi_b|H'|\Phi_a\rangle = \langle b_k b_l|v|a_k a_l\rangle - \langle b_l b_k|v|a_k a_l\rangle. \tag{1.54}$$

(4) Finally, if Φ_a and Φ_b differ in more than two functions, the matrix element $\langle\Phi_b|H'|\Phi_a\rangle$ is zero, since every term on the right-hand side of eq. (1.46) then contains a factor which vanishes owing to the orthogonality of the ϕ_{a_i}.

To summarize: H' can have non-zero matrix elements only between two determinantal functions which are the same or which differ in either one or two one-electron functions only, and these matrix elements are given by eqs. (1.48), (1.50) and (1.54).

1.5. Perturbation Theory

It will prove convenient for future reference to end this chapter with an account of the Rayleigh–Schrödinger perturbation theory, which is the elementary time-independent perturbation theory described in most text-books of quantum mechanics†. The method we shall use, however, is different from that generally encountered in introductory texts, and has the advantage that it provides a *formally* simple expression for the perturbed wave function and energy to infinite order.

As before, let us take the Hamiltonian to be

$$H = H_0 + H', \tag{1.55}$$

but now we shall treat H' as *any* perturbation, not necessarily the Coulomb interaction (indeed, it is not necessary for the system to contain many electrons). We suppose that Φ_n is an eigenfunction of H_0 corresponding to the eigenvalue E_n, that is,

$$H_0\Phi_n = E_n\Phi_n, \tag{1.56}$$

and wish to consider the effect of the perturbation on a particular *non-degenerate* state Φ_0, where

$$H_0\Phi_0 = E_0\Phi_0. \tag{1.57}$$

---

† See, for example, WM, §3.8. The principal purpose of the present section is to derive the expressions (1.78) and (1.87) for the first- and second-order energy corrections. Those readers who are already familiar with the latter may prefer to omit this section. However, the derivation used here has a certain intrinsic interest and reference will be made to parts of it in later chapters.

We denote by Ψ_0 the state into which Φ_0 changes under the action of the perturbation, so that Ψ_0 is an eigenfunction of H, corresponding to the eigenvalue E, say: that is,

$$H\Psi_0 = E\Psi_0. \tag{1.58}$$

In our future work Φ_0 and Ψ_0 will denote the ground states of the unperturbed and perturbed systems, respectively, but the present treatment is quite general.

We thus have

$$H'\Psi_0 = (H-H_0)\Psi_0 = (E-H_0)\Psi_0, \tag{1.59}$$

so that

$$\langle\Phi_0|H'|\Psi_0\rangle = E\langle\Phi_0|\Psi_0\rangle - \langle\Phi_0|H_0|\Psi_0\rangle. \tag{1.60}$$

Now, since H_0 is Hermitian, it follows from Appendix I, eq. (I.24), that

$$\langle\Phi_0|H_0|\Psi_0\rangle = E_0\langle\Phi_0|\Psi_0\rangle, \tag{1.61}$$

and substitution in (1.60) gives

$$E-E_0 = \frac{\langle\Phi_0|H'|\Psi_0\rangle}{\langle\Phi_0|\Psi_0\rangle}. \tag{1.62}$$

This expression is, of course, exact and independent of any particular perturbation method. It cannot be used immediately, however, because the right-hand side contains the perturbed wave function, which is unknown.

We now define a so-called *projection operator* R for the state Φ_0 by the equation

$$R\Psi = \Psi - \Phi_0\langle\Phi_0|\Psi\rangle, \tag{1.63}$$

where Ψ is any function of the same variables as Φ_0. This operator removes the Φ_0 component of the function Ψ. Thus, if

$$\Psi = \sum_{n=0}^{\infty} B_n\Phi_n \tag{1.64}$$

is the expansion of Ψ in terms of the functions Φ_n, assumed orthonormal, we find

$$\begin{aligned} R\Psi &= \Psi - \Phi_0 \sum_{n=0}^{\infty} B_n\langle\Phi_0|\Phi_n\rangle \\ &= \Psi - B_0\Phi_0. \end{aligned} \tag{1.65}$$

In particular,

$$R\Phi_0 = 0. \tag{1.66}$$

If we substitute $R\Psi$ for Ψ in eq. (1.63), we obtain

$$\begin{aligned} R^2\Psi &= R\Psi - \Phi_0\langle\Phi_0|R\Psi\rangle \\ &= R\Psi - \Phi_0[\langle\Phi_0|\Psi\rangle - \langle\Phi_0|\Phi_0\rangle\langle\Phi_0|\Psi\rangle] \\ &= R\Psi, \end{aligned} \tag{1.67}$$

owing to which property R is said to be *idempotent*. This also follows immediately from eqs. (1.65) and (1.66).

We shall in future write

$$\langle\Phi_0|\Psi_0\rangle = C, \tag{1.68}$$

a constant depending upon the normalization of Ψ_0. Eq. (1.63) gives

$$\begin{aligned} R(E_0-H_0)\Psi_0 &= (E_0-H_0)\Psi_0 - \Phi_0\langle\Phi_0|E_0-H_0|\Psi_0\rangle \\ &= (E_0-H_0)\Psi_0 - CE_0\Phi_0 + \Phi_0\langle\Phi_0|H_0|\Psi_0\rangle \\ &= (E_0-H_0)\Psi_0, \end{aligned} \tag{1.69}$$

using (1.61), and

$$\begin{aligned} (E_0-H_0)R\Psi_0 &= (E_0-H_0)(\Psi_0 - C\Phi_0) \\ &= (E_0-H_0)\Psi_0. \end{aligned} \tag{1.70}$$

In other words, R commutes with E_0-H_0.

Now,

$$(E_0-H_0)\Psi_0 = (E_0-H+H')\Psi_0 = (E_0-E+H')\Psi_0, \tag{1.71}$$

so that

$$(E_0-H_0)R\Psi_0 = R(E_0-H_0)\Psi_0 = R(E_0-E+H')\Psi_0, \tag{1.72}$$

and, therefore[†],

$$R\Psi_0 = \frac{R}{E_0-H_0}(E_0-E+H')\Psi_0 = \Psi_0 - \Phi_0\langle\Phi_0|\Psi_0\rangle, \tag{1.73}$$

again using (1.63). The perturbed wave function Ψ_0 thus satisfies the equation

$$\Psi_0 = C\Phi_0 + \frac{R}{E_0-H_0}(E_0-E+H')\Psi_0, \tag{1.74}$$

[†] H_0 is a differential operator, so that division by E_0-H_0 requires some justification. The

which may be iterated to give

$$\begin{aligned}\Psi_0 &= C\Phi_0+\frac{R}{E_0-H_0}(E_0-E+H')\left[C\Phi_0+\frac{R}{E_0-H_0}(E_0-E+H')\Psi_0\right]\\ &= C\Phi_0+\frac{CR}{E_0-H_0}(E_0-E+H')\Phi_0\\ &\quad+\left[\frac{R}{E_0-H_0}(E_0-E+H')\right]^2(C\Phi_0+\dots)\\ &= C\sum_{n=0}^{\infty}\left[\frac{R}{E_0-H_0}(E_0-E+H')\right]^n\Phi_0. \qquad (1.75)\end{aligned}$$

The perturbed energy can be obtained by substituting this expression in eq. (1.62), thus:

$$E-E_0=\sum_{n=0}^{\infty}\left\langle\Phi_0\middle|H'\left[\frac{R}{E_0-H_0}(E_0-E+H')\right]^n\middle|\Phi_0\right\rangle. \qquad (1.76)$$

It will be observed that the right-hand side of this equation also contains E, but this is eliminated when the terms are expanded.

We shall write

$$\Delta E=E-E_0=\Delta E^{(1)}+\Delta E^{(2)}+\Delta E^{(3)}+\dots, \qquad (1.77)$$

where $\Delta E^{(m)}$, the *m-th-order energy correction*, contains the m-th power of the perturbation H'. The first-order correction is the term of (1.76) with $n=0$, that is,

$$\Delta E^{(1)}=\langle\Phi_0|H'|\Phi_0\rangle. \qquad (1.78)$$

The second-order correction is

$$\Delta E^{(2)}=\left\langle\Phi_0\middle|H'\frac{R}{E_0-H_0}(E_0-E+H')\middle|\Phi_0\right\rangle. \qquad (1.79)$$

From eq. (1.66),

$$R(E_0-E+H')\Phi_0=RH'\Phi_0, \qquad (1.80)$$

so that (1.79) can equally well be written

$$\Delta E^{(2)}=\left\langle\Phi_0\middle|H'\frac{R}{E_0-H_0}H'\middle|\Phi_0\right\rangle. \qquad (1.81)$$

We may expand $H'\Phi_0$ in terms of the Φ_n, thus:

$$H'\Phi_0=\sum_{n=0}^{\infty}B_n\Phi_n. \qquad (1.82)$$

The coefficients B_n are obtained by multiplying both sides of this equation by Φ_m^* and integrating over the configuration space of the system. This gives

$$\langle\Phi_m|H'|\Phi_0\rangle=\sum_{n=0}^{\infty}B_n\langle\Phi_m|\Phi_n\rangle=B_m, \qquad (1.83)$$

and

$$H'\Phi_0=\sum_{n=0}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle\Phi_n, \qquad (1.84)$$

so that

$$RH'\Phi_0=\sum_{n=1}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle\Phi_n, \qquad (1.85)$$

which simply removes the Φ_0 term from (1.84). It then follows from eq. (1.81) that

$$\begin{aligned}\Delta E^{(2)}&=\left\langle\Phi_0\middle|H'\frac{1}{E_0-H_0}\sum_{n=1}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle\middle|\Phi_n\right\rangle\\ &=\sum_{n=1}^{\infty}\frac{\langle\Phi_0|H'|\Phi_n\rangle\langle\Phi_n|H'|\Phi_0\rangle}{E_0-E_n}, \qquad (1.86)\end{aligned}$$

or, on the assumption that H' is Hermitian,

$$\Delta E^{(2)}=\sum_{n=1}^{\infty}\frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{E_0-E_n}. \qquad (1.87)$$

---

operator $(E_0-H_0)^{-1}$ is defined as the inverse of E_0-H_0 by the equations

$$(E_0-H_0)(E_0-H_0)^{-1}\Psi=(E_0-H_0)^{-1}(E_0-H_0)\Psi=\Psi.$$

In practical application it may be equated to its binomial expansion: thus

$$\frac{1}{E_0-H_0}=\frac{1}{E_0}\left(1+\frac{H_0}{E_0}+\frac{H_0^2}{E_0^2}+\dots\right),$$

from which it follows that

$$\frac{1}{E_0-H_0}\Phi_n=\frac{1}{E_0-E_n}\Phi_n.$$

Clearly this breaks down when $n=0$, the coefficient on the right then being infinite.

Since E_0-H_0 commutes with R, we may write

$$\frac{R}{E_0-H_0}\Psi=\frac{1}{E_0-H_0}R\Psi=R\frac{1}{E_0-H_0}\Psi.$$

In other words, it is immaterial whether we operate with R or $(E_0-H_0)^{-1}$ first. However, owing to the above-mentioned difficulty regarding the effect of $(E_0-H_0)^{-1}$ on Φ_0, it is advisable to operate with R first, since this gets rid of the Φ_0 component.

This depends only upon H' and the *unperturbed* energy levels and wave functions.

The higher-order energy corrections may be found in the same way, but these become very complicated and Rayleigh–Schrödinger perturbation theory is rarely taken beyond the second order. None the less, we shall find that second-order perturbation theory fails for an electron gas when the perturbation H' is the Coulomb interaction of the electrons. It is then necessary to consider perturbation theory of infinite order, that is, the whole of the perturbation series (1.77). The simple approach described above is no longer adequate and more powerful perturbation techniques have to be used, as will be discussed in the following chapters.

CHAPTER 2

THE OCCUPATION NUMBER REPRESENTATION (SECOND QUANTIZATION)

2.1. Creation and Destruction Operators

In the previous chapter we obtained as much information as possible, without knowing the actual form of the one-electron functions ϕ_i, about the matrix elements of H with respect to a complete, orthonormal set of N-electron determinantal functions Φ_a. In the present chapter we shall use these results in order to establish a new mathematical formalism which will prove convenient in later developments of the theory. This formalism is generally called *second quantization*, but the name suggests that some new quantum-mechanical principle is involved, which is certainly not the case in the present context, and we should therefore prefer to call it the *occupation number representation* or something similar.

A determinantal function can be specified by the one-electron functions ϕ_i which it contains. The only ambiguity is that of sign, and this can be removed by ascribing a definite *order* to the functions ϕ_i. Let us therefore suppose that in any determinantal function Φ_a, defined as in eq. (1.16), the one-electron functions ϕ_{a_i} are always written in the same order, thus:

$$a_1 < a_2 < \ldots a_N.$$

The order may be that of increasing one-electron energy (where this is meaningful), but this is by no means necessary – so long as we choose a fixed order for the ϕ_i, *any* order will do. Then we may specify Φ_a completely by giving the subscripts of the functions it contains, thus:

$$\Phi_{a_1a_2\ldots a_N}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N) = \frac{1}{(N!)^{1/2}} \sum_P (-1)^p P \phi_{a_1}(\boldsymbol{x}_1)\phi_{a_2}(\boldsymbol{x}_2) \ldots \phi_{a_N}(\boldsymbol{x}_N). \tag{2.1}$$

We now introduce the *destruction operator* (or *annihilation operator*) c_{a_k} which removes a state ϕ_{a_k} from a determinantal function containing this state; that is to say, it converts an N-electron function containing ϕ_{a_k} into an $(N-1)$-electron function not containing ϕ_{a_k}. The normalizing factor is also changed, so that the new function is a normalized determinantal function appropriate to a system of $N-1$ electrons. The sign has to be chosen according to the position of ϕ_{a_k} in Φ_a, as described below.

The formal definition of c_{a_k} is as follows†:

if Φ^N_a contains ϕ_{a_k}, then

$$\begin{aligned} c_{a_k}\Phi^N_{a_1a_2\ldots a_k\ldots a_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N) &= \pm\Phi^{N-1}_{a_1a_2\ldots a_{k-1}a_{k+1}\ldots a_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_{N-1}) \\ &= \pm\frac{1}{[(N-1)!]^{1/2}}\sum_P (-1)^p P\phi_{a_1}(\mathbf{x}_1)\ldots\phi_{a_{k-1}}(\mathbf{x}_{k-1}) \\ &\quad\times\phi_{a_{k+1}}(\mathbf{x}_k)\ldots\phi_N(\mathbf{x}_{N-1}), \end{aligned} \tag{2.2}$$

the sign being *positive* if k is *odd* (that is, if an even number of functions precede ϕ_{a_k} in Φ^N) and *negative* if k is *even*.

On the other hand, if Φ^N_a does not contain a one-electron function ϕ_l, say (that is, if l is a number not appearing in the set $a_1, a_2, \ldots a_N$), then we *define* c_l so that

$$c_l\Phi^N_{a_1a_2\ldots a_N} = 0. \tag{2.3}$$

Similarly, the *creation operator* $c^\dagger_l$ adds a state ϕ_l to a determinantal function not containing this state; that is to say, it converts an N-electron function not containing ϕ_l into an $(N+1)$-electron function containing ϕ_l, again with appropriate normalizing constant and sign. The formal definition is as follows:

$$\begin{aligned} c^\dagger_l\Phi^N_{a_1a_2\ldots a_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N) &= \pm\Phi^{N+1}_{a_1a_2\ldots a_jla_{j+1}\ldots a_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_{N+1}) \\ &= \pm\frac{1}{[(N+1)!]^{1/2}}\sum_P (-1)^p P\phi_{a_1}(\mathbf{x}_1)\ldots\phi_{a_j}(\mathbf{x}_j)\phi_l(\mathbf{x}_{j+1}) \\ &\quad\times\phi_{a_{j+1}}(\mathbf{x}_{j+2})\ldots\phi_N(\mathbf{x}_{N+1}), \end{aligned} \tag{2.4}$$

---

† To avoid confusion we use superscripts N, $N-1$, etc., to indicate the number of one-electron functions, that is, the number of rows, in the determinant.

where

$$a_1 < a_2 < \ldots < a_j < l < a_{j+1} \ldots < a_N, \tag{2.5}$$

and the sign is *positive* when j is *even* (that is, when an even number of functions precede ϕ_l in Φ^{N+1}) and *negative* when j is *odd*.

If Φ^N_a contains ϕ_{a_k}, that is, if a_k is a member of the set $a_1, a_2, \ldots a_N$, then $c^\dagger_{a_k}$ operates upon Φ^N_a to give zero. Thus

$$c^\dagger_{a_k}\Phi^N_{a_1a_2\ldots a_k\ldots a_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N) = 0. \tag{2.6}$$

This is simply an expression of the Pauli principle, and follows immediately from the definition (2.4), since a determinantal function is zero if the same one-electron function appears twice in it.

It may help to remove some of the apparent arbitrariness from the above definitions if we remark that the sign convention is simply that which leaves the sign of the determinant unchanged when a one-electron function is destroyed or created *in the first position* (that is, in the first row). The positive or negative sign which must be included is consequently that which results from moving a given one-electron function *to* the first position when it is to be *destroyed*, or *from* the first position to its correctly ordered position when it has been *created*. Thus, for example,

$$\begin{aligned} c_{a_1}\Phi^N_{a_1a_2\ldots a_N} &= \Phi^{N-1}_{a_2a_3\ldots a_N} \\ &= \frac{1}{[(N-1)!]^{1/2}}\sum_P (-1)^p P\phi_{a_2}(\mathbf{x}_1) \\ &\quad\times\phi_{a_3}(\mathbf{x}_2)\ldots\phi_{a_N}(\mathbf{x}_{N-1}), \end{aligned} \tag{2.7}$$

with positive sign, but

$$\begin{aligned} c_{a_2}\Phi^N_{a_1a_2\ldots a_N} &= -c_{a_2}\Phi^N_{a_2a_1\ldots a_N} \\ &= -\Phi^{N-1}_{a_1a_3\ldots a_N} \\ &= -\frac{1}{[(N-1)!]^{1/2}}\sum_P (-1)^p P\phi_{a_1}(\mathbf{x}_1) \\ &\quad\times\phi_{a_3}(\mathbf{x}_2)\ldots\phi_{a_N}(\mathbf{x}_{N-1}), \end{aligned} \tag{2.8}$$

with negative sign, in agreement with the definition (2.2). Similarly,

$$\begin{aligned} c^\dagger_l\Phi^N_{a_1a_2\ldots a_N} &= \Phi^{N+1}_{la_1a_2\ldots a_N} \\ &= \frac{1}{[(N+1)!]^{1/2}}\sum_P (-1)^p P\phi_l(\mathbf{x}_1) \\ &\quad\times\phi_{a_1}(\mathbf{x}_2)\ldots\phi_{a_N}(\mathbf{x}_{N+1}), \end{aligned} \tag{2.9}$$

so that, for example, if $a_1 < l < a_2$,

$$\begin{aligned} c_l^\dagger \Phi^N_{a_1 a_2 \dots a_N} &= \Phi^{N+1}_{l a_1 a_2 \dots a_N} \\ &= -\Phi^{N+1}_{a_1 l a_2 \dots a_N}, \end{aligned} \tag{2.10}$$

whereas, if $a_2 < l < a_3$,

$$\begin{aligned} c_l^\dagger \Phi^N_{a_1 a_2 \dots a_N} &= \Phi^{N+1}_{l a_1 a_2 \dots a_N} \\ &= \Phi^{N+1}_{a_1 a_2 l a_3 \dots a_N}, \end{aligned} \tag{2.11}$$

in agreement with the definition (2.4).

In spite of this rationalization of the signs, however, the definitions remain somewhat cumbersome and they will be put in more compact form in the following section by means of the *occupation number* formalism. It will be found eventually that the definitions rarely have to be used explicitly, since the most important properties of the creation and destruction operators are expressed by their *commutation relations*, which will be derived in §2.3. Meanwhile, it may be profitable to give some simple, concrete examples of the foregoing.

Let us take the case $N = 3$ and consider

$$\begin{aligned} \Phi^3_{123}(x_1, x_2, x_3) &= \frac{1}{(3!)^{1/2}} \sum_P (-1)^p P \phi_1(x_1)\phi_2(x_2)\phi_3(x_3) \\ &= \frac{1}{(3!)^{1/2}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(x_3) \\ \phi_3(x_1) & \phi_3(x_2) & \phi_3(x_3) \end{vmatrix} . \end{aligned} \tag{2.12}$$

Then it follows from (2.2), (2.3) and (2.4) that

$$\begin{aligned} c_2 \Phi^3_{123}(x_1, x_2, x_3) &= -\Phi^2_{13}(x_1, x_2) \\ &= -\frac{1}{(2!)^{1/2}} \sum_P (-1)^p P \phi_1(x_1)\phi_3(x_2) \\ &= -\frac{1}{(2!)^{1/2}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) \\ \phi_3(x_1) & \phi_3(x_2) \end{vmatrix}, \end{aligned} \tag{2.13}$$

$$c_2^\dagger \Phi^3_{123}(x_1, x_2, x_3) = 0, \tag{2.14}$$

$$c_4 \Phi^3_{123}(x_1, x_2, x_3) = 0, \tag{2.15}$$

$$\begin{aligned} c_4^\dagger \Phi^3_{123}(x_1, x_2, x_3) &= -\Phi^4_{1234}(x_1, x_2, x_3, x_4) \\ &= -\frac{1}{(4!)^{1/2}} \sum_P (-1)^p P \phi_1(x_1) \\ &\quad \times \phi_2(x_2)\phi_3(x_3)\phi_4(x_4), \end{aligned} \tag{2.16}$$

$$c_2 c_1 \Phi^3_{123}(x_1, x_2, x_3) = c_2 \Phi^2_{23}(x_1, x_2) = \Phi^1_3(x_1) = \phi_3(x_1), \tag{2.17}$$

$$c_1 c_2 \Phi^3_{123}(x_1, x_2, x_3) = -c_1 \Phi^2_{13}(x_1, x_2) = -\Phi^1_3(x_1) = -\phi_3(x_1), \tag{2.18}$$

$$c_2^\dagger c_2 \Phi^3_{123}(x_1, x_2, x_3) = -c_2^\dagger \Phi^2_{13}(x_1, x_2) = \Phi^3_{123}(x_1, x_2, x_3), \tag{2.19}$$

$$c_3 c_4^\dagger \Phi^3_{123}(x_1, x_2, x_3) = -c_3 \Phi^4_{1234}(x_1, x_2, x_3, x_4) = -\Phi^3_{124}(x_1, x_2, x_3), \tag{2.20}$$

$$c_4^\dagger c_3 \Phi^3_{123}(x_1, x_2, x_3) = c_4^\dagger \Phi^2_{12}(x_1, x_2) = \Phi^3_{124}(x_1, x_2, x_3). \tag{2.21}$$

2.2. Occupation Numbers

Instead of specifying a determinantal function Φ by the subscripts of the functions appearing in it, as in eq. (2.1), we could equally well do this by giving the occupation numbers of these functions. An *occupation number* n_i is defined as having the value 1 or 0 according to whether the function ϕ_i does or does not appear in Φ. Thus, an N-electron determinantal function may be written†

$$\Phi^N(n_1, n_2, n_3, \dots), \tag{2.22}$$

where N specified occupation numbers have the value unity and the rest are zero. We have used this notation to avoid a complexity of subscripts, but it must be understood that Φ^N is not a *function* of the n_i. The superscript N (which is generally omitted, although we include it for the sake of clarity) indicates that Φ^N is a function of the N coordinates $x_1, x_2, \dots x_N$. The number of the n_i is infinite, although in this case only N of them have the value unity, and they merely indicate which of the one-electron functions ϕ_i appear in the determinantal function Φ^N.

If we wish to show in (2.22) which of the n_i are unity and which zero without making a separate statement, we may write 1_i if ϕ_i appears in Φ and 0_i if it does not. Thus, the function Φ_a in eq. (2.1) may be written

$$\Phi^N(0_1, 0_2, \dots, 1_{a_1}, \dots, 1_{a_N}, \dots). \tag{2.23}$$

---

† This is often written $|n_1, n_2, n_3, \dots\rangle$, but we shall retain the wave-mechanical notation.

As a specific example, the function $\Phi^3_{123}(x_1, x_2, x_3)$ in eq. (2.12) may be expressed as

$$\Phi^3(1_1, 1_2, 1_3, 0_4, 0_5, \ldots), \tag{2.24}$$

the occupation numbers of the states ϕ_1, ϕ_2, ϕ_3 being unity and all the rest zero.

The destruction and creation operators may be defined more compactly using the occupation number representation. Thus the destruction operator c_k is defined by

$$c_k\Phi^N(\ldots 1_k \ldots) = \theta_k\Phi^{N-1}(\ldots 0_k \ldots), \tag{2.25}$$

$$c_k\Phi^N(\ldots 0_k \ldots) = 0, \tag{2.26}$$

where, in (2.25), it is understood that all the occupation numbers apart from n_k remain unaltered. The factor θ_k is plus or minus one according to whether an even or odd number of functions precede ϕ_k in Φ^N. It is easy to see that we may write

$$\theta_k = (-1)^{\sum\limits_{j<k} n_j} \tag{2.27}$$

for, if an even number of functions precede ϕ_k, each with occupation number unity, the sum is an even number and θ_k is plus one, while if an odd number of functions precede ϕ_k, the sum is an odd number and θ_k is minus one.

Similarly, the creation operator $c_k^\dagger$ is defined by

$$c_k^\dagger\Phi^N(\ldots 0_k \ldots) = \theta_k\Phi^{N+1}(\ldots 1_k \ldots), \tag{2.28}$$

$$c_k^\dagger\Phi^N(\ldots 1_k \ldots) = 0. \tag{2.29}$$

Finally, the pairs of eqs. (2.25), (2.26) and (2.28), (2.29) can be combined, provided we include a factor on the right-hand side which has value 1 or 0 as the case may be. The definitions in their most compact form are then

$$c_k\Phi^N(\ldots n_k \ldots) = \theta_k n_k\Phi^{N-1}(\ldots 0_k \ldots), \tag{2.30}$$

$$c_k^\dagger\Phi^N(\ldots n_k \ldots) = \theta_k(1-n_k)\Phi^{N+1}(\ldots 1_k \ldots), \tag{2.31}$$

with θ_k given by eq. (2.27). It is easily verified that these definitions agree exactly with the more cumbersome ones given in the previous section.

2.3. Commutation Relations

Let us consider the function $\Phi^N(\ldots n_k \ldots n_l \ldots)$, with $l > k$, and operate upon it first with c_k and then with c_l. We have, from eq. (2.30),

$$c_lc_k\Phi^N(\ldots n_k \ldots n_l) = \theta_k n_k c_l\Phi^{N-1}(\ldots 0_k \ldots n_l \ldots). \tag{2.32}$$

Similarly,

$$c_kc_l\Phi^N(\ldots n_k \ldots n_l \ldots) = \theta_l n_l c_k\Phi^{N-1}(\ldots n_k \ldots 0_l \ldots). \tag{2.33}$$

We see that, if either $n_k = 0$ or $n_l = 0$, both these expressions are zero. Let us therefore consider the case $n_k = n_l = 1$. Then

$$\begin{aligned} c_lc_k\Phi^N(\ldots 1_k \ldots 1_l \ldots) &= \theta_k c_l\Phi^{N-1}(\ldots 0_k \ldots 1_l \ldots) \\ &= \theta_k\theta_l'\Phi^{N-2}(\ldots 0_k \ldots 0_l \ldots), \end{aligned} \tag{2.34}$$

and

$$\begin{aligned} c_kc_l\Phi^N(\ldots 1_k \ldots 1_l \ldots) &= \theta_l c_k\Phi^{N-1}(\ldots 1_k \ldots 0_l \ldots) \\ &= \theta_l\theta_k'\Phi^{N-2}(\ldots 0_k \ldots 0_l \ldots). \end{aligned} \tag{2.35}$$

Here $\theta_l' = (-1)^p$, where p is the number of occupied states preceding ϕ_l in the function $\Phi^{N-1}(\ldots 0_k \ldots 1_l \ldots)$, and $\theta_k' = (-1)^q$, where q is the number of occupied states preceding ϕ_k in the function $\Phi^{N-1}(\ldots 1_k \ldots 0_l \ldots)$. It is clear that $\theta_k = \theta_k'$, since no states preceding ϕ_k have been destroyed or created. However, $\theta_l = -\theta_l'$, since the destruction of the state ϕ_k in Φ^N by c_k reduces the number of occupied states preceding ϕ_l by unity.

We thus have

$$c_lc_k\Phi^N = -c_kc_l\Phi^N, \tag{2.36}$$

for any determinantal function Φ^N (the fact that in the above proof we took $l > k$ for convenience is immaterial, since l and k can be interchanged without affecting the result). The equation remains true if $n_k = 0$ or $n_l = 0$ or $k = l$, both sides then being zero. Hence, for all k and l, we have the *operator equation*

$$c_lc_k + c_kc_l = 0. \tag{2.37}$$

This differs from the usual commutation relation in that the sign is plus instead of minus. For this reason c_l and c_k are said to *anticommute*, and the left-hand side of eq. (2.37) is called the *anticommutator* of c_l and c_k.

We shall denote[†] the anticommutator by the symbol $\{c_l, c_k\}$, using curly brackets instead of the square brackets which denote the commutator – in general,

$$\{A, B\} = AB + BA,$$

$$[A, B] = AB - BA.$$

Eq. (2.37) can thus be written

$$\{c_l, c_k\} = 0. \tag{2.38}$$

In a similar way[‡] it can be shown that the same anticommutation relation applies to the creation operators $c_l^\dagger$, $c_k^\dagger$, that is,

$$\{c_l^\dagger, c_k^\dagger\} = 0, \tag{2.39}$$

and this again is true for all l and k, including $k = l$.

Now let us consider the effect of one destruction and one creation operator. Suppose, first of all, that again $l > k$. Then, from (2.30) and (2.31), we have

$$c_l^\dagger c_k \Phi^N(\ldots n_k \ldots n_l \ldots) = \theta_k n_k c_l^\dagger \Phi^{N-1}(\ldots 0_k \ldots n_l \ldots), \tag{2.40}$$

and

$$c_k c_l^\dagger \Phi^N(\ldots n_k \ldots n_l \ldots) = \theta_l (1 - n_l) c_k \Phi^{N+1}(\ldots n_k \ldots 1_l \ldots). \tag{2.41}$$

If either $n_k = 0$ or $n_l = 1$, both these expressions vanish. Let us therefore take the case $n_k = 1$, $n_l = 0$. Then

$$\begin{aligned} c_l^\dagger c_k \Phi^N(\ldots 1_k \ldots 0_l \ldots) &= \theta_k c_l^\dagger \Phi^{N-1}(\ldots 0_k \ldots 0_l \ldots) \\ &= \theta_k \theta_l' \Phi^N(\ldots 0_k \ldots 1_l \ldots), \end{aligned} \tag{2.42}$$

[†] The symbol $[c_l, c_k]_+$ is also frequently used.

[‡] In fact, this can be shown more simply using the notation of §2.1. Thus,

$$\begin{aligned} c_l^\dagger c_k^\dagger \Phi^N_{a_1 a_2 \ldots a_N} &= c_l^\dagger \Phi^{N+1}_{k a_1 a_2 \ldots a_N} \\ &= \Phi^{N+2}_{l k a_1 a_2 \ldots a_N}, \end{aligned}$$

and

$$\begin{aligned} c_k^\dagger c_l^\dagger \Phi^N_{a_1 a_2 \ldots a_N} &= \Phi^{N+2}_{k l a_1 a_2 \ldots a_N} \\ &= -\Phi^{N+2}_{l k a_1 a_2 \ldots a_N}, \end{aligned}$$

from which eq. (2.39) follows immediately. The other commutation relations can also be found by this method, although not quite so directly.

and

$$\begin{aligned} c_k c_l^\dagger \Phi^N(\ldots 1_k \ldots 0_l \ldots) &= \theta_l c_k \Phi^{N+1}(\ldots 1_k \ldots 1_l \ldots) \\ &= \theta_l \theta_k'' \Phi^N(\ldots 0_k \ldots 1_l \ldots). \end{aligned} \tag{2.43}$$

Here $\theta_l' = (-1)^p$, where p is the number of occupied states preceding ϕ_l in Φ^{N-1}, and $\theta_k'' = (-1)^q$, where q is the number of occupied states preceding ϕ_k in Φ^{N+1}. The number of states preceding ϕ_k is unaltered in any of these operations, so that $\theta_k'' = \theta_k$, but $\theta_l' = -\theta_l$, since $n_k = 0$ in Φ^{N-1} and $n_k = 1$ in Φ^N. Hence

$$c_l^\dagger c_k \Phi^N(\ldots 1_k \ldots 0_l \ldots) = -c_k c_l^\dagger \Phi^N(\ldots 1_k \ldots 0_l \ldots). \tag{2.44}$$

This remains true if $n_k = 0$ or $n_l = 1$ in Φ^N, both sides of the equation being zero in these cases. So long as $l \neq k$, therefore, we have the operator equation

$$c_l^\dagger c_k + c_k c_l^\dagger = 0,$$

or

$$\{c_l^\dagger, c_k\} = 0, \quad l \neq k. \tag{2.45}$$

Now let us consider the case $l = k$. If $n_k = 1$, we find

$$\begin{aligned} c_k^\dagger c_k \Phi^N(\ldots 1_k \ldots) &= \theta_k c_k^\dagger \Phi^{N-1}(\ldots 0_k \ldots) \\ &= \theta_k^2 \Phi^N(\ldots 1_k \ldots) \\ &= \Phi^N(\ldots 1_k \ldots), \end{aligned} \tag{2.46}$$

and

$$c_k c_k^\dagger \Phi^N(\ldots 1_k \ldots) = 0, \tag{2.47}$$

while, if $n_k = 0$,

$$c_k^\dagger c_k \Phi^N(\ldots 0_k \ldots) = 0, \tag{2.48}$$

and

$$\begin{aligned} c_k c_k^\dagger \Phi^N(\ldots 0_k \ldots) &= \theta_k c_k \Phi^{N+1}(\ldots 1_k \ldots) \\ &= \theta_k^2 \Phi^N(\ldots 0_k \ldots) \\ &= \Phi^N(\ldots 0_k \ldots). \end{aligned} \tag{2.49}$$

Both pairs of equations lead to the operator equation

$$c_k^\dagger c_k + c_k c_k^\dagger = 1,$$

or

$$\{c_k^\dagger, c_k\} = 1, \tag{2.50}$$

one or other of the terms on the left-hand side giving zero result.

Combining (2.45) and (2.50), we may write, for all l and k, including $l=k$,

$$\{c_l^\dagger, c_k\} = \delta_{lk}. \tag{2.51}$$

It is clear, of course, that also

$$\{c_k, c_l^\dagger\} = \delta_{lk}, \tag{2.52}$$

since the order of adding the two terms in the anticommutator is immaterial.

To summarize: the commutation, or anticommutation, relations for the destruction and creation operators are

$$\{c_l, c_k\} = 0, \tag{2.53}$$

$$\{c_l^\dagger, c_k^\dagger\} = 0, \tag{2.54}$$

$$\{c_l^\dagger, c_k\} = \delta_{lk}, \tag{2.55}$$

and these are true for all values of l and k.

Finally, we note that eqs. (2.46) and (2.48) can be combined in the single equation

$$c_k^\dagger c_k \Phi^N(\ldots n_k \ldots) = n_k \Phi^N(\ldots n_k \ldots), \tag{2.56}$$

giving the operator equation

$$c_k^\dagger c_k = n_k. \tag{2.57}$$

For this reason the operator $c_k^\dagger c_k$ is called the *number operator* for the one-electron function ϕ_k.

2.4. The Vacuum State

If we operate with the destruction operator c_1 upon the first-order determinantal function $\Phi^1(1_1, 0_2, 0_3, \ldots)$, in which only the one-electron state ϕ_1 is occupied, the result is a fictitious 'zeroth-order determinant' in which no one-electron state is occupied. This is called the *vacuum state* or *empty state*, and we shall denote it by Φ_{vac}. Although clearly an artificial concept, it none the less has the convenient property that any determinantal function can be constructed by operating upon Φ_{vac} with the appropriate creation operators. Thus, for example,

$$c_1^\dagger \Phi_{\text{vac}}(0_1, 0_2, 0_3, \ldots) = \Phi^1(1_1, 0_2, 0_3, \ldots) = \phi(x_1), \tag{2.58}$$

$$\begin{aligned} c_2^\dagger c_1^\dagger \Phi_{\text{vac}}(0_1, 0_2, 0_3, \ldots) &= -\Phi^2(1_1, 1_2, 0_3, \ldots) \\ &= \frac{1}{\sqrt{2}}[\phi_2(x_1)\phi_1(x_2) - \phi_1(x_1)\phi_2(x_2)], \end{aligned} \tag{2.59}$$

$$\begin{aligned} c_3^\dagger c_1^\dagger c_2^\dagger \Phi_{\text{vac}}(0_1, 0_2, 0_3, \ldots) &= \Phi^3(1_1, 1_2, 1_3, 0_4 \ldots) \\ &= \frac{1}{(3!)^{1/2}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(x_3) \\ \phi_3(x_1) & \phi_3(x_2) & \phi_3(x_3) \end{vmatrix}, \end{aligned} \tag{2.60}$$

and so on.

It is, of course, unnecessary for the one-electron states to be created in any particular order, so long as the sign of the resulting determinantal function is correctly chosen. If the creation operators appear in the same order, from left to right, as that chosen for the one-electron states, the sign is always positive. For example,

$$c_1^\dagger c_2^\dagger c_4^\dagger \Phi_{\text{vac}} = \Phi^3(1_1, 1_2, 0_3, 1_4, 0_5, \ldots), \tag{2.61}$$

$$c_2^\dagger c_3^\dagger c_4^\dagger \Phi_{\text{vac}} = \Phi^3(0_1, 1_2, 1_3, 1_4, 0_5, \ldots). \tag{2.62}$$

This is because, at every stage, no occupied state precedes the one being created.

2.5. The Hamiltonian

We now wish to express the Hamiltonian

$$H = H_0 + H' \tag{2.63}$$

where

$$H_0 = \sum_{i=1}^{N} f(x_i), \tag{2.64}$$

$$H' = \tfrac{1}{2} \sum_{i \neq j}^{N} \sum^{N} v(x_i, x_j), \tag{2.65}$$

in terms of the destruction and creation operators.

We shall show that, when operating upon wave functions expressed as sums of determinants of the one-electron functions ϕ_i,

$$H_0 = \sum_{i,j} \langle i|f|j\rangle c_i^\dagger c_j, \tag{2.66}$$

where

$$\langle i|f|j\rangle = \int \phi_i^*(x) f(x) \phi_j(x)\, \mathrm{d}x \tag{2.67}$$

and the sum is over *all* values of i and j.

Furthermore, we shall show that

$$H' = \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k, \tag{2.68}$$

where

$$\langle ij|v|kl\rangle = \iint \phi_i^*(\boldsymbol{x}_1)\phi_j^*(\boldsymbol{x}_2)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\phi_k(\boldsymbol{x}_1)\phi_l(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2, \tag{2.69}$$

and the sum is over *all* values of i, j, k and l. The order of the product $c_l c_k$ should be noted.

We shall proceed in the most direct manner by comparing the matrix elements of (2.64) and (2.65) with those of (2.66) and (2.68) for an arbitrary pair of N-electron determinantal functions Φ_a and Φ_b. As pointed out in §1.4, if the corresponding matrix elements of two operators are the same, the operators are equivalent. However, although straightforward, this is a tedious process, and we shall defer the general treatment, for N-electron systems, to the following section. Here we shall simply make the above statements plausible by considering the simplest possible systems to which they can apply, namely, one-electron systems for H_0 and two-electron systems for H' (clearly, H' does not exist for a one-electron system, since more than one electron must be present to have an interaction).

2.5.1. *Matrix elements of H_0 for a one-electron system*

For a one-electron system only a single occupation number has value unity. Suppose that

$$\Phi_a = \Phi^1(\ldots 1_k \ldots) = \phi_k(\boldsymbol{x}_1), \tag{2.70}$$

$$\Phi_b = \Phi^1(\ldots 1_l \ldots) = \phi_l(\boldsymbol{x}_1). \tag{2.71}$$

Then, from (2.64),

$$\langle\Phi_a|H_0|\Phi_b\rangle = \langle\Phi_a|f(\boldsymbol{x}_1)|\Phi_b\rangle = \langle k|f|l\rangle. \tag{2.72}$$

Also, from (2.66),

$$\langle\Phi_a|H_0|\Phi_b\rangle = \left\langle\Phi_a\left|\sum_{i,j}\langle i|f|j\rangle c_i^\dagger c_j\right|\Phi_b\right\rangle$$

$$= \sum_{i,j}\langle i|f|j\rangle\langle\Phi_a|c_i^\dagger c_j|\Phi_b\rangle$$

$$= \sum_{i,j}\langle i|f|j\rangle \int \Phi_a^* c_i^\dagger c_j \Phi_b\,\mathrm{d}\tau'$$

$$= \langle k|f|l\rangle, \tag{2.73}$$

since $c_j\Phi_b = 0$ unless $j = l$, $c_l\Phi_b = \Phi_{\text{vac}}$, and $c_i^\dagger\Phi_{\text{vac}} = \Phi_a$ only if $i = k$.

This shows that the two forms of H_0 have the same matrix elements, and are therefore equivalent, for a one-electron system.

2.5.2. *Matrix elements of H' for a two-electron system*

For a two-electron system only two occupation numbers have value unity. Suppose that

$$\Phi_a = \Phi^2(\ldots 1_m \ldots 1_n \ldots)$$

$$= \frac{1}{\sqrt{2}}[\phi_m(\boldsymbol{x}_1)\phi_n(\boldsymbol{x}_2) - \phi_n(\boldsymbol{x}_1)\phi_m(\boldsymbol{x}_2)], \tag{2.74}$$

$$\Phi_b = \Phi^2(\ldots 1_p \ldots 1_q \ldots)$$

$$= \frac{1}{\sqrt{2}}[\phi_p(\boldsymbol{x}_1)\phi_q(\boldsymbol{x}_2) - \phi_q(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_2)]. \tag{2.75}$$

Then, from (2.65),

$$\langle\Phi_a|H'|\Phi_b\rangle = \tfrac{1}{2}\sum_{i\neq j}^{2}\sum^{2} \int \Phi_a^* v(\boldsymbol{x}_i,\boldsymbol{x}_j)\Phi_b\,\mathrm{d}\tau'$$

$$= \int \Phi_a^* v(\boldsymbol{x}_1,\boldsymbol{x}_2)\Phi_b\,\mathrm{d}\tau'$$

$$= \tfrac{1}{2}\iint [\phi_m^*(\boldsymbol{x}_1)\phi_n^*(\boldsymbol{x}_2) - \phi_n^*(\boldsymbol{x}_1)\phi_m^*(\boldsymbol{x}_2)]$$

$$\times v(\boldsymbol{x}_1,\boldsymbol{x}_2)[\phi_p(\boldsymbol{x}_1)\phi_q(\boldsymbol{x}_2) - \phi_q(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_2)]\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2$$

$$= \tfrac{1}{2}[\langle mn|v|pq\rangle - \langle mn|v|qp\rangle - \langle nm|v|pq\rangle + \langle nm|v|qp\rangle]$$

$$= \langle mn|v|pq\rangle - \langle mn|v|qp\rangle, \tag{2.76}$$

where, in the second line, we have used the fact that $v(\boldsymbol{x}_1,\boldsymbol{x}_2) = v(\boldsymbol{x}_2,\boldsymbol{x}_1)$, and, in the final line, that $\langle mn|v|pq\rangle = \langle nm|v|qp\rangle$, as shown in eq. (1.51).

Now, from (2.68), we have

$$\langle\Phi_a|H'|\Phi_b\rangle = \left\langle\Phi_a\left|\tfrac{1}{2}\sum_{i,j,k,l}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k\right|\Phi_b\right\rangle$$

$$= \tfrac{1}{2}\sum_{i,j,k,l}\langle ij|v|kl\rangle \int \Phi_a^* c_i^\dagger c_j^\dagger c_l c_k \Phi_b\,\mathrm{d}\tau', \tag{2.77}$$

and the integral vanishes unless

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_b = \pm\Phi_a, \tag{2.78}$$

that is, unless $l, k = p, q$ and $i, j = m, n$.

If $k = p, l = q, j = n, i = m$, we have, from (2.74) and (2.75),

$$\begin{aligned} c_m^\dagger c_n^\dagger c_q c_p \Phi_b &= c_m^\dagger c_n^\dagger c_q c_p \Phi^2(\ldots 1_p \ldots 1_q \ldots) \\ &= c_m^\dagger c_n^\dagger c_q \Phi^1(\ldots 1_q \ldots) \\ &= c_m^\dagger c_n^\dagger \Phi_{\text{vac}} \\ &= c_m^\dagger \Phi^1(\ldots 1_n \ldots) \\ &= \Phi^2(\ldots 1_m \ldots 1_n \ldots) = \Phi_a, \end{aligned} \tag{2.79}$$

and the value of the integral in (2.77) is unity. Then, using the commutation relations for the operators, we have immediately, for

$$k = q, l = p, j = n,\ i = m, \qquad \int = -1,$$

$$k = p, l = q, j = m, i = n, \qquad \int = -1,$$

$$k = q, l = p, j = m, i = n, \qquad \int = +1.$$

Hence, from (2.68),

$$\begin{aligned} \langle\Phi_a|H'|\Phi_b\rangle &= \tfrac{1}{2}[\langle mn|v|pq\rangle - \langle mn|v|qp\rangle - \langle nm|v|pq\rangle + \langle nm|v|qp\rangle] \\ &= \langle mn|v|pq\rangle - \langle mn|v|qp\rangle, \end{aligned} \tag{2.80}$$

as in eq. (2.76).

This shows that the two forms of H' have the same matrix elements, and hence are equivalent, for a two-electron system.

2.6. Matrix Elements of H for an N-electron System

In most texts on second quantization or perturbation theory the expressions for H_0 and H' given in eqs. (2.66) and (2.68) are simply stated, with little or no attempt at derivation. Even an elementary verification of (2.68) such as that of the preceding section is rarely included, although some justification of (2.66) may be given. Many readers may be content to accept these expressions as stated, and for others the incomplete verifications of the previous section may suffice. For such, the present section may be omitted. However, for the more sceptical reader, we present the following straightforward and rather tedious, but completely rigorous, verification of the expression for the Hamiltonian in the occupation number representation. If it does nothing more, this work will provide beginners with a useful exercise in the manipulation of the creation and destruction operators.

The method used will be a direct extension to N-electron systems of that of the previous section. The matrix elements of H_0 and H', as expressed in (2.66) and (2.68), with respect to a set of N-electron determinantal functions, will be compared with those already found in ch. 1.

2.6.1. *Matrix elements of* H_0

If Φ_a and Φ_b are N-electron determinantal functions and H_0 is given by (2.66), we have

$$\begin{aligned} \langle\Phi_b|H_0|\Phi_a\rangle &= \int \Phi_b^*\Big[\sum_{i,j} \langle i|f|j\rangle c_i^\dagger c_j\Big]\Phi_a \,\mathrm{d}\tau' \\ &= \sum_{i,j} \langle i|f|j\rangle \int \Phi_b^* c_i^\dagger c_j \Phi_a \,\mathrm{d}\tau'. \end{aligned} \tag{2.81}$$

(1) As in §1.4.1, let us first consider a diagonal element, with $\Phi_b = \Phi_a$. Owing to the orthogonality of the determinantal functions Φ_a, the integral

$$\int \Phi_a^* c_i^\dagger c_j \Phi_a \,\mathrm{d}\tau'$$

vanishes unless

$$c_i^\dagger c_j \Phi_a = \pm\Phi_a. \tag{2.82}$$

This can only be true if $i = j$ and Φ_a contains ϕ_j, in which case the sign on the right-hand side is positive. We have, in fact, from eq. (2.57) and the normalization of the Φ_a,

$$\int \Phi_a^* c_i^\dagger c_j \Phi_a \,\mathrm{d}\tau' = n_j. \tag{2.83}$$

Thus, from (2.81) and (2.83),

$$\begin{aligned} \langle\Phi_a|H_0|\Phi_a\rangle &= \sum_{i,j} \langle i|f|j\rangle \int \Phi_a^* c_i^\dagger c_j \Phi_a \,\mathrm{d}\tau' \\ &= \sum_j \langle j|f|j\rangle n_j. \end{aligned} \tag{2.84}$$

However, for the function Φ_a, which is $\Phi_{a_1,a_2,\dots a_N}(x_1, x_2, \dots x_N)$ in our earlier notation, we have

$$n_j = \begin{cases} 1, & \text{for } j = a_1, a_2, \dots, a_N, \\ 0, & \text{otherwise.} \end{cases}$$

Hence (2.84) becomes

$$\langle\Phi_a|H_0|\Phi_a\rangle = \sum_{i=1}^{N} \langle a_i|f|a_i\rangle, \tag{2.85}$$

which is exactly the result we obtained in eq. (1.41). The diagonal matrix elements of (2.64) and (2.66) are thus the same.

(2) Next, let us suppose that Φ_a and Φ_b differ in only one function, so that, for example,

$$\Phi_a = \Phi^N(\dots 0_k \dots 1_l \dots), \tag{2.86}$$

$$\Phi_b = \Phi^N(\dots 1_k \dots 0_l \dots), \tag{2.87}$$

all the other occupation numbers being the same in the two determinantal functions. In other words, Φ_a and Φ_b are the same, except that in Φ_b the function ϕ_k is substituted for the ϕ_l appearing in Φ_a. Also, we shall assume that *the occupation numbers of all the one-electron functions lying between ϕ_k and ϕ_l are zero.*

The latter assumption is to ensure that, when Φ_a and Φ_b are written out as in eq. (2.1), ϕ_k and ϕ_l will occupy the same positions[†], as was assumed in calculating the matrix elements in §1.4.1 (e.g., $a_j = b_j$ for $j \neq k$, but $a_k \neq b_k$). This obviates any difficulty with the sign of the matrix elements when comparing with those found in ch. 1. However, no generality is lost here, for a one-electron function can be changed to any required position in a determinantal function by a permutation which will at most change the sign of the matrix element. Precisely the same permutation would be required in the calculation of ch. 1, in order to produce the ordering assumed there, if the one-electron functions were originally ordered according to the scheme of §2.1.

With Φ_a and Φ_b given by (2.86) and (2.87), the integral

$$\int \Phi_b^* c_i^\dagger c_j \Phi_a \, d\tau'$$

vanishes unless $j = l$ and $i = k$, when it has the value unity. Thus, from eq. (2.81),

$$\langle\Phi_b|H_0|\Phi_a\rangle = \langle k|f|l\rangle, \tag{2.88}$$

and this is exactly the result found in eq. (1.43), if we simply write a_k and b_k instead of l and k for the subscripts of the functions in which Φ_a and Φ_b differ.

(3) Again, as found in §1.4.1, the matrix element must be zero if Φ_a and Φ_b differ in two or more functions, that is, in four or more occupation numbers, for then the operator $c_i^\dagger c_j$ can never transform Φ_a into Φ_b, and the integrals in eq. (2.81) must all vanish.

This completes the proof that the matrix elements of (2.64) are identical with those of (2.66), for an N-electron system, and the two forms of H_0 are therefore equivalent.

2.6.2. *Matrix elements of H'*

If Φ_a and Φ_b are N-electron determinantal functions and H' is given by (2.68), we have

$$\langle\Phi_b|H'|\Phi_a\rangle = \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl\rangle \int \Phi_b^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \, d\tau'. \tag{2.89}$$

(1) As in §1.4.2, we first consider a diagonal element, with $\Phi_b = \Phi_a$. Owing to the orthogonality of the Φ_a, the integral

$$\int \Phi_a^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \, d\tau'$$

vanishes unless

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = \pm\Phi_a. \tag{2.90}$$

For a non-zero value, therefore, Φ_a must contain ϕ_l and ϕ_k, with $k \neq l$, and we must have either $i = k, j = l$ or $i = l, j = k$.

[†] As a concrete example let us take $N = 3$ and suppose that Φ_a contains ϕ_1, ϕ_2 and ϕ_5, while Φ_b contains ϕ_1, ϕ_4 and ϕ_5. We then have

$$\Phi_a^3(1_1, 1_2, 0_3, 0_4, 1_5, 0_6, \dots) = \frac{1}{(3!)^{1/2}} \sum_P (-1)^p P \phi_1(x_1)\phi_2(x_2)\phi_5(x_3),$$

and

$$\Phi_b^3(1_1, 0_2, 0_3, 1_4, 1_5, 0_6, \dots) = \frac{1}{(3!)^{1/2}} \sum_P (-1)^p P \phi_1(x_1)\phi_4(x_2)\phi_5(x_3).$$

We see that ϕ_2 occupies the same position in Φ_a as ϕ_4 does in Φ_b. This is so only because the occupation number of ϕ_3, lying between ϕ_2 and ϕ_4, is zero in both cases.

From the commutation relations (2.53), (2.55), we find, if $j \neq l$,

$$c_i^\dagger c_j^\dagger c_l c_k = -c_i^\dagger c_l c_j^\dagger c_k, \tag{2.91}$$

and, if $k \neq j$,

$$c_i^\dagger c_j^\dagger c_l c_k = -c_i^\dagger c_j^\dagger c_k c_l = c_i^\dagger c_k c_j^\dagger c_l. \tag{2.92}$$

Hence, if $i = l, j = k, k \neq l$, we have, from (2.91),

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = -c_i^\dagger c_i c_j^\dagger c_j \Phi_a = -n_i n_j \Phi_a. \tag{2.93}$$

Alternatively, if $i = k, j = l, k \neq l$, we have, from (2.92),

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_i^\dagger c_i c_j^\dagger c_j \Phi_a = n_i n_j \Phi_a. \tag{2.94}$$

Thus, from (2.89),

$$\langle \Phi_a | H' | \Phi_a \rangle = \tfrac{1}{2} \sum_{i,j} n_i n_j [\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \tag{2.95}$$

Here it is unnecessary to specify $i \neq j$, as the factor in square brackets vanishes when $i = j$.

As before, for the function Φ_a, defined in eq. (2.1), we have

$$n_j = \begin{cases} 1, & \text{for } \; j = a_1, a_2, \ldots, a_N, \\ 0, & \text{otherwise,} \end{cases}$$

and similarly for n_i. Hence (2.95) becomes

$$\langle \Phi_a | H' | \Phi_a \rangle = \tfrac{1}{2} \sum_{i,j}^{N} [\langle a_i a_j | v | a_i a_j \rangle - \langle a_i a_j | v | a_j a_i \rangle], \tag{2.96}$$

which is exactly the result found in eq. (1.48). The diagonal elements of (2.65) and (2.68) are thus identical.

(2) Next, let us suppose that Φ_a and Φ_b differ in only one function, as in eqs. (2.86) and (2.87); say

$$\Phi_a = \Phi^N(\ldots 1_p \ldots 0_q \ldots), \tag{2.97}$$

$$\Phi_b = \Phi^N(\ldots 0_p \ldots 1_q \ldots), \tag{2.98}$$

all the other occupation numbers being the same in the two determinants. Again, in order to make an immediate comparison with the results of ch. 1, we shall assume that the occupation numbers of all one-electron functions lying between ϕ_p and ϕ_q are zero.

The integral

$$\int \Phi_b^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \, d\tau'$$

vanishes unless

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = \pm \Phi_b. \tag{2.99}$$

The action of the operator on Φ_a must then be to destroy ϕ_p and create ϕ_q, for a non-zero value. It follows that either l or k must be p and either i or j must be q, while the remaining pair of one destruction and one creation operator must simply destroy and create the same function in Φ_a. We note also that we must have $l \neq k$, $i \neq j$, since the same function can neither be created nor destroyed twice in succession. There are four possibilities:

(a) $k = p, j = q, i = l$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_i^\dagger c_q^\dagger c_i c_p \Phi_a = -c_i^\dagger c_i c_q^\dagger c_p \Phi_a = -n_i \Phi_b; \tag{2.100}$$

(b) $k = p, i = q, j = l$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_q^\dagger c_j^\dagger c_j c_p \Phi_a = c_j^\dagger c_j c_q^\dagger c_p \Phi_a = n_j \Phi_b; \tag{2.101}$$

(c) $l = p, j = q, i = k$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_i^\dagger c_q^\dagger c_p c_i \Phi_a = c_i^\dagger c_i c_q^\dagger c_p \Phi_a = n_i \Phi_b; \tag{2.102}$$

(d) $l = p, i = q, j = k$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_q^\dagger c_j^\dagger c_p c_j \Phi_a = -c_j^\dagger c_j c_q^\dagger c_p \Phi_a = -n_j \Phi_b. \tag{2.103}$$

The results are the same whether $p > q$ or $q > p$, so long as the occupation numbers of states lying between ϕ_p and ϕ_q are zero.

Eq. (2.89) thus becomes

$$\begin{aligned} \langle \Phi_b | H' | \Phi_a \rangle &= \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl \rangle \int \Phi_b^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \, d\tau' \\ &= \tfrac{1}{2} \sum_j \langle qj|v|pj \rangle n_j - \tfrac{1}{2} \sum_i \langle iq|v|pi \rangle n_i \\ &\quad + \tfrac{1}{2} \sum_i \langle iq|v|ip \rangle n_i - \tfrac{1}{2} \sum_j \langle qj|v|jp \rangle n_j \\ &= \sum_i [\langle iq|v|ip \rangle - \langle qi|v|ip \rangle] n_i, \end{aligned} \tag{2.104}$$

again using the fact, proved in eq. (1.51), that

$$\langle ij|v|kl\rangle = \langle ji|v|lk\rangle. \tag{2.105}$$

The occupation number n_i is, strictly speaking, that of ϕ_i in Φ_b; but this is the same as the occupation number of ϕ_i in Φ_a, except when $i = p$ or $i = q$. However, the square bracket in eq. (2.104) vanishes when $i = p$ or $i = q$, so that we may take

$$n_i = \begin{cases} 1, & \text{for} \quad i = a_1, a_2, \ldots, a_N, \\ 0, & \text{otherwise,} \end{cases}$$

as before. Then, if we write a_k for p and b_k for q, the states in which Φ_a and Φ_b differ, (2.104) becomes

$$\langle\Phi_b|H'|\Phi_a\rangle = \sum_{i=1}^{N} [\langle a_i b_k|v|a_i a_k\rangle - \langle b_k a_i|v|a_i a_k\rangle], \tag{2.106}$$

which agrees with eq. (1.50).

We have thus shown that the matrix elements of (2.65) and (2.68) are the same between two N-electron determinantal functions which differ in a single one-electron function only.

(3) Now let us suppose that Φ_a and Φ_b differ in two functions, say Φ_a contains ϕ_p and ϕ_q but not ϕ_r and ϕ_s, while Φ_b contains ϕ_r and ϕ_s but not ϕ_p and ϕ_q, so that

$$\Phi_a = \Phi^N(\ldots 1_p \ldots 0_r \ldots 1_q \ldots 0_s \ldots), \tag{2.107}$$

$$\Phi_b = \Phi^N(\ldots 0_p \ldots 1_r \ldots 0_q \ldots 1_s \ldots), \tag{2.108}$$

all the other occupation numbers being the same. We also assume that the occupation numbers of all states between ϕ_p and ϕ_r and between ϕ_q and ϕ_s are zero in both Φ_a and Φ_b. This ensures that when Φ_a and Φ_b are written out as in eq. (2.1) the states ϕ_p, ϕ_q and the states ϕ_r, ϕ_s will occupy the same positions, as was assumed in ch. 1. It may be seen that, in order for this to be so, we must have either $p, r < q, s$ or $q, s < p, r$, but it is immaterial whether $p < r$ or $r < p$ or whether $q < s$ or $s < q$.

As before, the integral

$$\int \Phi_b^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \,\mathrm{d}\tau'$$

vanishes unless

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = \pm\Phi_b. \tag{2.109}$$

For a non-zero value, therefore, the destruction operators must destroy the states ϕ_p and ϕ_q, and the creation operators must create the states ϕ_r and ϕ_s. Also we note that p, q, r and s are all different. There are again four possibilities:

(a) $i = r, j = s, l = p, k = q$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_r^\dagger c_s^\dagger c_p c_q \Phi_a = -c_r^\dagger c_p c_s^\dagger c_q \Phi_a = -\Phi_b; \tag{2.110}$$

(b) $i = r, j = s, l = q, k = p$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_r^\dagger c_s^\dagger c_q c_p \Phi_a = c_r^\dagger c_p c_s^\dagger c_q \Phi_a = \Phi_b; \tag{2.111}$$

(c) $i = s, j = r, l = p, k = q$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_s^\dagger c_r^\dagger c_p c_q \Phi_a = c_r^\dagger c_p c_s^\dagger c_q \Phi_a = \Phi_b; \tag{2.112}$$

(d) $i = s, j = r, l = q, k = p$, so that

$$c_i^\dagger c_j^\dagger c_l c_k \Phi_a = c_s^\dagger c_r^\dagger c_q c_p \Phi_a = -c_s^\dagger c_q c_r^\dagger c_p \Phi_a = -\Phi_b. \tag{2.113}$$

Eq. (2.89) becomes in this case

$$\begin{aligned} \langle\Phi_b|H'|\Phi_a\rangle &= \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl\rangle \int \Phi_b^* c_i^\dagger c_j^\dagger c_l c_k \Phi_a \,\mathrm{d}\tau' \\ &= -\tfrac{1}{2}\langle rs|v|qp\rangle + \tfrac{1}{2}\langle rs|v|pq\rangle + \tfrac{1}{2}\langle sr|v|qp\rangle - \tfrac{1}{2}\langle sr|v|pq\rangle \\ &= \langle rs|v|pq\rangle - \langle sr|v|pq\rangle. \end{aligned} \tag{2.114}$$

If we simply write a_k instead of p, a_l instead of q, b_k instead of r and b_l instead of s, which was the notation used in ch. 1, we obtain

$$\langle\Phi_b|H'|\Phi_a\rangle = \langle b_k b_l|v|a_k a_l\rangle - \langle b_l b_k|v|a_k a_l\rangle, \tag{2.115}$$

which is exactly the result obtained in eq. (1.54). This shows that the matrix elements of (2.65) and (2.68) are the same between two N-electron determinantal functions which differ in two one-electron states.

(4) Eq. (2.89) shows immediately, as was found in §1.4.1, that the matrix element $\langle\Phi_b|H'|\Phi_a\rangle$ must be zero if Φ_a and Φ_b differ in more than two one-electron functions, for then two destruction and two creation operators could never convert Φ_a into $\pm\Phi_b$.

This completes the proof that the matrix elements of (2.65) are identical with those of (2.68), for an N-electron system, and the two forms of H' are therefore equivalent.

To summarize: we have proved, by direct comparison of their matrix elements, that the operators

$$H=\sum_{i=1}^{N} f(\boldsymbol{x}_i)+\tfrac{1}{2}\sum_{i,j\neq i}^{N} v(\boldsymbol{x}_i,\boldsymbol{x}_j) \tag{2.116}$$

and

$$H=\sum_{i,j}\langle i|f|j\rangle c_i^\dagger c_j+\tfrac{1}{2}\sum_{i,j,k,l}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k \tag{2.117}$$

are equivalent for an N-electron system. It is worth noting, however, that there are certain differences between these operators. For example, (2.116) is meaningful when applied to an N-electron wave function in any analytical form, while (2.117) is meaningful only when applied to a wave function expressed as a sum of determinantal functions, the creation and destruction operators being defined in relation to the one-electron functions which are the elements of these determinants. On the other hand, (2.117) is in one sense more general than (2.116), since it does not depend upon N – it is the same whatever the number of electrons in the system. N now enters the problem only via the determinantal functions from which the total wave function is constructed – for an N-electron system these are N-th-order determinants.

The foregoing proof is undoubtedly exhausting and unlikely to be followed in detail by many readers. None the less, even a cursory perusal of it may serve at least one useful purpose. It demonstrates quite clearly that, in spite of the use of apocalyptic nouns such as *creation* and *destruction*, the so-called second quantization formalism, in this context at least, involves *no new principle*. Nothing, in fact, is being created or destroyed – these operators merely afford a convenient means of expressing one determinantal function in terms of another.

2.7. To Prove that c_i and $c_i^\dagger$ are Hermitian Conjugates

The use of the symbol $c_i^\dagger$ for the creation operator of the state ϕ_i implies that it is the Hermitian conjugate[†] of the destruction operator c_i. It can easily be shown that this is so.

Let us suppose that

$$\Phi_a=\Phi^{N+1}(\ldots 1_i\ldots) \tag{2.118}$$

and

$$\Phi_b=\Phi^{N}(\ldots 0_i\ldots), \tag{2.119}$$

[†] See Appendix I.

all the other occupation numbers of these two functions being the same. Also, to make the sign definite, let the number of occupied states preceding ϕ_i be *even*. Then

$$\int \Phi_b^* c_i \Phi_a \,\mathrm{d}\tau' = 1. \tag{2.120}$$

Now let c_i^{H} be the Hermitian conjugate of c_i. Then, by definition,

$$\int \Phi_b^* c_i \Phi_a \,\mathrm{d}\tau' = \int (c_i^{\mathrm{H}}\Phi_b)^* \Phi_a \,\mathrm{d}\tau' = \left(\int \Phi_a^* c_i^{\mathrm{H}} \Phi_b \,\mathrm{d}\tau'\right)^* = 1, \tag{2.121}$$

using (2.120).

Since there is no restriction on Φ_b, except that it must not contain ϕ_i, it follows that

$$c_i^{\mathrm{H}}\Phi_b=\Phi_a. \tag{2.122}$$

In other words, c_i^{H} is the creation operator for the state ϕ_i, or

$$c_i^{\mathrm{H}}=c_i^\dagger, \tag{2.123}$$

as we wished to prove.

This result enables us to express the matrix elements of the previous section in a more symmetrical, and sometimes more convenient, form. For example, eq. (I.5) of Appendix I tells us that

$$(c_i^\dagger)^\dagger=c_i, \tag{2.124}$$

so that, for any determinantal functions Φ_m, Φ_n,

$$\int \Phi_m^* c_i^\dagger c_j \Phi_n \,\mathrm{d}\tau' = \int (c_i\Phi_m)^* c_j \Phi_n \,\mathrm{d}\tau', \tag{2.125}$$

that is,

$$\langle \Phi_m|c_i^\dagger c_j|\Phi_n\rangle = \langle c_i\Phi_m|c_j\Phi_n\rangle. \tag{2.126}$$

In the same way, if $i\neq j$,

$$\langle \Phi_m|c_i^\dagger c_j|\Phi_n\rangle = -\langle \Phi_m|c_j c_i^\dagger|\Phi_n\rangle = -\langle c_j^\dagger\Phi_m|c_i^\dagger\Phi_n\rangle. \tag{2.127}$$

Also, from eq. (I.11) of Appendix I, we find

$$\langle \Phi_m|c_i^\dagger c_j^\dagger c_k c_l|\Phi_n\rangle = \langle c_j c_i\Phi_m|c_k c_l\Phi_n\rangle, \tag{2.128}$$

so that this matrix element vanishes unless

$$c_j c_i\Phi_m=\pm c_k c_l\Phi_n, \tag{2.129}$$

and Φ_n contains ϕ_k and ϕ_l.

2.8. Field Operators

The equations of the many-body theory, in the occupation number formalism, are often expressed in terms of the so-called *field operators*, $\boldsymbol{\psi}(\boldsymbol{x})$ and $\boldsymbol{\psi}^\dagger(\boldsymbol{x})$, defined by

$$\boldsymbol{\psi}(\boldsymbol{x}) = \sum_i \phi_i(\boldsymbol{x})c_i, \tag{2.130}$$

$$\boldsymbol{\psi}^\dagger(\boldsymbol{x}) = \sum_i \phi_i^*(\boldsymbol{x})c_i^\dagger. \tag{2.131}$$

The commutation relations for these operators are

$$\{\boldsymbol{\psi}(\boldsymbol{x}), \boldsymbol{\psi}(\boldsymbol{x}')\} = 0, \tag{2.132}$$

$$\{\boldsymbol{\psi}^\dagger(\boldsymbol{x}), \boldsymbol{\psi}^\dagger(\boldsymbol{x}')\} = 0, \tag{2.133}$$

$$\{\boldsymbol{\psi}(\boldsymbol{x}), \boldsymbol{\psi}^\dagger(\boldsymbol{x}')\} = \delta(\boldsymbol{x}-\boldsymbol{x}'), \tag{2.134}$$

and can easily be obtained from those for the creation and destruction operators, eqs. (2.53), (2.54) and (2.55). Thus

$$\begin{aligned}\{\boldsymbol{\psi}(\boldsymbol{x}), \boldsymbol{\psi}(\boldsymbol{x}')\} &= \sum_{i,j} \phi_i(\boldsymbol{x})\phi_j(\boldsymbol{x}')c_ic_j + \sum_{i,j} \phi_j(\boldsymbol{x}')\phi_i(\boldsymbol{x})c_jc_i \\ &= \sum_{i,j} \phi_i(\boldsymbol{x})\phi_j(\boldsymbol{x}')\{c_i, c_j\} = 0,\end{aligned} \tag{2.135}$$

and eq. (2.133) follows in the same way. In the case of (2.134) we have

$$\begin{aligned}\{\boldsymbol{\psi}(\boldsymbol{x}), \boldsymbol{\psi}^\dagger(\boldsymbol{x}')\} &= \sum_{i,j} \phi_i(\boldsymbol{x})\phi_j^*(\boldsymbol{x}')c_ic_j^\dagger + \sum_{i,j} \phi_j^*(\boldsymbol{x}')\phi_i(\boldsymbol{x})c_j^\dagger c_i \\ &= \sum_{i,j} \phi_i(\boldsymbol{x})\phi_j^*(\boldsymbol{x}')\{c_i, c_j^\dagger\} \\ &= \sum_i \phi_i(\boldsymbol{x})\phi_i^*(\boldsymbol{x}') = \delta(\boldsymbol{x}-\boldsymbol{x}'),\end{aligned} \tag{2.136}$$

from the completeness condition (1.15).

We will now show that H_0, eq. (2.66), may be written

$$H_0 = \int \boldsymbol{\psi}^\dagger(\boldsymbol{x})f(\boldsymbol{x})\boldsymbol{\psi}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{2.137}$$

The right-hand side becomes

$$\begin{aligned}\int \sum_i \phi_i^*(\boldsymbol{x})c_i^\dagger f(\boldsymbol{x}) \sum_j \phi_j(\boldsymbol{x})c_j\mathrm{d}\boldsymbol{x} &= \sum_{i,j} c_i^\dagger c_j \int \phi_i^*(\boldsymbol{x})f(\boldsymbol{x})\phi_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \sum_{i,j} \langle i|f|j\rangle c_i^\dagger c_j,\end{aligned} \tag{2.138}$$

which is H_0.

Similarly, H', eq. (2.68), may be written

$$H' = \tfrac{1}{2}\int \boldsymbol{\psi}^\dagger(\boldsymbol{x}_1)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2)v(\boldsymbol{x}_1, \boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_1)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2. \tag{2.139}$$

The right-hand side becomes

$$\begin{aligned}&\tfrac{1}{2}\int \sum_i \phi_i^*(\boldsymbol{x}_1)c_i^\dagger \sum_j \phi_j^*(\boldsymbol{x}_2)c_j^\dagger v(\boldsymbol{x}_1, \boldsymbol{x}_2) \sum_l \phi_l(\boldsymbol{x}_2)c_l \sum_k \phi_k(\boldsymbol{x}_1)c_k\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &= \tfrac{1}{2}\sum_{i,j,k,l} c_i^\dagger c_j^\dagger c_l c_k \int \phi_i^*(\boldsymbol{x}_1)\phi_j^*(\boldsymbol{x}_2)v(\boldsymbol{x}_1, \boldsymbol{x}_2)\phi_k(\boldsymbol{x}_1)\phi_l(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &= \tfrac{1}{2}\sum_{i,j,k,l} \langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k,\end{aligned} \tag{2.140}$$

which is H'.

It should be noted that (2.137) is the occupation number representation of a sum of one-electron operators $f(\boldsymbol{x}_i)$, and (2.139) that of a sum of two-electron operators $\frac{1}{2}v(\boldsymbol{x}_i, \boldsymbol{x}_j)$, whether or not these occur in the Hamiltonian. For example, consider the *particle density*[†], including spin,

$$\rho(\boldsymbol{x}) = \sum_{i=1}^{N} \delta(\boldsymbol{x}-\boldsymbol{x}_i). \tag{2.141}$$

It follows that the particle density operator is given by[‡]

$$\begin{aligned}\rho(\boldsymbol{x}) &= \int \boldsymbol{\psi}^\dagger(\boldsymbol{x}')\delta(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{\psi}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}' \\ &= \boldsymbol{\psi}^\dagger(\boldsymbol{x})\boldsymbol{\psi}(\boldsymbol{x}).\end{aligned} \tag{2.142}$$

It should be noted that, in deriving the above properties of the field operators, we have not required to interpret these operators physically. This again emphasizes the fact that no new physical principle is being invoked. However, it is sometimes a useful aid to intuition to use, even if only loosely, the physical interpretation of these operators, which we will now describe.

[†] $\rho(\boldsymbol{x})$ is the number density of electrons at point $\boldsymbol{r}$ with spin coordinate ζ, $\delta(\boldsymbol{x}-\boldsymbol{x}_i)$ being a shorthand notation for $\delta(\boldsymbol{r}-\boldsymbol{r}_i)\delta_{\zeta\zeta_i}$. This means that the integral of $\rho(\boldsymbol{x})$, including a sum over the two values of ζ, throughout a given region is equal to the number of electrons in the region. Thus, if a number ν of the points $\boldsymbol{r}_i$ lie in a region Ω, we have

$$\int_\Omega \rho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_i \sum_{\zeta=\pm1} \delta_{\zeta\zeta_i}\int_\Omega \delta(\boldsymbol{r}-\boldsymbol{r}_i)\,\mathrm{d}\boldsymbol{r} = \nu.$$

[‡] It is necessary here to change the variable of integration to $\boldsymbol{x}'$ owing to the presence of $\boldsymbol{x}$ in $\delta(\boldsymbol{x}-\boldsymbol{x}_i)$, which replaces $f(\boldsymbol{x}_i)$ in this case.

The operators $\boldsymbol{\psi}^\dagger(\boldsymbol{x})$ and $\boldsymbol{\psi}(\boldsymbol{x})$ are said to create and destroy, respectively, a particle at $\boldsymbol{x}$ (that is, at position $\boldsymbol{r}$ with spin coordinate ζ). This is most easily demonstrated by operating upon Φ_{vac} with $\boldsymbol{\psi}^\dagger(\boldsymbol{x})$, which gives

$$\begin{aligned}\boldsymbol{\psi}^\dagger(\boldsymbol{x})\Phi_{\mathrm{vac}} &= \sum_i \phi_i^*(\boldsymbol{x})c_i^\dagger\Phi_{\mathrm{vac}} \\ &= \sum_i \phi_i^*(\boldsymbol{x})\phi_i(\boldsymbol{x}_1) = \delta(\boldsymbol{x}-\boldsymbol{x}_1), \qquad (2.143)\end{aligned}$$

using the completeness condition (1.15) and the fact that

$$c_i^\dagger\Phi_{\mathrm{vac}} = \Phi^1(\ldots 1_i \ldots) = \phi_i(\boldsymbol{x}_1). \qquad (2.144)$$

The function $\delta(\boldsymbol{x}-\boldsymbol{x}_1)$ describes a one-electron state in which the electron is located at the point $\boldsymbol{x}$, since the probability of $\boldsymbol{x}_1$ being different from $\boldsymbol{x}$ in this state is zero. A similar result is obtained if $\boldsymbol{\psi}^\dagger(\boldsymbol{x})$ operates upon a determinantal function of any order (see Problem 2.6).

Now let us operate with $\boldsymbol{\psi}(\boldsymbol{x}')$ upon $\boldsymbol{\psi}^\dagger(\boldsymbol{x})\Phi_{\mathrm{vac}}$. Using the commutation relation (2.136), we obtain

$$\begin{aligned}\boldsymbol{\psi}(\boldsymbol{x}')\boldsymbol{\psi}^\dagger(\boldsymbol{x})\Phi_{\mathrm{vac}} &= \delta(\boldsymbol{x}-\boldsymbol{x}')\Phi_{\mathrm{vac}} - \boldsymbol{\psi}^\dagger(\boldsymbol{x})\boldsymbol{\psi}(\boldsymbol{x}')\Phi_{\mathrm{vac}} \\ &= \delta(\boldsymbol{x}-\boldsymbol{x}')\Phi_{\mathrm{vac}}, \qquad (2.145)\end{aligned}$$

since all destruction operators give zero when operating on Φ_{vac}. We have seen that $\boldsymbol{\psi}^\dagger(\boldsymbol{x})\Phi_{\mathrm{vac}}$ has a single electron at $\boldsymbol{x}$, and (2.145) shows that $\boldsymbol{\psi}(\boldsymbol{x}')$ operates on this to give zero if $\boldsymbol{x}' \neq \boldsymbol{x}$ and a constant† times Φ_{vac} when $\boldsymbol{x}' = \boldsymbol{x}$. We thus deduce that $\boldsymbol{\psi}(\boldsymbol{x})$ destroys the electron at $\boldsymbol{x}$ in the state $\boldsymbol{\psi}^\dagger(\boldsymbol{x})\Phi_{\mathrm{vac}}$.

In conclusion, it is perhaps worth mentioning that it is in connection with the field operators that the significance of the term *second quantization* emerges. These operators may be regarded as describing a *matter field*, whose quanta are the electrons, in the same way as photons are the quanta of the electromagnetic field. This *quantization of the matter field* is called *second* quantization, the quantization of the motion of the individual electrons being *first* quantization. We shall not pursue this concept, however, since it plays no part in the present work and would give rise to unnecessary confusion if described in detail.

---

† The constant is admittedly infinite, a difficulty which stems from the fact that the function $\delta(\boldsymbol{x}-\boldsymbol{x}_1)$ is not normalizable, since

$$\int |\delta(\boldsymbol{x}-\boldsymbol{x}_1)|^2 \mathrm{d}\boldsymbol{x}_1 = \delta(0),$$

which is infinite.

CHAPTER 3

THE HARTREE–FOCK METHOD AND THE FREE-ELECTRON GAS

3.1. The Hartree–Fock Method†

We have seen in ch. 1 that, if the electrons did not interact, so that the total Hamiltonian were

$$H_0 = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i)\right], \qquad (3.1)$$

the Schrödinger equation,

$$H_0\Psi = E\Psi, \qquad (3.2)$$

would be separable, and the eigenfunctions would be N-th-order determinants of one-electron functions whose orbital factors were eigenfunctions of the equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\boldsymbol{r})\psi = \epsilon\psi. \qquad (3.3)$$

However, the electrons *do* interact, the correct Hamiltonian being

$$H = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i)\right] + \tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N}\frac{e^2}{r_{ij}}, \qquad (3.4)$$

and the Schrödinger equation

$$H\Psi = E\Psi \qquad (3.5)$$

---

† There is an elementary account of the Hartree–Fock method and its application to a free-electron gas in WM, chs. 6 and 7.

is not separable. The foregoing does not apply, therefore. None the less, as we have already emphasized, we can always express the eigenfunctions of eq. (3.5) as infinite sums of the determinantal eigenfunctions of eq. (3.2).

This is a reasonable thing to do if the interaction term is to be treated as a perturbation, H_0 being the Hamiltonian of the unperturbed system, but it is not obligatory. The eigenfunctions of (3.5) can be expressed as infinite sums of determinants constructed from *any* set of orthogonal one-electron functions, not only those whose orbital factors satisfy (3.3), and it may be convenient to choose some other set.

For example, since a single determinant is easier to handle than an infinite sum of them, one is attracted to the idea of *approximating* to the ground-state wave function of the interacting system by a single determinant (the energy of the ground state is what will principally interest us). The problem then is to choose the one-electron functions so that this determinant gives the best approximation to the energy. According to the variation principle the best functions are, in fact, those which minimize the energy.

Let us rewrite the Hamiltonian as

$$H = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\nabla_i^2 + V(\boldsymbol{r}_i) + F(\boldsymbol{r}_i)\right] + \left[\tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N}\frac{e^2}{r_{ij}} - \sum_{i=1}^{N}F(\boldsymbol{r}_i)\right], \tag{3.6}$$

or

$$H = \sum_{i=1}^{N}\left[f(\boldsymbol{r}_i) + F(\boldsymbol{r}_i)\right] + \left[\tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N} v(\boldsymbol{r}_i, \boldsymbol{r}_j) - \sum_{i=1}^{N}F(\boldsymbol{r}_i)\right], \tag{3.7}$$

using our previous notation†, where the second term in square brackets is a modified interaction. We now take as an approximation to the ground-state wave function a single N-th-order determinant of one-electron functions whose orbital factors are eigenfunctions of the equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) + F(\boldsymbol{r})\right]\psi(\boldsymbol{r}) = \epsilon\psi(\boldsymbol{r}), \tag{3.8}$$

---

† Although previously we have written $f(\boldsymbol{x}_i)$, $v(\boldsymbol{x}_i, \boldsymbol{x}_j)$, to emphasize that the results applied generally to operators which might depend upon the spin, here we are dealing specifically with a Hamiltonian which is spin-independent, and so it will be less confusing simply to write $f(\boldsymbol{r}_i)$, $v(\boldsymbol{r}_i, \boldsymbol{r}_j)$.

the operator F being chosen so as to minimize the total energy. N functions have to be chosen, two of them, with opposite spins, corresponding to each of the lowest $\frac{1}{2}N$ eigenvalues (or, strictly speaking, *eigenfunctions*, since there will usually be degeneracy) of eq. (3.8). Let us designate these functions by $\phi_1, \phi_2, \ldots, \phi_N$.

This determinant is the unperturbed ground-state wave function when the second term in square brackets in eq. (3.6) is treated as a perturbation. The wave functions of the excited unperturbed states are also N-th-order determinants containing one or more functions ϕ_i formed from eigenfunctions of (3.8) with eigenvalues $\epsilon_i > \epsilon_{N/2}$.

We now intend to show that the best choice of F, according to the variation principle, is such that

$$\langle q|F|p\rangle = \sum_{i=1}^{N}\left[\langle iq|v|ip\rangle - \langle qi|v|ip\rangle\right]. \tag{3.9}$$

Suppose that Φ_0 is the ground-state unperturbed wave function, described above, and let Φ_q differ from it only in that the function $\phi_p (p \leqslant N)$ in Φ_0 has been replaced by $\phi_q (q > N)$; that is, Φ_q represents an unperturbed state in which a single electron has been excited from state ϕ_p to state ϕ_q. Also, for simplicity, we shall assume that ϕ_q occupies the same position in Φ_q as ϕ_p does in Φ_0.

We shall show first that the matrix element of H, eq. (3.7), between these two functions vanishes if F satisfies eq. (3.9). According to eq. (2.88) (or eq. (1.43)),

$$\left\langle \Phi_q \left| \sum_{i=1}^{N}\left[f(\boldsymbol{r}_i) + F(\boldsymbol{r}_i)\right] \right| \Phi_0 \right\rangle = \langle q|f+F|p\rangle = 0, \tag{3.10}$$

since $q \neq p$ and ϕ_q and ϕ_p are eigenfunctions of $f+F$ (that is, $f+F$ is diagonal with respect to the chosen one-electron functions). Also, according to eq. (2.104) (or eq. (1.50)),

$$\left\langle \Phi_q \left| \tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N} v(\boldsymbol{r}_i, \boldsymbol{r}_j) \right| \Phi_0 \right\rangle = \sum_{i=1}^{N}\left[\langle iq|v|ip\rangle - \langle qi|v|ip\rangle\right], \tag{3.11}$$

while, according to eqs. (2.88) and (3.9),

$$\left\langle \Phi_q \left| \sum_{i=1}^{N} F(\boldsymbol{r}_i) \right| \Phi_0 \right\rangle = \langle q|F|p\rangle = \sum_{i=1}^{N}\left[\langle iq|v|ip\rangle - \langle qi|v|ip\rangle\right]. \tag{3.12}$$

The matrix element of the second term of (3.7) in square brackets thus also vanishes, and we have, finally,

$$\langle \Phi_q | H | \Phi_0 \rangle = 0. \tag{3.13}$$

It should be noted that, so long as the functions ϕ_i are eigenfunctions of $f+F$, it is *necessary*, as well as sufficient, that F satisfy eq. (3.9) in order that (3.13) may follow. It is worth emphasizing that the *Hamiltonian* is not affected by the choice of F, but only the one-electron functions ϕ_i.

We have thus shown that a choice of F leading to eq. (3.9) causes the Hamiltonian to have zero matrix elements between the unperturbed ground-state determinant and determinants which differ from this in one row only. We shall now show that this choice of F gives the best single-determinantal ground-state wave function according to the variation principle[†].

We wish to choose those one-electron functions which make the value of

$$E = \frac{\langle \Phi_0 | H | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} \tag{3.14}$$

an absolute minimum. In other words, if any one of our chosen functions is altered, the value of this expression must increase. Let us make a change in one of the functions, say ϕ_p, appearing in Φ_0, by adding to it $\eta\phi_q (q > N)$, where η is a real constant. That is to say[‡], let us substitute $\Phi_0 + \eta\Phi_q$ for Φ_0 in (3.14), so that E becomes a function of η, thus:

$$\begin{aligned} E(\eta) &= \frac{\langle \Phi_0 + \eta\Phi_q | H | \Phi_0 + \eta\Phi_q \rangle}{\langle \Phi_0 + \eta\Phi_q | \Phi_0 + \eta\Phi_q \rangle} \\ &= \frac{\langle \Phi_0 | H | \Phi_0 \rangle + \eta \langle \Phi_0 | H | \Phi_q \rangle + \eta \langle \Phi_q | H | \Phi_0 \rangle + \eta^2 \langle \Phi_q | H | \Phi_q \rangle}{\langle \Phi_0 | \Phi_0 \rangle + \eta \langle \Phi_0 | \Phi_q \rangle + \eta \langle \Phi_q | \Phi_0 \rangle + \eta^2 \langle \Phi_q | \Phi_q \rangle} \\ &= [\langle \Phi_0 | H | \Phi_0 \rangle + \eta \{ \langle \Phi_0 | H | \Phi_q \rangle + \langle \Phi_q | H | \Phi_0 \rangle \} \\ &\quad + \eta^2 \langle \Phi_q | H | \Phi_q \rangle] (1+\eta^2)^{-1}, \end{aligned} \tag{3.15}$$

on the assumption that Φ_0 and Φ_q are orthogonal, as will be shown in the following section, as well as normalized. Differentiating with respect to η, and then putting $\eta = 0$, we find

$$\left[\frac{\mathrm{d}E(\eta)}{\mathrm{d}\eta} \right]_{\eta=0} = \langle \Phi_0 | H | \Phi_q \rangle + \langle \Phi_q | H | \Phi_0 \rangle, \tag{3.16}$$

and we have seen, eq. (3.13), that the matrix elements[†] vanish.

For our chosen one-electron functions, therefore, we have

$$\left[\frac{\mathrm{d}E(\eta)}{\mathrm{d}\eta} \right]_{\eta=0} = 0, \tag{3.17}$$

which proves that these functions minimize[‡] (3.14) and are consequently the best functions to use in the determinantal wave function according to the variation principle. The use of these functions in a single determinant as an approximation to the ground-state wave function is known as the Fock or Hartree–Fock method.

3.2. The One-electron Hartree–Fock Equation

In order to determine the explicit form of the Hartree–Fock equation (3.8), let us first write out eq. (3.9) in full, that is,

---

[†] The variational method used here is the elementary one described in WM, ch. 3.

[‡] In our previous notation

$$\begin{vmatrix} \phi_1(x_1) \ldots\ldots\ldots\ldots \phi_1(x_N) \\ \vdots \qquad\qquad \vdots \\ \phi_p(x_1) + \eta\phi_q(x_1) \ldots \phi_p(x_N) + \eta\phi_q(x_N) \\ \vdots \qquad\qquad \vdots \\ \phi_N(x_1) \ldots\ldots\ldots\ldots \phi_N(x_N) \end{vmatrix} = \begin{vmatrix} \phi_1(x_1) \ldots \phi_1(x_N) \\ \vdots \qquad \vdots \\ \phi_p(x_1) \ldots \phi_p(x_N) \\ \vdots \qquad \vdots \\ \phi_N(x_1) \ldots \phi_N(x_N) \end{vmatrix} + \eta \begin{vmatrix} \phi_1(x_1) \ldots \phi_1(x_N) \\ \vdots \qquad \vdots \\ \phi_q(x_1) \ldots \phi_q(x_N) \\ \vdots \qquad \vdots \\ \phi_N(x_1) \ldots \phi_N(x_N) \end{vmatrix}$$

$$= (N!)^{1/2} (\Phi_0 + \eta\Phi_q).$$

[†] Since H is Hermitian, we have

$$\langle \Phi_0 | H | \Phi_q \rangle = \langle \Phi_q | H | \Phi_0 \rangle^* = 0,$$

from (3.13).

[‡] Eq. (3.17), in fact, proves only that $E(\eta)$ has a *stationary value* at $\eta = 0$. That this is a minimum is suggested by the fact that the functions ϕ_p chosen for the ground state Φ_0 are those corresponding to the *lowest* eigenvalues ϵ_p of eq. (3.8). More rigorously, it can be shown that the second derivative of $E(\eta)$ is positive when $\eta = 0$ (see Problem 3.1).

$$
\begin{aligned}
&\int \phi_q^*(\boldsymbol{x})F(\boldsymbol{r})\phi_p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\
&\quad=\sum_{i=1}^{N}\Big[\iint \phi_i^*(\boldsymbol{x}_1)\phi_q^*(\boldsymbol{x}_2)v(\boldsymbol{r}_1,\boldsymbol{r}_2)\phi_i(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2\\
&\qquad-\iint \phi_q^*(\boldsymbol{x}_1)\phi_i^*(\boldsymbol{x}_2)v(\boldsymbol{r}_1,\boldsymbol{r}_2)\phi_i(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2\Big]\\
&\quad=\sum_{i=1}^{N}\Big[\int \phi_q^*(\boldsymbol{x}_2)\Big\{\int |\phi_i(\boldsymbol{x}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{x}_1\Big\}\phi_p(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_2\\
&\qquad-\int \phi_q^*(\boldsymbol{x}_2)\Big\{\int \phi_i^*(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_1)v(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{x}_1\Big\}\phi_i(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_2\\
&\quad=\int \phi_q^*(\boldsymbol{x})\sum_{i=1}^{N}\Big[\int |\phi_i(\boldsymbol{x}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r})\phi_p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_1\\
&\qquad-\int \phi_i^*(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\phi_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_1\Big]\mathrm{d}\boldsymbol{x},
\end{aligned}
\tag{3.18}
$$

where we have first interchanged $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in one integral and finally substituted $\boldsymbol{x}$ for $\boldsymbol{x}_2$ (this is perfectly legitimate as the value of a definite integral is independent of the symbols used for the variables of integration). We deduce that

$$
\begin{aligned}
F(\boldsymbol{r})\phi_p(\boldsymbol{x})=&\sum_{i=1}^{N}\int |\phi_i(\boldsymbol{x}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r})\phi_p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_1\\
&-\sum_{i=1}^{N}\int \phi_i^*(\boldsymbol{x}_1)\phi_p(\boldsymbol{x}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\phi_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}_1.
\end{aligned}
\tag{3.19}
$$

Remembering that ϕ_p is the product of ψ_p and a spin function, and that the integral implies a sum over the two values of the spin variable, we may eliminate the spin immediately and obtain

$$
\begin{aligned}
F(\boldsymbol{r})\psi_p(\boldsymbol{r})=&\sum_{i=1}^{N}\int |\psi_i(\boldsymbol{r}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r})\psi_p(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1\\
&-\sum_{\substack{i=1\\(\text{spin}\,i=\text{spin}\,p)}}^{N}\int \psi_i^*(\boldsymbol{r}_1)\psi_p(\boldsymbol{r}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\psi_i(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1.
\end{aligned}
\tag{3.20}
$$

In the latter term the sum is only over those states $\psi_i(i \leqslant N)$ which are associated with the same spin function as ψ_p, because the integral in the second term on the right-hand side of eq. (3.19) vanishes if ϕ_i and ϕ_p have different spin factors.

Eq. (3.20) can be simplified still further in the case of a *non-ferromagnetic system*, where there are equal numbers of electrons with opposite spins. We may then assume that, in the ground-state determinant, the set of functions $\phi_i(i \leqslant N)$ contains the set $\psi_i(i \leqslant \frac{1}{2}N)$ *twice* with opposite spin factors – in other words, the set $\psi_i(i \leqslant \frac{1}{2}N)$ and the set $\psi_i(\frac{1}{2}N < i \leqslant N)$ are the same. The equation thus becomes

$$
\begin{aligned}
F(\boldsymbol{r})\psi_p(\boldsymbol{r})=2&\sum_{i=1}^{N/2}\int |\psi_i(\boldsymbol{r}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r})\psi_p(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1\\
&-\sum_{i=1}^{N/2}\int \psi_i^*(\boldsymbol{r}_1)\psi_p(\boldsymbol{r}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\psi_i(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1,
\end{aligned}
\tag{3.21}
$$

and it is no longer necessary to specify spin $i=$ spin p, since the sum includes all the orbital states associated with either spin.

The Hartree–Fock equation for the function ψ_p, from eq. (3.8), is

$$
\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})+F(\boldsymbol{r})\right]\psi_p(\boldsymbol{r})=\epsilon_p\psi_p(\boldsymbol{r}),
\tag{3.22}
$$

and, using eq. (3.21), this becomes

$$
\begin{aligned}
&\left[-\frac{\hbar^2}{2m}\nabla^2+V(\boldsymbol{r})+2\sum_{i=1}^{N/2}\int |\psi_i(\boldsymbol{r}_1)|^2v(\boldsymbol{r}_1,\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1\right]\psi_p(\boldsymbol{r})\\
&\qquad-\sum_{i=1}^{N/2}\psi_i(\boldsymbol{r})\int \psi_i^*(\boldsymbol{r}_1)\psi_p(\boldsymbol{r}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1=\epsilon_p\psi_p(\boldsymbol{r}).
\end{aligned}
\tag{3.23}
$$

There is one such equation for each of the $\frac{1}{2}N$ different functions ψ_p appearing in the ground-state determinant[†]. Since the operator $F(\boldsymbol{r})$ depends upon all these functions, the set of $\frac{1}{2}N$ simultaneous equations must be solved by a process of iteration, which is continued until a sufficient degree of self-consistency is attained.

The last term on the left-hand side of eq. (3.23) is called the *exchange term*. Without this term the equation would reduce to the simpler *Hartree equation*, which is derived from a total wave function consisting of a single product, rather than a determinant, of one-electron functions. The

---

[†] We may assume that eq. (3.23) applies to *all* functions ψ_p, even though those with $p > N$ are not required in the ground state. This means that eq. (3.9) applies for *all* p and q, and not only for $p \leqslant N$, $q > N$, which is what we have previously required.

Hartree potential, the last term in square brackets, is just the potential energy of electron p due to its Coulomb interaction with the charge cloud of all the electrons. The operator $F(\boldsymbol{r})$ may be written as the sum

$$F(\boldsymbol{r}) = V_{\mathrm{C}}(\boldsymbol{r}) + F_{\mathrm{X}}(\boldsymbol{r}), \tag{3.24}$$

where V_{C} is the Hartree potential and F_{X} is the *exchange operator*, an integral operator defined by

$$F_{\mathrm{X}}(\boldsymbol{r})\psi_p(\boldsymbol{r}) = -\sum_{i=1}^{N/2} \psi_i(\boldsymbol{r}) \int \psi_i^*(\boldsymbol{r}_1)\psi_p(\boldsymbol{r}_1)v(\boldsymbol{r}_1, \boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1. \tag{3.25}$$

It is easy to see that V_{C} is Hermitian, since it is just a Coulomb potential, the same for all ψ_p, but it is not so easy to see that F_{X} is also Hermitian. However, the property of F defined in eq. (3.9) tells us immediately that *it* is Hermitian; thus, because v is real,

$$\begin{aligned}\langle q|F|p\rangle^* &= \sum_{i=1}^{N} [\langle iq|v|ip\rangle^* - \langle qi|v|ip\rangle^*] \\ &= \sum_{i=1}^{N} [\langle ip|v|iq\rangle - \langle ip|v|qi\rangle] \\ &= \sum_{i=1}^{N} [\langle ip|v|iq\rangle - \langle pi|v|iq\rangle] \\ &= \langle p|F|q\rangle,\end{aligned} \tag{3.26}$$

which is the condition for Hermiticity† (see Appendix I). Since F and V_{C} are Hermitian, it follows that F_{X} must also be.

This is an important result, for it means that, in eq. (3.22), the operator in square brackets, *which is the same for all* ψ_p, is Hermitian, so that the functions ψ_p may be assumed to form a complete, orthogonal set‡. It follows that the determinantal functions Φ_0 and Φ_q, for example, are orthogonal, as was assumed in the previous section.

---

† Strictly speaking, this proof of the Hermiticity of F is only rigorous if ϕ_p and ϕ_q are arbitrary functions or arbitrary members of a complete, orthogonal set of functions. The latter is certainly the case if F is Hermitian!

‡ This point was not brought out in the elementary treatment of the Hartree–Fock method in WM, ch. 6, where the expression of the exchange operator as a multiplying operator, like the Hartree potential, caused it apparently to depend upon ψ_p (present notation).

3.3. The Hamiltonian and the Hartree–Fock Ground-state Energy

If $c_j^\dagger$, c_j are the creation and destruction operators for the Hartree–Fock function ϕ_j, we know from the work of the previous chapter that, in the occupation number representation, the Hamiltonian (3.4) can be written as in eq. (2.117) or, using (3.7), in the form

$$\begin{aligned}H = &\sum_{i,j} \langle i|f+F|j\rangle c_i^\dagger c_j \\ &+ \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k - \sum_{i,j} \langle i|F|j\rangle c_i^\dagger c_j.\end{aligned} \tag{3.27}$$

However, ϕ_j (or its orbital factor) is an eigenfunction of eq. (3.22), that is,

$$(f+F)\phi_j = \epsilon_j\phi_j, \tag{3.28}$$

and we may assume that these eigenfunctions form a complete, orthonormal set (this is, in fact, necessary for the validity of the above expression for H). It follows that

$$\langle i|f+F|j\rangle = \epsilon_j\langle i|j\rangle = \epsilon_j\delta_{ij}. \tag{3.29}$$

Also, from eq. (3.9), we have

$$\begin{aligned}\langle i|F|j\rangle &= \sum_{k=1}^{N} [\langle ki|v|kj\rangle - \langle ik|v|kj\rangle] \\ &= \sum_{k} [\langle ki|v|kj\rangle - \langle ik|v|kj\rangle]n_k,\end{aligned} \tag{3.30}$$

where n_k is the occupation number of the function ϕ_k *in the ground-state determinant* Φ_0. Eq. (3.27) thus becomes

$$\begin{aligned}H = &\sum_i \epsilon_i c_i^\dagger c_i + \tfrac{1}{2} \sum_{i,j,k,l} \langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k \\ &- \sum_{i,j,k} [\langle ki|v|kj\rangle - \langle ik|v|kj\rangle]n_k c_i^\dagger c_j.\end{aligned} \tag{3.31}$$

The ground-state energy in the Hartree–Fock approximation is, therefore,

$$\begin{aligned}E &= \langle\Phi_0|H|\Phi_0\rangle \\ &= \sum_{i=1}^{N} \epsilon_i + \langle\Phi_0|H'|\Phi_0\rangle \\ &\quad - \sum_{i,j,k} [\langle ki|v|kj\rangle - \langle ik|v|kj\rangle]\langle\Phi_0|n_k c_i^\dagger c_j|\Phi_0\rangle,\end{aligned} \tag{3.32}$$

where H' is the second term on the right-hand side of eq. (3.31). From eq. (2.95), we have

$$\begin{aligned}\langle\Phi_0|H'|\Phi_0\rangle &= \tfrac{1}{2}\sum_{i,j} n_i n_j[\langle ij|v|ij\rangle - \langle ij|v|ji\rangle] \\ &= \tfrac{1}{2}\sum_{i,j=1}^{N} [\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \qquad (3.33)\end{aligned}$$

Also,

$$\langle\Phi_0|n_k c_i^\dagger c_j|\Phi_0\rangle = \begin{cases}\delta_{ij}, & \text{for } i,j,k \leqslant N,\\ 0, & \text{otherwise,}\end{cases}$$

so that the last term on the right-hand side of eq. (3.32) becomes

$$-\sum_{i,k=1}^{N} [\langle ki|v|ki\rangle - \langle ik|v|ki\rangle] = -\sum_{i,j=1}^{N} [\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \qquad (3.34)$$

Substituting (3.33) and (3.34) in (3.32) gives, finally,

$$E = \sum_{i=1}^{N} \epsilon_i - \tfrac{1}{2}\sum_{i,j=1}^{N} [\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \qquad (3.35)$$

This expression can be put in another form which is sometimes useful. From eq. (3.28), we have

$$\epsilon_i = \langle i|f|i\rangle + \langle i|F|i\rangle, \qquad (3.36)$$

and eq. (3.9) gives

$$\langle i|F|i\rangle = \sum_{j=1}^{N} [\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \qquad (3.37)$$

Hence, eq. (3.35) may be written

$$E = \tfrac{1}{2}\sum_{i=1}^{N} [\epsilon_i + \langle i|f|i\rangle]. \qquad (3.38)$$

3.4. The Free-electron Gas

The only realistic model of a many-electron system for which the Hartree–Fock equation can be solved analytically is the free-electron gas, which approximates to the valence-electron gas in many metals. A system of electrons moving under no external forces is called a free-electron gas, even though the electrons are interacting with each other. It is customary to assume the presence of a uniform distribution of positive charge of density equal in magnitude, but opposite in sign, to the average charge density of the electrons, so that the system as a whole is electrically neutral.

Let us assume that we have N electrons in a cubic box, with edges of length L and volume $\Omega = L^3$, which also contains a distribution of positive charge of uniform density Ne/Ω. The eigenfunctions of the Hartree–Fock equation for this system (see WM, ch. 7), subject to periodic boundary conditions, are the functions

$$\phi_{k\sigma}(\boldsymbol{x}) = \psi_k(\boldsymbol{r})\chi_\sigma(\zeta), \qquad (3.39)$$

where the spin factor $\chi_\sigma(\zeta)$ is† $\alpha(\zeta)$ when $\sigma = \frac{1}{2}$ and $\beta(\zeta)$ when $\sigma = -\frac{1}{2}$, and the normalized function ψ_k is

$$\psi_k(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega}}\exp(i\boldsymbol{k}\cdot\boldsymbol{r}), \qquad (3.40)$$

the *wave vector* $\boldsymbol{k}$ being given by

$$\boldsymbol{k} = \frac{2\pi}{L}(n_1\boldsymbol{e}_1 + n_2\boldsymbol{e}_2 + n_3\boldsymbol{e}_3), \qquad (3.41)$$

where n_1, n_2, n_3 are integers, positive, negative or zero, and $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ are unit vectors along three mutually orthogonal cube edges.

The wave vectors of the occupied one-electron states in the Hartree–Fock ground-state determinant lie within a sphere in '$\boldsymbol{k}$-space', called the *Fermi sphere*, of radius‡ k_F. Each orbital state is doubly-occupied by electrons with opposite spins. A sum over the different *orbital* states in the ground-state determinant is thus a sum over all $\boldsymbol{k}$ vectors lying within the Fermi sphere. The Hartree–Fock eq. (3.23) therefore becomes

$$\begin{aligned}&\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) + 2\sum_{k'}\int |\psi_{k'}(\boldsymbol{r})|^2 v(\boldsymbol{r}_1,\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1\right]\psi_k(\boldsymbol{r}) \\ &\qquad -\sum_{k'}\psi_{k'}(\boldsymbol{r})\int \psi_{k'}^*(\boldsymbol{r}_1)\psi_k(\boldsymbol{r}_1)v(\boldsymbol{r}_1,\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}_1 = \epsilon_k\psi_k(\boldsymbol{r}), \qquad (3.42)\end{aligned}$$

where

$$v(\boldsymbol{r}_1,\boldsymbol{r}) = \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}|}.$$

Now $V(\boldsymbol{r})$ is the potential energy of an electron at $\boldsymbol{r}$ due to the uniform

---

† We use the symbol σ to denote the spin quantum number, which is more usual in this work than the symbol m_s used in WM, ch. 5.

‡ This was denoted by k_0 in WM.

background of positive charge, and this is exactly cancelled by the Hartree potential (the last term in square brackets), which is the potential energy of an electron at $\boldsymbol{r}$ due to a uniform distribution of *negative* charge, with density $-Ne/\Omega$, due to *all* the electrons, since, on the assumption of (3.40),

$$|\psi_{\boldsymbol{k}}(\boldsymbol{r})|^2 = 1/\Omega, \tag{3.43}$$

for all $\boldsymbol{k}$. Eq. (3.42) thus reduces to

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + F_{\mathrm{X}}(\boldsymbol{r})\right]\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \epsilon_{\boldsymbol{k}}\psi_{\boldsymbol{k}}(\boldsymbol{r}), \tag{3.44}$$

where $F_{\mathrm{X}}(\boldsymbol{r})$ is the exchange operator defined in eq. (3.25). It was shown in WM, §7.5, that the function $\psi_{\boldsymbol{k}}$ given in (3.40) is a self-consistent eigenfunction of this equation, and the eigenvalue is

$$\epsilon_k = \frac{\hbar^2k^2}{2m} - \frac{e^2k_{\mathrm{F}}}{2\pi}\left[2 + \frac{k_{\mathrm{F}}^2 - k^2}{kk_{\mathrm{F}}}\log\left|\frac{k_{\mathrm{F}}+k}{k_{\mathrm{F}}-k}\right|\right]. \tag{3.45}$$

It is easy to see that the same $\psi_{\boldsymbol{k}}$ is also an eigenfunction of the simpler Hartree equation (that is, (3.44) without $F_{\mathrm{X}}(\boldsymbol{r})$), and that the eigenvalue in this case is just $\hbar^2k^2/2m$. Since we shall be concerned only with the *total* energy of a free-electron gas, and not with the distribution of one-electron energies, we need not refer to the expression (3.45) again, but shall simply use the one-electron functions (3.39) as the basis set from which the determinants for this system will be constructed. That this is an orthonormal set is easily verified, for

$$\int \phi^*_{\boldsymbol{k}'\sigma'}(\boldsymbol{x})\phi_{\boldsymbol{k}\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{\zeta=\pm 1}\chi_{\sigma'}(\zeta)\chi_{\sigma}(\zeta)\int \psi^*_{\boldsymbol{k}'}(\boldsymbol{r})\psi_{\boldsymbol{k}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$$

$$= \frac{\delta_{\sigma\sigma'}}{\Omega}\int \exp\,[i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}]\,\mathrm{d}\boldsymbol{r} = \delta_{\sigma\sigma'}\delta_{\boldsymbol{k}\boldsymbol{k}'}. \tag{3.46}$$

The Hamiltonian of this system, in the notation of ch. 1, is essentially

$$H = H_0 + H'$$

$$= -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N}\frac{e^2}{r_{ij}},$$

the first term representing the kinetic energy and the second the Coulomb interaction of the electrons. We have not explicitly included terms representing the interaction of the electrons with the uniform distribution of positive charge and the self-energy of this charge, but these terms will be cancelled out of our final Hamiltonian.

We denote by $c_{\boldsymbol{k}\sigma}$ and $c^{\dagger}_{\boldsymbol{k}\sigma}$ the destruction and creation operators for the state $\phi_{\boldsymbol{k}\sigma}$, and, with

$$f(\boldsymbol{r}_i) = -\frac{\hbar^2}{2m}\nabla_i^2 \tag{3.47}$$

and

$$v(\boldsymbol{r}_i, \boldsymbol{r}_j) = \frac{e^2}{r_{ij}}, \tag{3.48}$$

the Hamiltonian expressed in the occupation number formalism becomes, from (2.117),

$$\begin{aligned} H = &\sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \sigma_1,\sigma_2}}\langle \boldsymbol{k}_1\sigma_1|f|\boldsymbol{k}_2\sigma_2\rangle c^{\dagger}_{\boldsymbol{k}_1\sigma_1}c_{\boldsymbol{k}_2\sigma_2}\\ &+\tfrac{1}{2}\sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4\\ \sigma_1,\sigma_2,\sigma_3,\sigma_4}}\langle \boldsymbol{k}_1\sigma_1,\boldsymbol{k}_2\sigma_2|v|\boldsymbol{k}_3\sigma_3,\boldsymbol{k}_4\sigma_4\rangle c^{\dagger}_{\boldsymbol{k}_1\sigma_1}c^{\dagger}_{\boldsymbol{k}_2\sigma_2}c_{\boldsymbol{k}_4\sigma_4}c_{\boldsymbol{k}_3\sigma_3}, \end{aligned} \tag{3.49}$$

where the sums over $\boldsymbol{k}_1$, etc., are over all the wave vectors (3.41) and the sums over σ_1, etc., are over the two values $\pm\frac{1}{2}$.

3.4.1. *The kinetic energy term*

Expression (3.49) simplifies somewhat when the matrix elements are evaluated. Let us consider the kinetic energy term first. We have

$$\begin{aligned} \langle \boldsymbol{k}_1\sigma_1|f|\boldsymbol{k}_2\sigma_2\rangle &= \int \phi^*_{\boldsymbol{k}_1\sigma_1}(\boldsymbol{x})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\phi_{\boldsymbol{k}_2\sigma_2}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \frac{1}{\Omega}\sum_{\zeta=\pm 1}\chi_{\sigma_1}(\zeta)\chi_{\sigma_2}(\zeta)\int \exp\,(-\mathrm{i}\boldsymbol{k}_1\cdot\boldsymbol{r})\\ &\quad\times\left(-\frac{\hbar^2}{2m}\nabla^2\right)\exp\,(\mathrm{i}\boldsymbol{k}_2\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\\ &= -\frac{\hbar^2}{2m\Omega}\delta_{\sigma_1\sigma_2}\int \exp\,(-\mathrm{i}\boldsymbol{k}_1\cdot\boldsymbol{r})\,\nabla^2\exp\,(i\boldsymbol{k}_2\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\\ &= \frac{\hbar^2k_2^2}{2m\Omega}\delta_{\sigma_1\sigma_2}\int \exp\,[\mathrm{i}(\boldsymbol{k}_2-\boldsymbol{k}_1)\cdot\boldsymbol{r}]\,\mathrm{d}\boldsymbol{r}\\ &= \frac{\hbar^2k_2^2}{2m}\delta_{\sigma_1\sigma_2}\delta_{\boldsymbol{k}_1\boldsymbol{k}_2}. \end{aligned} \tag{3.50}$$

The matrix element therefore vanishes unless $\sigma_1 = \sigma_2$ and $\boldsymbol{k}_1 = \boldsymbol{k}_2$, when its value is $\hbar^2 k_2^2/2m$. The kinetic energy term thus reduces to

$$H_0 = \sum_{\boldsymbol{k},\sigma} \frac{\hbar^2 k^2}{2m} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma}, \tag{3.51}$$

where $c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma}$ is just the number operator for the state $\phi_{\boldsymbol{k}\sigma}$.

3.4.2. *The interaction term*

The matrix element in H' is

$$\begin{aligned}
&\langle \boldsymbol{k}_1\sigma_1, \boldsymbol{k}_2\sigma_2 | v | \boldsymbol{k}_3\sigma_3, \boldsymbol{k}_4\sigma_4 \rangle \\
&\quad = \iint \phi_{\boldsymbol{k}_1\sigma_1}^*(\boldsymbol{x}_1)\phi_{\boldsymbol{k}_2\sigma_2}^*(\boldsymbol{x}_2)\frac{e^2}{r_{12}}\phi_{\boldsymbol{k}_3\sigma_3}(\boldsymbol{x}_1)\phi_{\boldsymbol{k}_4\sigma_4}(\boldsymbol{x}_2)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\
&\quad = \sum_{\zeta_1,\zeta_2=\pm 1} \chi_{\sigma_1}(\zeta_1)\chi_{\sigma_2}(\zeta_2)\chi_{\sigma_3}(\zeta_1)\chi_{\sigma_4}(\zeta_2) \\
&\qquad \times \iint \psi_{\boldsymbol{k}_1}^*(\boldsymbol{r}_1)\psi_{\boldsymbol{k}_2}^*(\boldsymbol{r}_2)\frac{e^2}{r_{12}}\psi_{\boldsymbol{k}_3}(\boldsymbol{r}_1)\psi_{\boldsymbol{k}_4}(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2 \\
&\quad = \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4}\frac{e^2}{\Omega^2}\iint \frac{\exp[\mathrm{i}(\boldsymbol{k}_3-\boldsymbol{k}_1)\cdot\boldsymbol{r}_1]\exp[\mathrm{i}(\boldsymbol{k}_4-\boldsymbol{k}_2)\cdot\boldsymbol{r}_2]}{r_{12}}\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2.
\end{aligned} \tag{3.52}$$

The integral may be written

$$I = \int \exp[\mathrm{i}(\boldsymbol{k}_3-\boldsymbol{k}_1)\cdot\boldsymbol{r}_1]\,\mathrm{d}\boldsymbol{r}_1 \int \frac{\exp[\mathrm{i}(\boldsymbol{k}_4-\boldsymbol{k}_2)\cdot\boldsymbol{r}_2]}{r_{12}}\,\mathrm{d}\boldsymbol{r}_2,$$

and (see WM, p. 170)

$$\int \frac{\exp[\mathrm{i}(\boldsymbol{k}_4-\boldsymbol{k}_2)\cdot\boldsymbol{r}_2]}{r_{12}}\,\mathrm{d}\boldsymbol{r}_2 = \frac{4\pi\exp[\mathrm{i}(\boldsymbol{k}_4-\boldsymbol{k}_2)\cdot\boldsymbol{r}_1]}{|\boldsymbol{k}_4-\boldsymbol{k}_2|^2}, \tag{3.53}$$

provided $\boldsymbol{k}_4 \neq \boldsymbol{k}_2$.

Hence

$$\begin{aligned}
I &= \frac{4\pi}{|\boldsymbol{k}_4-\boldsymbol{k}_2|^2}\int \exp[\mathrm{i}(\boldsymbol{k}_3-\boldsymbol{k}_1+\boldsymbol{k}_4-\boldsymbol{k}_2)\cdot\boldsymbol{r}_1]\,\mathrm{d}\boldsymbol{r}_1 \\
&= \begin{cases} \dfrac{4\pi\Omega}{|\boldsymbol{k}_4-\boldsymbol{k}_2|^2}, & \text{if} \quad \boldsymbol{k}_1-\boldsymbol{k}_3 = \boldsymbol{k}_4-\boldsymbol{k}_2 \neq \boldsymbol{0}, \\ 0, & \text{if} \quad \boldsymbol{k}_1-\boldsymbol{k}_3 \neq \boldsymbol{k}_4-\boldsymbol{k}_2. \end{cases}
\end{aligned} \tag{3.54}$$

If we write

$$\boldsymbol{k}_1-\boldsymbol{k}_3 = \boldsymbol{k}_4-\boldsymbol{k}_2 = \boldsymbol{k} \neq \boldsymbol{0},$$

then

$$I = \frac{4\pi\Omega}{k^2}. \tag{3.55}$$

If $\boldsymbol{k} = \boldsymbol{0}$, however,

$$I = \int \mathrm{d}\boldsymbol{r}_1 \int \frac{\mathrm{d}\boldsymbol{r}_2}{r_{12}}. \tag{3.56}$$

The only non-zero matrix elements, with $\boldsymbol{k} \neq \boldsymbol{0}$, are those of the type

$$\langle \boldsymbol{k}_3+\boldsymbol{k}, \sigma_1; \boldsymbol{k}_4-\boldsymbol{k}, \sigma_2 | v | \boldsymbol{k}_3\sigma_1, \boldsymbol{k}_4\sigma_2 \rangle = \frac{4\pi e^2}{\Omega k^2}. \tag{3.57}$$

When $\boldsymbol{k} = 0$, we have

$$\langle \boldsymbol{k}_3\sigma_1, \boldsymbol{k}_4\sigma_2 | v | \boldsymbol{k}_3\sigma_1, \boldsymbol{k}_4\sigma_2 \rangle = \frac{e^2}{\Omega^2}\int \mathrm{d}\boldsymbol{r}_1 \int \frac{\mathrm{d}\boldsymbol{r}_2}{r_{12}}. \tag{3.58}$$

This expression represents *twice* the self-potential energy of one *electronic* charge uniformly distributed throughout the volume Ω (there is, of course, a factor $\frac{1}{2}$ in H'). However, we have assumed, in order to ensure the charge neutrality of the system as a whole, that to each electron there is a single *protonic* charge uniformly distributed throughout the volume Ω. The effect of this positive charge distribution is to cancel the effect of those terms in the Hamiltonian with $\boldsymbol{k} = \boldsymbol{0}$. This will be seen more clearly below, but meanwhile we shall drop those terms with $\boldsymbol{k} = \boldsymbol{0}$.

Putting $\boldsymbol{k}_3 = \boldsymbol{p}$, $\boldsymbol{k}_4 = \boldsymbol{q}$, $\sigma_1 = \sigma$, $\sigma_2 = \sigma'$, the interaction term becomes

$$\begin{aligned}
H' &= \tfrac{1}{2}\sum_{\substack{\boldsymbol{k},\boldsymbol{p},\boldsymbol{q}\\ \sigma,\sigma'}} \langle \boldsymbol{p}+\boldsymbol{k}, \sigma; \boldsymbol{q}-\boldsymbol{k}, \sigma' | v | \boldsymbol{p}\sigma, \boldsymbol{q}\sigma' \rangle c_{\boldsymbol{p}+\boldsymbol{k},\sigma}^\dagger c_{\boldsymbol{q}-\boldsymbol{k},\sigma'}^\dagger c_{\boldsymbol{q}\sigma'} c_{\boldsymbol{p}\sigma} \\
&= \sum_{\substack{\boldsymbol{k},\boldsymbol{p},\boldsymbol{q}\\ \sigma,\sigma'}} \frac{2\pi e^2}{\Omega k^2} c_{\boldsymbol{p}+\boldsymbol{k},\sigma}^\dagger c_{\boldsymbol{q}-\boldsymbol{k},\sigma'}^\dagger c_{\boldsymbol{q}\sigma'} c_{\boldsymbol{p}\sigma},
\end{aligned} \tag{3.59}$$

where the sums are over all the wave vectors (3.41) except $\boldsymbol{k} = \boldsymbol{0}$.

3.4.3. *Alternative treatment of the interaction term*

Another treatment of the interaction term, which will be used in the following chapter for the modified interaction considered there, consists in first expanding $1/r_{ij}$ in Fourier series within the cubic box of volume Ω (see WM, p. 285). We then find

$$H' = \sum_{i,j\neq i}^{N} {\sum_{\boldsymbol{k}}}' \frac{2\pi e^2}{\Omega k^2}\exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)], \tag{3.60}$$

where the prime denotes that $\boldsymbol{k} = \boldsymbol{0}$ is excluded to account for the uniform background of positive charge. We may thus write

$$v(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_{\boldsymbol{k}}{}' \frac{4\pi e^2}{\Omega k^2} \exp\left[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)\right]. \tag{3.61}$$

and, using (3.52),

$$\begin{aligned}
&\langle \boldsymbol{k}_1\sigma_1, \boldsymbol{k}_2\sigma_2 | v | \boldsymbol{k}_3\sigma_3, \boldsymbol{k}_4\sigma_4\rangle \\
&\quad = \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4} \sum_{\boldsymbol{k}}{}' \frac{4\pi e^2}{\Omega^3 k^2} \iint \exp\left[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)\right] \\
&\qquad \times \exp\left[\mathrm{i}(\boldsymbol{k}_3 - \boldsymbol{k}_4)\cdot\boldsymbol{r}_1\right] \exp\left[\mathrm{i}(\boldsymbol{k}_4 - \boldsymbol{k}_2)\cdot\boldsymbol{r}_2\right] \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 \\
&\quad = \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4} \sum_{\boldsymbol{k}}{}' \frac{4\pi e^2}{\Omega^3 k^2} \int \exp\left[\mathrm{i}\boldsymbol{r}_1\cdot(\boldsymbol{k} - \boldsymbol{k}_1 + \boldsymbol{k}_3)\right] \mathrm{d}\boldsymbol{r}_1 \\
&\qquad \times \int \exp\left[\mathrm{i}\boldsymbol{r}_2\cdot(-\boldsymbol{k} - \boldsymbol{k}_2 + \boldsymbol{k}_4)\right] \mathrm{d}\boldsymbol{r}_2 \\
&\quad = \begin{cases} \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4} \dfrac{4\pi e^2}{\Omega k^2}, & \text{if} \quad \boldsymbol{k}_1 - \boldsymbol{k}_3 = \boldsymbol{k}_2 - \boldsymbol{k}_4 = \boldsymbol{k} \neq \boldsymbol{0}, \\ 0, & \text{if} \quad \boldsymbol{k}_1 - \boldsymbol{k}_3 \neq \boldsymbol{k}_2 - \boldsymbol{k}_4. \end{cases}
\end{aligned} \tag{3.62}$$

This leads immediately to the expression (3.59) for H', and $\boldsymbol{k} = \boldsymbol{0}$ is automatically excluded, because it is excluded from (3.60).

3.5. First-order Perturbation Theory for a Free-electron Gas

It should be clear from what has gone before that, for a free-electron gas, the Hartree–Fock ground-state energy is precisely the same as that given by first-order perturbation theory when the whole Coulomb interaction is treated as a perturbation. In the latter case the unperturbed Hamiltonian consists of the kinetic energy term H_0 only, and the wave function Φ_0 of this unperturbed ground state is a single determinant of free-electron functions with $\boldsymbol{k}$-vectors lying within the Fermi sphere – in other words, the Hartree–Fock ground-state wave function.

The unperturbed ground-state energy is just the kinetic energy

$$\begin{aligned}
\langle \Phi_0 | H_0 | \Phi_0 \rangle &= \sum_{\boldsymbol{k},\sigma} \frac{\hbar^2 k^2}{2m} \langle \Phi_0 | c^{\dagger}_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} | \Phi_0 \rangle \\
&= 2 \sum_{\substack{\boldsymbol{k} \\ (k<k_{\mathrm{F}})}} \frac{\hbar^2 k^2}{2m},
\end{aligned} \tag{3.63}$$

since $c_{\boldsymbol{k}\sigma}\Phi_0 = 0$ if $k > k_{\mathrm{F}}$, and each orbital state $\boldsymbol{k}$ is occupied by two electrons with opposite spins. Replacing the sum by an integral, remembering that the density of orbital states in $\boldsymbol{k}$-space is $\Omega/8\pi^3$, we have

$$\begin{aligned}
\langle \Phi_0 | H_0 | \Phi_0 \rangle &= \frac{\Omega}{4\pi^3} \int_0^{k_{\mathrm{F}}} \frac{\hbar^2 k^2}{2m} 4\pi k^2 \,\mathrm{d}k \\
&= \frac{\hbar^2 \Omega}{10\pi^2 m} k_{\mathrm{F}}^5.
\end{aligned} \tag{3.64}$$

Now, twice the number of orbital states within the Fermi sphere is equal to N, the number of electrons, so that

$$\frac{\Omega}{4\pi^3}\frac{4\pi}{3} k_{\mathrm{F}}^3 = N,$$

and

$$k_{\mathrm{F}}^3 = \frac{3\pi^2 N}{\Omega}. \tag{3.65}$$

Finally, then,

$$\langle \Phi_0 | H_0 | \Phi_0 \rangle = \frac{\hbar^2 \Omega}{10\pi^2 m} \left(\frac{3\pi^2 N}{\Omega} \right)^{5/3} = \frac{3\hbar^2}{10m} \left(\frac{9\pi}{4} \right)^{2/3} \frac{N}{r_{\mathrm{s}}^2}, \tag{3.66}$$

the *atomic radius* r_{s} being defined by

$$\frac{4\pi}{3} r_{\mathrm{s}}^3 = \frac{\Omega}{N}. \tag{3.67}$$

The mean kinetic energy per electron is called the *Fermi energy*, and we denote it by E_{F}. If energy is measured in rydbergs and length in Bohr units[†], we obtain from (3.66)

$$E_{\mathrm{F}} = \frac{2.21}{r_{\mathrm{s}}^2} \text{ ry}. \tag{3.68}$$

The first-order perturbation correction due to the interaction term H' is

$$\begin{aligned}
\Delta E^{(1)} &= \langle \Phi_0 | H' | \Phi_0 \rangle \\
&= \sum_{\substack{\boldsymbol{k},\boldsymbol{p},\boldsymbol{q} \\ \sigma,\sigma'}} \frac{2\pi e^2}{\Omega k^2} \langle \Phi_0 | c^{\dagger}_{\boldsymbol{p}+\boldsymbol{k},\sigma} c^{\dagger}_{\boldsymbol{q}-\boldsymbol{k},\sigma'} c_{\boldsymbol{q}\sigma'} c_{\boldsymbol{p}\sigma} | \Phi_0 \rangle,
\end{aligned} \tag{3.69}$$

from (1.78) and (3.59). Now,

$$c_{\boldsymbol{p}\sigma}\Phi_0 = 0 \quad \text{unless} \quad p < k_{\mathrm{F}}, \tag{3.70}$$

[†] 1 rydberg = $me^4/2\hbar^2$ = 13.60 eV,
1 Bohr unit = $\hbar^2/me^2$ = 0.5292 Å.

$$c_{q\sigma'}\Phi_0 = 0 \quad \text{unless} \quad q < k_{\mathrm{F}}, \tag{3.71}$$

and the matrix element vanishes unless

$$c^{\dagger}_{p+k,\sigma}c^{\dagger}_{q-k,\sigma'}c_{q\sigma'}c_{p\sigma}\Phi_0 = \pm\Phi_0, \tag{3.72}$$

or, using the commutation relations,

$$-c^{\dagger}_{p+k,\sigma}c_{q\sigma'}c^{\dagger}_{q-k,\sigma'}c_{p\sigma}\Phi_0 = \pm\Phi_0. \tag{3.73}$$

Since $\boldsymbol{k} = \mathbf{0}$ is excluded from the sum in (3.69), it follows that, with $p,q < k_{\mathrm{F}}$, the matrix element vanishes unless

$$\boldsymbol{p} = \boldsymbol{q} - \boldsymbol{k}, \quad \boldsymbol{q} = \boldsymbol{p} + \boldsymbol{k}, \quad \sigma = \sigma',$$

so that the left-hand side of (3.73) becomes

$$-c^{\dagger}_{q\sigma}c_{q\sigma}c^{\dagger}_{p\sigma}c_{p\sigma}\Phi_0 = -\Phi_0, \tag{3.74}$$

and the matrix element has the value -1. Thus

$$\Delta E^{(1)} = -\frac{4\pi e^2}{\Omega}\sum_{\substack{p,q \neq p \\ (p,q < k_{\mathrm{F}})}} \frac{1}{|\boldsymbol{p}-\boldsymbol{q}|^2}, \tag{3.75}$$

an extra factor 2 arising because of the sum over σ, each orbital state being doubly occupied. It should be noted that, since $\boldsymbol{p} = \boldsymbol{q} - \boldsymbol{k}$ and $\boldsymbol{k} \neq \mathbf{0}$, we never have $\boldsymbol{p} = \boldsymbol{q}$, so that all the terms are finite. Replacing the sums by integrals, we find

$$\Delta E^{(1)} = -\frac{4\pi e^2}{\Omega}\left(\frac{\Omega}{8\pi^3}\right)^2 \iint \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}}{|\boldsymbol{p}-\boldsymbol{q}|^2}, \tag{3.76}$$

both integrals being throughout the Fermi sphere. The integration is elementary but tedious (see WM, ch. 7, especially eqs. (7.43) and (7.49)) and gives

$$\Delta E^{(1)} = -\frac{e^2\Omega}{4\pi^3}k_{\mathrm{F}}^4 = -\frac{e^2\Omega}{4\pi^3}\left(\frac{3\pi^2 N}{\Omega}\right)^{4/3}$$

$$= -\frac{3e^2N}{4}\left(\frac{9}{4\pi^2}\right)^{1/3}\frac{1}{r_{\mathrm{s}}}. \tag{3.77}$$

This is called the *exchange energy*.

Denoting the *average* exchange energy per electron by E_{X}, and expressing energy in rydbergs and length in Bohr units, we have

$$E_{\mathrm{X}} = -\frac{0.916}{r_{\mathrm{s}}}\,\text{ry}. \tag{3.78}$$

The average ground-state energy per electron of a free-electron gas, as given by first-order perturbation theory (or by the Hartree–Fock method, which is the reason for the symbol E_{HF}), is thus

$$E_{\mathrm{HF}} = E_{\mathrm{F}} + E_{\mathrm{X}} = \frac{2.21}{r_{\mathrm{s}}^2} - \frac{0.916}{r_{\mathrm{s}}}\,\text{ry}. \tag{3.79}$$

3.6. Second-order Perturbation Theory for a Free-electron Gas

According to the Rayleigh–Schrödinger perturbation theory, the second-order correction $\Delta E^{(2)}$ to the energy of the ground state of a free-electron gas, due to the interaction term, is given by (eq. (1.87))

$$\Delta E^{(2)} = -\sum_{n=1}^{\infty} \frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{E_n - E_0}, \tag{3.80}$$

where E_0 and E_n are the kinetic energies of the ground state Φ_0 and the excited state Φ_n of the unperturbed system, and H' is given by (3.59). Thus

$$\langle\Phi_n|H'|\Phi_0\rangle = \sum_{\substack{k,p,q \\ \sigma,\sigma'}} \frac{2\pi e^2}{\Omega k^2}\langle\Phi_n|c^{\dagger}_{p+k,\sigma}c^{\dagger}_{q-k,\sigma'}c_{q\sigma'}c_{p\sigma}|\Phi_0\rangle. \tag{3.81}$$

Since the one-electron functions, which are the elements of Φ_0 and Φ_n, are eigenfunctions of the Hartree–Fock equations for this system, we may deduce immediately (see eq. (3.13) and relevant discussion) that the matrix element vanishes if Φ_n and Φ_0 differ in only one row (that is, in a single one-electron function).

It is also clear that the matrix element vanishes if Φ_n and Φ_0 differ in more than two rows, since two destruction and two creation operators could not then operate upon Φ_0 to give $\pm\Phi_n$. Thus, if (3.81) is not to vanish, Φ_n and Φ_0 *must differ in two rows and only two rows*. Let us consider this case.

We shall suppose that Φ_n is obtained from Φ_0 by exciting two electrons, initially in states $\boldsymbol{k}_1\sigma_1$, $\boldsymbol{k}_2\sigma_2$, to states $\boldsymbol{k}_3\sigma_3$, $\boldsymbol{k}_4\sigma_4$ lying outside the Fermi sphere. That is to say, Φ_n is

$$\Phi_n(\ldots 0_{k_1\sigma_1}\ldots 0_{k_2\sigma_2}\ldots 1_{k_3\sigma_3}\ldots 1_{k_4\sigma_4}\ldots),$$

the occupation numbers of the other states being the same as in Φ_0. We then find

$$\Delta E^{(2)} = -\tfrac{1}{4}\sum_{\substack{k_1k_2k_3k_4 \\ \sigma_1\sigma_2\sigma_3\sigma_4}} \frac{\left|\sum_{kpq\sigma\sigma'}(2\pi e^2/\Omega k^2)\,\langle\Phi_n|c^{\dagger}_{p+k,\sigma}c^{\dagger}_{q-k,\sigma'}c_{q\sigma'}c_{p\sigma}|\Phi_0\rangle\right|^2}{(\hbar^2/2m)(k_3^2+k_4^2-k_1^2-k_2^2)}, \tag{3.82}$$

where $k_1, k_2 < k_F$ and $k_3, k_4 > k_F$. The factor $\frac{1}{4}$ is necessary, because Φ_n is unchanged if either $\boldsymbol{k}_1\sigma_1$, $\boldsymbol{k}_2\sigma_2$ are interchanged or $\boldsymbol{k}_3\sigma_3$, $\boldsymbol{k}_4\sigma_4$ are interchanged, so that the sum includes all the *different* Φ_n four times. We note that the matrix element

$$\langle \Phi_n | c^\dagger_{\boldsymbol{p}+\boldsymbol{k},\sigma} c^\dagger_{\boldsymbol{q}-\boldsymbol{k},\sigma'} c_{\boldsymbol{q}\sigma'} c_{\boldsymbol{p}\sigma} | \Phi_0 \rangle \tag{3.83}$$

must vanish unless the operators destroy $\boldsymbol{k}_1\sigma_1$, $\boldsymbol{k}_2\sigma_2$ and create $\boldsymbol{k}_3\sigma_3$, $\boldsymbol{k}_4\sigma_4$.

The expression (3.82) is awkward, but not impossible, and, if indeed second-order perturbation theory were sufficient here, as is so often the case, there would be no need for the elaborate perturbation techniques described in later chapters. Unfortunately, this is not so, and we shall show below that $\Delta E^{(2)}$ diverges.

The expression (3.82) is generally written as the sum of two terms, and this may be done in two different ways, both of which have their uses. In the present chapter†, $\Delta E^{(2)}$ will be split into the contribution from the interaction of pairs of electrons with antiparallel spins ($\sigma_1 \neq \sigma_2$) and that from the interaction of electrons with parallel spins ($\sigma_1 = \sigma_2$). Only the former, which we shall denote by $\Delta E^{(2)}_{\uparrow\downarrow}$, will be explicitly calculated, since it diverges and we only wish to demonstrate the divergence of $\Delta E^{(2)}$. It is somewhat easier to handle than the parallel spin part which, as will be seen in ch. 9, also diverges, and its divergence does not cancel that of $\Delta E^{(2)}_{\uparrow\downarrow}$.

3.6.1. *Antiparallel spins*

We shall first consider a particular Φ_n with $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, $\boldsymbol{k}_3$, $\boldsymbol{k}_4$ fixed and $\sigma_1 = \frac{1}{2}$, $\sigma_3 = \frac{1}{2}$. This implies that $\sigma_2 = -\frac{1}{2}$, since $\sigma_1 \neq \sigma_2$, and we shall see that the only non-zero matrix elements are those with $\sigma_4 = -\frac{1}{2}$ also. In fact, the matrix element (3.83) is zero except in the following cases:

(1) $\boldsymbol{p} = \boldsymbol{k}_1$, $\boldsymbol{q} = \boldsymbol{k}_2$, $\boldsymbol{p}+\boldsymbol{k} = \boldsymbol{k}_3$, $\boldsymbol{q}-\boldsymbol{k} = \boldsymbol{k}_4$, $\sigma = \frac{1}{2}$, $\sigma' = -\frac{1}{2} = \sigma_4$.
This gives

$$\boldsymbol{k} = \boldsymbol{k}_3 - \boldsymbol{k}_1 = \boldsymbol{k}_2 - \boldsymbol{k}_4, \tag{3.84}$$

which requires that

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3 + \boldsymbol{k}_4. \tag{3.85}$$

In other words, for a non-zero matrix element, *momentum must be conserved.*

(2) $\boldsymbol{q} = \boldsymbol{k}_1$, $\boldsymbol{p} = \boldsymbol{k}_2$, $\boldsymbol{q}-\boldsymbol{k} = \boldsymbol{k}_3$, $\boldsymbol{p}+\boldsymbol{k} = \boldsymbol{k}_4$, $\sigma' = \frac{1}{2}$, $\sigma = -\frac{1}{2} = \sigma_4$.

---

† The alternative way of expressing $\Delta E^{(2)}$ will be described in §9.2.

This gives

$$\boldsymbol{k} = \boldsymbol{k}_1 - \boldsymbol{k}_3 = \boldsymbol{k}_4 - \boldsymbol{k}_2, \tag{3.86}$$

again implying momentum conservation.

Case (2) is obtained from case (1) by inverting the order of *both* the destruction *and* the creation operators, so that the sign of the matrix element is the same in both cases and its numerical value is unity. Also, we have $k^2 = |\boldsymbol{k}_3 - \boldsymbol{k}_1|^2$ in both cases, so that the sum over $\boldsymbol{k}_3$, say, can be expressed as a sum over $\boldsymbol{k}$. We may consider $\boldsymbol{k}_4$ to be fixed by momentum conservation, which gives $\boldsymbol{k}_4 = \boldsymbol{k}_2 - \boldsymbol{k}_1 + \boldsymbol{k}_3$.

The numerator in (3.82) thus has the value $(4\pi e^2/\Omega k^2)^2$, and there are four combinations of σ_1 and σ_3, namely $\sigma_1 = \pm\frac{1}{2}$, $\sigma_3 = \pm\frac{1}{2}$, which give the same value, the values of σ_2 and σ_4 being automatically fixed. This cancels the factor $\frac{1}{4}$ in (3.82).

With $\boldsymbol{k} = \boldsymbol{k}_3 - \boldsymbol{k}_1 = \boldsymbol{k}_2 - \boldsymbol{k}_4$, we have

$$\begin{aligned} k_3^2 + k_4^2 - k_1^2 - k_2^2 &= |\boldsymbol{k}+\boldsymbol{k}_1|^2 + |\boldsymbol{k}-\boldsymbol{k}_2|^2 - k_1^2 - k_2^2 \\ &= 2\boldsymbol{k}\cdot(\boldsymbol{k}_1 - \boldsymbol{k}_2 + \boldsymbol{k}). \end{aligned} \tag{3.87}$$

Finally, then,

$$\Delta E^{(2)}_{\uparrow\downarrow} = -\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}} \frac{m(4\pi e^2/\Omega k^2)^2}{\hbar^2 \boldsymbol{k}\cdot(\boldsymbol{k}_1 - \boldsymbol{k}_2 + \boldsymbol{k})},$$

or

$$\Delta E^{(2)}_{\uparrow\downarrow} = -\frac{16\pi^2 e^4 m}{\hbar^2\Omega^2} \sum_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{k}} \frac{1}{k^4 \boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})}, \tag{3.88}$$

where $p, q < k_F$ and $|\boldsymbol{p}+\boldsymbol{k}|, |\boldsymbol{q}-\boldsymbol{k}| > k_F$.

Expressing this as an integral, in the usual way, we have

$$\Delta E^{(2)}_{\uparrow\downarrow} = -\frac{16\pi^2 e^4 m}{\hbar^2\Omega^2}\left(\frac{\Omega}{8\pi^3}\right)^3 \iiint \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{k}}{k^4 \boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})}. \tag{3.89}$$

It is customary to measure energy in rydbergs and k, p and q in units of k_F, given by (3.65), which leads to

$$\Delta E^{(2)}_{\uparrow\downarrow} = -\frac{3N}{16\pi^5}\int \frac{\mathrm{d}\boldsymbol{k}}{k^4}\int \mathrm{d}\boldsymbol{q}\int \frac{\mathrm{d}\boldsymbol{p}}{\boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})}\,\mathrm{ry}. \tag{3.90}$$

the region of integration now being

$$p, q < 1 \quad \text{and} \quad |\boldsymbol{p}+\boldsymbol{k}|, |\boldsymbol{q}-\boldsymbol{k}| > 1. \tag{3.91}$$

Now, the greatest contribution to the integral will be from the region where k is very small, and in this region we must have $p \approx 1$ and $q \approx 1$, since $p < 1$ and $|\boldsymbol{p}+\boldsymbol{k}| > 1$, etc. Thus, for very small k, $\boldsymbol{p}$ and $\boldsymbol{q}$ must lie in a shell of approximately unit radius and thickness of the order of k. In other words, the greatest contribution to the integral is from states $\boldsymbol{p}$ and $\boldsymbol{q}$ near the Fermi surface. In order to investigate this further, let us define x and y by

$$\boldsymbol{k}\cdot\boldsymbol{p} = kpx, \qquad \boldsymbol{k}\cdot\boldsymbol{q} = -kpy. \tag{3.92}$$

Now, from $|\boldsymbol{p}+\boldsymbol{k}| > 1$, we have

$$p^2+k^2+2kpx > 1. \tag{3.93}$$

If k is very small, so that k^2 can be neglected and $p \approx 1$, this becomes, approximately,

$$kx > 1-p, \tag{3.94}$$

so that the limits on p are

$$1 > p > 1-kx. \tag{3.95}$$

Similarly, we find that

$$1 > q > 1-ky. \tag{3.96}$$

Clearly, both x and y must be positive, since k is positive.

Thus, in the region where k is very small,

$$\begin{aligned}\int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{\boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})} &= \int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{k(px+qy+k)} \\ &\approx \int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{k(x+y)}.\end{aligned} \tag{3.97}$$

Now, if θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{p}$, so that $x = \cos\theta$, the volume element for the $\boldsymbol{p}$ integration, using spherical polar coordinates and the fact that the integrand is independent of the azimuthal angle, is

$$2\pi p^2 \sin\theta\, \mathrm{d}\theta\, \mathrm{d}p \approx -2\pi\, \mathrm{d}x\, \mathrm{d}p. \tag{3.98}$$

Since x is positive, so that $0 < \theta < \frac{1}{2}\pi$, we then have

$$\begin{aligned}\int \frac{\mathrm{d}\boldsymbol{p}}{k(x+y)} &\approx 2\pi \int_0^1 \mathrm{d}x \int_{1-kx}^1 \frac{\mathrm{d}p}{k(x+y)} \\ &= 2\pi \int_0^1 \frac{x\,\mathrm{d}x}{x+y}.\end{aligned} \tag{3.99}$$

The $\boldsymbol{q}$ integration is similar, giving finally

$$\int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{k(x+y)} = (2\pi)^2 k \int_0^1 \mathrm{d}y \int_0^1 \frac{xy\,\mathrm{d}x}{x+y}, \tag{3.100}$$

for very small k. The double integral on the right is independent of k, so that the contribution to $\Delta E^{(2)}_{\uparrow\downarrow}$, eq. (3.90), from small values of k is proportional to

$$\int \frac{\mathrm{d}\boldsymbol{k}}{k^3} = 4\pi \int \frac{\mathrm{d}k}{k} = 4\pi \log k. \tag{3.101}$$

This tends to $-\infty$ as $k \to 0$.

We have thus shown that $\Delta E^{(2)}_{\uparrow\downarrow}$, and consequently $\Delta E^{(2)}$, diverge logarithmically at $k = 0$. The part of $\Delta E^{(2)}$ due to interactions between electrons with parallel spins will be seen in ch. 9 to consist of two terms, one of which is the same as $\Delta E^{(2)}_{\uparrow\downarrow}$ and the other of which does not diverge at $k = 0$.

Second-order perturbation theory therefore does not work for a free-electron gas, if the perturbation is taken to be the whole of the Coulomb interaction of the electrons. In fact, it will be seen, also in ch. 9, that the perturbation energy of every order after the first diverges. None the less, perturbation theory *can* be applied in this case, at least when the electron density is high, but the elaborate techniques required will take up much of the remainder of the book. First, as a slight digression, we shall, in the following chapter, describe an attempt to circumvent the difficulty with the second-order perturbation energy by using the concept of *plasma oscillations*.

CHAPTER 4

PLASMA OSCILLATIONS IN A FREE-ELECTRON GAS

4.1. Résumé of the Plasma Theory†

Any highly-ionized gas, that is, a system of positively charged ions and virtually free electrons, is called a *plasma*. Such plasmas occur in gas discharge tubes, and have been extensively studied for many years. Clearly, a metal may also be regarded as a plasma, since it consists of a lattice of positive ions and a gas of almost-free valence electrons. The motion of the relatively heavy ions may generally be neglected in comparison with the motion of the electrons, so that, as an idealization of a plasma, we may take the free-electron gas described in §3.4, where the ions are replaced by a uniform distribution of positive charge. We shall be concerned exclusively with such an idealized plasma in the present work.

According to classical electromagnetic theory a free-electron gas will exhibit oscillations of the electron density, analogous to sound waves, which are known as *plasma oscillations*. If dispersion is neglected, the angular frequency ω_{p} of these oscillations is

$$\omega_{\mathrm{p}} = \left(\frac{4\pi N e^2}{\Omega m}\right)^{1/2}, \tag{4.1}$$

---

† An introductory account of the theory of plasma oscillations in an electron gas, and its application to metals, is given in WM, ch. 10. Although the classical theory and the principal results of the quantum theory are discussed there at some length, the quantum-mechanical theory itself is not presented in detail – in particular, the unitary transformation of the Hamiltonian is not described. In order to reduce overlap to a minimum, we shall here, apart from the initial summary, concentrate upon those aspects of the theory not included, or only briefly described, in WM.

where N is the total number of electrons and Ω is the volume of the plasma. There is, in fact, a certain amount of dispersion, the angular frequency ω_k corresponding to the wave number k (i.e. to the wavelength $2\pi/k$) being given approximately by

$$\begin{aligned}\omega_k^2 &= \omega_{\mathrm{p}}^2 + k^2\langle v_i^2\rangle_{\mathrm{av}}\\ &= \omega_{\mathrm{p}}^2 + \frac{2}{m}E_{\mathrm{F}}k^2,\end{aligned} \tag{4.2}$$

where $\langle v_i^2\rangle_{\mathrm{av}}$ is the average of the squared velocities of the electrons and E_{F} is the Fermi energy (eq. (3.68)). The dispersion is small, however, because there is an upper limit, denoted by k_{c}, to the wave number, beyond which plasma oscillation does not occur.

Following Bohm and Pines†, we express the particle density (regardless of spin – compare eq. (2.141))

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} \delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{4.3}$$

in Fourier series

$$\rho(\boldsymbol{r}) = \frac{1}{\Omega}\sum_{\boldsymbol{k}} \rho_{\boldsymbol{k}} \exp(i\boldsymbol{k}\cdot\boldsymbol{r}) \tag{4.4}$$

within the box of volume Ω. This defines‡ the *density fluctuations* $\rho_{\boldsymbol{k}}$ as

$$\rho_{\boldsymbol{k}} = \int \rho(\boldsymbol{r}) \exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \sum_i \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_i), \tag{4.5}$$

so that $\rho_0 = N$. Bohm and Pines found that, within the *random phase approximation*, which will be described later, the equation of motion of $\rho_{\boldsymbol{k}}$ is

$$\ddot{\rho}_{\boldsymbol{k}} + \omega_{\mathrm{p}}^2\rho_{\boldsymbol{k}} = -\sum_i (\boldsymbol{k}\cdot\boldsymbol{v}_i)^2 \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_i), \tag{4.6}$$

---

† In addition to the elementary account in WM, more detailed accounts of the theory, together with references to the original papers, may be found in the following review articles:

Pines, D., 1955, Solid State Physics **1**, 367 (Academic Press, New York).

Raimes, S., 1957, Reports on Progress in Physics **20**, 1 (Institute of Physics, London).

A very full discussion of the theory is given in the following book:

Pines, D., 1963, *Elementary Excitations in Solids* (Benjamin, New York).

‡ Other definitions, differing from ours by a multiplying constant, are often used – for example, the factor $1/\Omega$ in (4.4) may be absorbed in the $\rho_{\boldsymbol{k}}$. Bohm and Pines took the volume Ω to be unity, so that, although their expression for $\rho_{\boldsymbol{k}}$ is that given in (4.5), in their work $\rho_0 = N/\Omega$, the average electron density.

or, alternatively,

$$\ddot{\rho}_k = -\sum_i \left[(\boldsymbol{k}\cdot\boldsymbol{v}_i)^2 + \omega_p^2\right] \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_i), \tag{4.7}$$

where $\boldsymbol{v}_i = \dot{\boldsymbol{r}}_i$. This shows clearly that, for very small k, the ρ_k will oscillate approximately with angular frequency ω_p. For larger k, however, these oscillations will be damped by the random thermal motion of the electrons, which gives rise to the term on the right-hand side of (4.6). Eq. (4.7) shows that plasma oscillation might be expected so long as

$$\omega_p \gg kv_0, \tag{4.8}$$

where v_0 is the speed of an electron at the Fermi surface. As a very rough upper bound to the value of k_c we may thus write

$$k_c \approx \omega_p/v_0, \tag{4.9}$$

which gives the value $k_c^{-1} \approx 1$ Å for the density of the valence electrons occurring in metals. This extreme value of k, when inserted in eq. (4.2), leads to a value of ω_k differing from ω_p by only 26 per cent.

All the foregoing is based upon classical electron theory. Although we expect the qualitative results still to hold, it is clear that, for an electron gas of the density occurring in metals, quantum mechanics must be used, and the method of doing this will be described in the following section. As a preliminary, we might apply elementary quantum theory to the results described above. The plasma oscillations may be represented approximately by a finite set of harmonic oscillators with angular frequency ω_p, one for each density fluctuation ρ_k with $k < k_c$. According to quantum theory the ground-state, or zero-point, energy of such an oscillator is $\frac{1}{2}\hbar\omega_p$, and the excitation energy, that is, the energy required to raise the oscillator from one state to that of next higher energy, is $\hbar\omega_p$.

The magnitude of the quantum $\hbar\omega_p$, which we shall call the *plasmon energy*†, is crucial to the success of the plasma theory as a computational

---

† In WM, ch. 10, the term *plasmon* was used essentially as a short-hand form of *plasma oscillation.* This is still quite common, although it is now more usual to use the term to mean a *quantum* of plasma oscillation, just as a phonon is a quantum of sound. Such quanta are known as *elementary excitations* and may often be treated as particles. In particle language we would refer to the *creation or emission of a plasmon* rather than to the *excitation of a plasma oscillation.* It is seldom that particle language is used consistently, however. For example, the state of the system in which all plasma oscillators are in their ground states should, strictly speaking, be referred to as the state in which no plasmons are present, but it is quite customary to say that the *plasmons* are in their ground states. Although this is unlikely to lead to confusion, we shall avoid such inconsistency by not using the particle language.

tool, as we shall now explain. From eq. (4.1) and the definition of r_s in eq. (3.67), we find

$$\hbar\omega_p = \frac{2\sqrt{3}}{r_s^{3/2}}\,\mathrm{ry}, \tag{4.10}$$

where r_s is measured in Bohr units. For the valence-electron density occurring in sodium, with $r_s = 4$ Bohr units, this gives $\hbar\omega_p = 0.433$ ry $=$ 5.9 eV. For the valence-electron densities occurring in all metals $\hbar\omega_p$ is found to lie between about 3 and 25 eV.

Now, at normal temperatures T, only relatively few electrons in states near the Fermi surface are excited, and their excitation energies are of the order of kT, where k is Boltzmann's constant. Since states lying well below the Fermi surface are all occupied, according to the Pauli principle, it follows that thermally excited electrons have only energy of the order of kT, which is about 0.025 eV at room temperature, to give away. This energy is very small compared with the plasmon energy $\hbar\omega_p$, which means that thermal excitation of plasma oscillations can usually be neglected. In other words, plasma oscillations may generally be assumed to remain in their ground states unless excited by some other method, such as the passage of a fast charged particle through the metal.

The foregoing is a most important result, for it means that the plasma oscillations take no active part in many electronic processes and may often be ignored. Since plasma oscillation is an organized motion of a large number of electrons, it may be inferred that the *long-range* part of the Coulomb interaction is responsible for the plasma oscillations, so that in calculations where the latter may be ignored the former may also be ignored. We shall see in the following section that the remaining interaction has an effective range of about 1 Å, and this is so short that reasonable results might often be expected by ignoring this interaction also, that is, by treating the electrons as non-interacting particles. The plasma theory thus offers a justification of the independent-particle approximation, which has been used with much success in the theory of metals. Even when the short-range interaction cannot be ignored, as in the calculation of the energy of the system, we shall see that it can be treated by elementary perturbation theory without divergence difficulties being encountered.

4.2. Quantum-mechanical Theory

We are considering a box of volume Ω containing N electrons together with a uniform distribution, or *background*, of positive charge, as described in §3.4. Using the expression (3.68) for the Coulomb interaction, the Hamiltonian for the system becomes

$$H=\sum_{i=1}^{N}\frac{p_i^2}{2m}+\sum_{i,j\neq i}^{N}\sum_{\boldsymbol{k}}{}'\,\frac{2\pi e^2}{\Omega k^2}\exp\left[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)\right], \tag{4.11}$$

where $\boldsymbol{p}_i$ $(=\mathrm{i}\hbar\nabla_i)$ is the momentum operator for electron i. Alternatively, using the definition (4.5) of the density fluctuation ρ_k, we may write†

$$H=\sum_{i=1}^{N}\frac{p_i^2}{2m}+\sum_{\boldsymbol{k}}\frac{2\pi e^2}{\Omega k^2}(\rho_k^*\rho_k-N). \tag{4.12}$$

The energy levels of the system are the eigenvalues of the Schrödinger equation

$$H\Psi=E\Psi, \tag{4.13}$$

and may thus be found, in principle, without explicitly considering plasma oscillations. However, as explained in the foregoing section, we expect plasma oscillations to occur, and this leads us to suspect the possibility of transforming the Hamiltonian (4.12) in such a way that these oscillations show themselves explicitly as a finite set of simple harmonic oscillator Hamiltonians, one for each $\boldsymbol{k}$ with $k<k_\mathrm{c}$, that is, Hamiltonians of the type‡

$$\tfrac{1}{2}(P_k^\dagger P_k+\omega^2Q_k^\dagger Q_k), \tag{4.14}$$

where P_k and Q_k are the canonically conjugate momentum and coordinate operators relating to the oscillator.

To find the necessary transformation operator is a matter of inspired guesswork, with a knowledge of the result which is desired. In fact, it is found convenient to start, not with the Hamiltonian (4.12), but with a so-called *model Hamiltonian*, which contains an extra term

$$H_1=\sum_{k<k_\mathrm{c}}(\tfrac{1}{2}P_k^\dagger P_k-M_kP_k^\dagger\rho_k), \tag{4.15}$$

where

$$M_k=\left(\frac{4\pi e^2}{\Omega k^2}\right)^{1/2}. \tag{4.16}$$

(*N.B.* Here and in future we write $\Sigma_{k<k_\mathrm{c}}$ to mean $\Sigma_{\boldsymbol{k}(k<k_\mathrm{c})}$.)

This introduces immediately a set of P_k, and implicitly a set of Q_k, with $k<k_\mathrm{c}$. In order to ensure that (4.15) is Hermitian we impose the conditions‡

$$P_k^\dagger=P_{-k},\qquad Q_k^\dagger=Q_{-k}, \tag{4.17}$$

the latter being implicit. This means that P_k and Q_k are not Hermitian. However, we note that the same condition is satisfied by ρ_k, since, from (4.5),

$$\rho_k^\dagger=\rho_k^*=\rho_{-k}. \tag{4.18}$$

The interpretation of this additional term need not concern us – we may regard its inclusion as a formal step to be justified by results. Of course, it is not permissible to take this step unconditionally, for in so doing we have increased the number of degrees of freedom of the system by the number of the oscillator variables, namely, the number of $\boldsymbol{k}$-vectors

---

† Here and in the remainder of this chapter we drop the prime on the summation sign. In *any* sum over wave vectors $\boldsymbol{k}$ it is to be understood that $\boldsymbol{k}=\boldsymbol{0}$ must be excluded.

‡ To see that this is indeed a simple harmonic oscillator Hamiltonian, let us take as an example

$$Q_k=x=Q_k^\dagger,\quad P_k=-\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}=P_k^\dagger,$$

which are canonically conjugate operators. Then (4.14) becomes

$$-\frac{\hbar^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\tfrac{1}{2}\omega^2x^2,$$

which has the form of the simple harmonic oscillator Hamiltonian encountered in elementary wave mechanics (cf. WM, eq. (1.53), with $m=1$). In this simple illustration we have chosen Q_k and P_k to be Hermitian (see Appendix I). If this were always so, we could write (4.14) as

$$\tfrac{1}{2}(P_k^2+\omega^2Q_k^2).$$

However, the more general form we have used does not require P_k and Q_k to be Hermitian (although (4.14) itself is Hermitian, of course), and this is desirable, since we expect P_k and Q_k to be related to ρ_k, which, as eq. (4.18) shows, is not Hermitian.

‡ The first term is Hermitian by the rules of Appendix I, whether (4.17) applies or not, and for the second term we have

$$\left(\sum_{k<k_\mathrm{c}}M_kP_k^\dagger\rho_k\right)^\dagger=\sum_{k<k_\mathrm{c}}M_kP_k\rho_k^*$$
$$=\sum_{k<k_\mathrm{c}}M_kP_{-k}\,\rho_{-k}^*=\sum_{k<k_\mathrm{c}}M_kP_k^\dagger\rho_k.$$

from (4.17), (4.18) and the spherically symmetrical distribution of the $\boldsymbol{k}$-vectors.

lying within a sphere of radius k_c in $\boldsymbol{k}$-space, which is

$$\frac{\Omega}{8\pi^3}\frac{4\pi}{3}k_c^3 = \frac{\Omega k_c^3}{6\pi^2}. \tag{4.19}$$

Our intention, however, is merely to re-describe part of the motion of the electrons in terms of plasma oscillations, the number of degrees of freedom remaining fixed at $3N$. We therefore impose on the wave function Ψ the *subsidiary conditions*

$$P_k\Psi = 0, \quad k < k_c, \tag{4.20}$$

We assume that P_k and Q_k satisfy the commutation relations

$$[P_k, Q_{k'}] = -\mathrm{i}\hbar\delta_{kk'} \tag{4.21}$$

(this, in fact, is what we mean by saying that P_k and Q_k are canonically conjugate), so that P_k can be represented by the differential operator $-i\hbar\partial/\partial Q_k$. In this representation, the subsidiary conditions (4.20) become

$$\frac{\partial\Psi}{\partial Q_k} = 0, \quad k < k_c, \tag{4.22}$$

and merely ensure that Ψ does not contain any of the Q_k, but is a function of the electronic coordinates only. It is clear then that the additional term in the Hamiltonian cannot affect the eigenvalues, for it operates on Ψ to give zero; thus

$$(H + H_1)\Psi = H\Psi = E\Psi, \tag{4.23}$$

with the same E as in (4.13).

It seems, therefore, that the inclusion of the extra term is correct[†] but, so far, it is also pointless, for it gives us no new information. To remedy this we shall make a unitary, or canonical, transformation of the Hamiltonian $H + H_1$. As shown in Appendix II, it is legitimate to make *any* such transformation, but the problem is to find one which will give significant and useful results.

The unitary operator we choose is

$$U = \exp(\mathrm{i}S/\hbar), \tag{4.24}$$

[†] There is, in fact, a slight difficulty regarding the normalization of Ψ. Since Ψ does not contain Q_k, the integral of $|\Psi|^2$ with respect to Q_k diverges, so that, strictly speaking, Ψ is not normalizable. We assume, however, that this formal difficulty may be overcome by applying some type of 'box normalization', ultimately allowing the dimensions of the box (that is, the range of values of the Q_k) to increase without limit.

where

$$S = \sum_{k<k_c} M_k Q_k \rho_k. \tag{4.25}$$

We have

$$U^\dagger = \exp(-\mathrm{i}S^\dagger/\hbar) \tag{4.26}$$

and

$$S^\dagger = \sum_{k<k_c} M_k Q_k^\dagger \rho_k^* = \sum_{k<k_c} M_k Q_{-k}\rho_{-k} = S \tag{4.27}$$

by (4.17), so that

$$U^\dagger = U^{-1}, \tag{4.28}$$

and the operator U is, indeed, unitary. The transformation we require is such that any operator O becomes O_{new}, say, where

$$O_{\text{new}} = U^{-1}OU, \tag{4.29}$$

and the wave function Ψ becomes Ψ_{new}, where

$$\Psi_{\text{new}} = U^{-1}\Psi = \exp(-\mathrm{i}S/\hbar)\Psi. \tag{4.30}$$

Now, let us see how the various operators transform. First, we have

$$(\boldsymbol{r}_i)_{\text{new}} = U^{-1}\boldsymbol{r}_i U = \boldsymbol{r}_i, \tag{4.31}$$

and

$$(Q_k)_{\text{new}} = U^{-1}Q_k U = Q_k, \tag{4.32}$$

since $\boldsymbol{r}_i$ and Q_k commute with U. It also follows that

$$(\rho_k)_{\text{new}} = U^{-1}\rho_k U = \rho_k. \tag{4.33}$$

In other words, $\boldsymbol{r}_i$, Q_k and ρ_k are unchanged by this transformation. This is not so for P_k and $\boldsymbol{p}_i$, however. Owing to (4.21), we have

$$[P_k, U] = -\mathrm{i}\hbar\frac{\partial U}{\partial Q_k}, \tag{4.34}$$

so that

$$\begin{aligned}(P_k)_{\text{new}} &= U^{-1}P_k U = P_k + U^{-1}[P_k, U]\\ &= P_k - \mathrm{i}\hbar U^{-1}\frac{\partial U}{\partial Q_k}\\ &= P_k + M_k\rho_k.\end{aligned} \tag{4.35}$$

Similarly, if (p_{ix}, p_{iy}, p_{iz}) are the Cartesian components of $\boldsymbol{p}_i$, and (x_i, y_i, z_i) those of $\boldsymbol{r}_i$, we have

$$[p_{ix}, U] = -\mathrm{i}\hbar \frac{\partial U}{\partial x_i}, \tag{4.36}$$

etc., so that

$$\begin{aligned}(p_{ix})_{\text{new}} &= U^{-1} p_{ix} U = p_{ix} + U^{-1}[p_{ix}, U] \\ &= p_{ix} - \mathrm{i}\hbar U^{-1} \frac{\partial U}{\partial x_i} = p_{ix} + \frac{\partial S}{\partial x_i} \\ &= p_{ix} - \mathrm{i} \sum_{k<k_c} M_k Q_k k_x \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i).\end{aligned} \tag{4.37}$$

It follows that

$$(\boldsymbol{p}_i)_{\text{new}} = \boldsymbol{p}_i - i \sum_{k<k_c} M_k Q_k \boldsymbol{k} \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_i). \tag{4.38}$$

We note that the subsidiary conditions (4.20) become

$$(U^{-1} P_k U) U^{-1}\Psi = 0, \quad k < k_c, \tag{4.39}$$

or

$$(P_k + M_k \rho_k)\Psi_{\text{new}} = 0, \quad k < k_c, \tag{4.40}$$

so that the field momenta P_k are directly related to the density fluctuations ρ_k for long wavelengths, which we expect to oscillate with roughly the plasma frequency ω_p.

We now wish to find

$$H_{\text{new}} = U^{-1}(H + H_1)U \equiv \mathscr{H}, \tag{4.41}$$

say, where

$$\begin{aligned}H + H_1 = \sum_i \frac{p_i^2}{2m} &+ \tfrac{1}{2} \sum_k M_k^2 (\rho_k^* \rho_k - N) \\ &+ \sum_{k<k_c} (\tfrac{1}{2} P_k^\dagger P_k - M_k P_k^\dagger \rho_k).\end{aligned} \tag{4.42}$$

First, we have

$$\boldsymbol{p}_i^2 = p_{ix}^2 + p_{iy}^2 + p_{iz}^2, \tag{4.43}$$

and, from (4.37),

$$\begin{aligned}(p_{ix})^2_{\text{new}} &= \Big[p_{ix} - \mathrm{i} \sum_{k<k_c} M_k Q_k k_x \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i)\Big]^2 \\ &= p_{ix}^2 - \mathrm{i} \sum_{k<k_c} M_k Q_k k_x [p_{ix} \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) + \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) p_{ix}] \\ &\quad - \sum_{k,l<k_c} M_k M_l Q_k Q_l k_x l_x \exp[-\mathrm{i}(\boldsymbol{k}+\boldsymbol{l})\cdot\boldsymbol{r}_i].\end{aligned} \tag{4.44}$$

Now,

$$[p_{ix}, \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i)] = -\mathrm{i}\hbar \frac{\partial}{\partial x_i} \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) = -\hbar k_x \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i), \tag{4.45}$$

so that

$$\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) p_{ix} = p_{ix} \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) + \hbar k_x \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) \tag{4.46}$$

and

$$\begin{aligned}(p_{ix})^2_{\text{new}} = p_{ix}^2 &- \mathrm{i} \sum_{k<k_c} M_k Q_k k_x (2p_{ix} + \hbar k_x) \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) \\ &- \sum_{k,l<k_c} M_k M_l Q_k Q_l k_x l_x \exp[-\mathrm{i}(\boldsymbol{k}+\boldsymbol{l})\cdot\boldsymbol{r}_i].\end{aligned} \tag{4.47}$$

This gives

$$\begin{aligned}\Big(\sum_i \frac{p_i^2}{2m}\Big)_{\text{new}} = \sum_i \frac{p_i^2}{2m} &- \frac{\mathrm{i}}{2m} \sum_i \sum_{k<k_c} M_k Q_k \boldsymbol{k}\cdot(2\boldsymbol{p}_i + \hbar\boldsymbol{k}) \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i) \\ &- \frac{1}{2m} \sum_i \sum_{k,l<k_c} M_k M_l Q_k Q_l \boldsymbol{k}\cdot\boldsymbol{l} \exp[-\mathrm{i}(\boldsymbol{k}+\boldsymbol{l})\cdot\boldsymbol{r}_i].\end{aligned} \tag{4.48}$$

The second term on the right-hand side of (4.42), the Coulomb interaction, commutes with U and is thus unchanged in the transformation. It will be convenient to write it in the following way:

$$\tfrac{1}{2} \sum_k M_k^2 (\rho_k^* \rho_k - N) = H_{\text{s.r.}} + \tfrac{1}{2} \sum_{k<k_c} M_k^2 (\rho_k^* \rho_k - N), \tag{4.49}$$

where

$$H_{\text{s.r.}} = \tfrac{1}{2} \sum_{k>k_c} M_k^2 (\rho_k^* \rho_k - N), \tag{4.50}$$

which we shall show to be a short-range interaction.

Passing to the third term on the right-hand side of (4.42), we have, from (4.35),

$$(P_k^\dagger P_k)_{\text{new}} = (P_k^\dagger + M_k \rho_k^*)(P_k + M_k \rho_k)$$
$$= P_k^\dagger P_k + M_k(P_k^\dagger \rho_k + P_k \rho_k^*) + M_k^2 \rho_k^* \rho_k. \qquad (4.51)$$

Now,

$$\sum_{k<k_\mathrm{c}} M_k(P_k^\dagger \rho_k + P_k \rho_k^*) = \sum_{k<k_\mathrm{c}} M_k(P_k^\dagger \rho_k + P_{-k} \rho_{-k}^*)$$
$$= 2 \sum_{k<k_\mathrm{c}} M_k P_k^\dagger \rho_k, \qquad (4.52)$$

using (4.17), (4.18) and the spherically symmetrical distribution of the **k**-vectors. Hence

$$\Big(\sum_{k<k_\mathrm{c}} \tfrac{1}{2} P_k^\dagger P_k\Big)_{\text{new}} = \sum_{k<k_\mathrm{c}} \tfrac{1}{2} P_k^\dagger P_k$$
$$+ \sum_{k<k_\mathrm{c}} M_k P_k^\dagger \rho_k + \tfrac{1}{2} \sum_{k<k_\mathrm{c}} M_k^2 \rho_k^* \rho_k. \qquad (4.53)$$

Finally,

$$(M_k P_k^\dagger \rho_k)_{\text{new}} = M_k \rho_k (P_k^\dagger + M_k \rho_k^*), \qquad (4.54)$$

and

$$\Big(-\sum_{k<k_\mathrm{c}} M_k P_k^\dagger \rho_k\Big)_{\text{new}} = -\sum_{k<k_\mathrm{c}} M_k P_k^\dagger \rho_k - \sum_{k<k_\mathrm{c}} M_k^2 \rho_k^* \rho_k. \qquad (4.55)$$

The new Hamiltonian $\mathscr{H}$ is the sum of (4.48), (4.49), (4.53) and (4.55), which gives

$$\mathscr{H} = \sum_i \frac{p_i^2}{2m} - \frac{i}{2m} \sum_i \sum_{k<k_\mathrm{c}} M_k Q_k \boldsymbol{k} \cdot (2\boldsymbol{p}_i + \hbar \boldsymbol{k}) \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}_i)$$
$$+ \frac{1}{2m} \sum_i \sum_{k,l<k_\mathrm{c}} M_k M_l Q_k Q_{-l} \boldsymbol{k} \cdot \boldsymbol{l} \exp[-\mathrm{i}(\boldsymbol{k}-\boldsymbol{l}) \cdot \boldsymbol{r}_i]$$
$$+ H_{\text{s.r.}} - \sum_{k<k_\mathrm{c}} \frac{2\pi e^2 N}{\Omega k^2} + \sum_{k<k_\mathrm{c}} \tfrac{1}{2} P_k^\dagger P_k. \qquad (4.56)$$

We have changed **l** to −**l** in the third term on the right-hand side. The latter may be expressed as the sum of two parts – those terms with $\boldsymbol{k} = \boldsymbol{l}$ and the rest. The first part becomes

$$\frac{N}{2m} \sum_{k<k_\mathrm{c}} M_k^2 Q_k Q_{-k} k^2 = \frac{2\pi N e^2}{\Omega m} \sum_{k<k_\mathrm{c}} Q_k Q_{-k}$$
$$= \tfrac{1}{2}\omega_\mathrm{p}^2 \sum_{k<k_\mathrm{c}} Q_k^\dagger Q_k. \qquad (4.57)$$

Hence, (4.56) may be written

$$\mathscr{H} = \sum_i \frac{p_i^2}{2m} + \sum_{k<k_\mathrm{c}} \tfrac{1}{2}(P_k^\dagger P_k + \omega_\mathrm{p}^2 Q_k^\dagger Q_k)$$
$$- \sum_{k<k_\mathrm{c}} \frac{2\pi e^2 N}{\Omega k^2} + H_{\text{s.r.}} + H_{\text{int}} + K, \qquad (4.58)$$

where

$$H_{\text{int}} = -\frac{\mathrm{i}}{2m} \sum_i \sum_{k<k_\mathrm{c}} M_k Q_k \boldsymbol{k} \cdot (2\boldsymbol{p}_i + \hbar \boldsymbol{k}) \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}_i) \qquad (4.59)$$

and

$$K = \frac{1}{2m} \sum_i \sum_{\substack{k,l<k_\mathrm{c} \\ k\neq l}} M_k M_l Q_k Q_{-l} \boldsymbol{k} \cdot \boldsymbol{l} \exp[-\mathrm{i}(\boldsymbol{k}-\boldsymbol{l}) \cdot \boldsymbol{r}_i]. \qquad (4.60)$$

It is important to realize that no approximations have been made in getting from (4.11) to (4.58). The new Hamiltonian is perfectly accurate (subject, of course, to the subsidiary conditions) and contains the required oscillator Hamiltonians and short-range interaction. However, it also contains the extremely awkward terms H_{int} and K, and the viability of the method depends upon our being able to deal with these in a fairly simple manner.

In fact, it is an excellent approximation to drop K altogether. This is because the general term of the sum over **k** and **l** in K contains

$$\rho_{k-l} = \sum_i \exp[-\mathrm{i}(\boldsymbol{k}-\boldsymbol{l}) \cdot \boldsymbol{r}_i],$$

which, with $\boldsymbol{k} \neq \boldsymbol{l}$, is expected to be very small compared with ρ_0 (= N), the terms containing ρ_0 having been retained in (4.57). The reason is simply that ρ_{k-l}, $\boldsymbol{k} \neq \boldsymbol{l}$, is the sum over all electrons of phase factors of the type $\exp[-\mathrm{i}(\boldsymbol{k}-\boldsymbol{l}) \cdot \boldsymbol{r}_i]$, that is to say, complex numbers of modulus unity. Now, the electrons are randomly distributed, so that ρ_{k-l} is the sum of a large number of unit vectors with random directions in the complex plane. Clearly this sum must be small compared with N, since for each unit vector there will be one in almost the opposite direction which will very nearly cancel it. This reason for the neglect of K is known as the *random phase approximation*, and makes its appearance in the classical theory of plasma oscillations as well as in various guises throughout the whole theory of many-electron systems.

The interaction term H_{int} cannot immediately be dismissed for the same reason, owing to the presence of $\boldsymbol{p}_i$ – for each phase factor $\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_i)$ there will be one which is very nearly its negative, but it is unlikely that the corresponding momenta will be the same. Bohm and Pines demonstrated the effect of this term by a further unitary transformation, which almost completely eliminated H_{int}. This transformation is too complicated for inclusion here, but its principal effects on the rest of the Hamiltonian are to change the first two terms of (4.58) to

$$\sum_i \frac{p_i^2}{2m}\left(1-\frac{\beta^3}{6}\right)+\sum_{k<k_c}\tfrac{1}{2}(P_k^\dagger P_k+\omega_k^2 Q_k^\dagger Q_k), \tag{4.61}$$

where

$$\beta = k_c/k_F, \tag{4.62}$$

and ω_k is given, as far as terms of the second order in k, by the classical dispersion relation (4.2). It will be shown that, for sodium, $\beta \approx 0.7$, so that the change in the particle kinetic energy is about 6% or 0.01 ry per electron. Also, although we have already seen that the dispersion is slight, the change in the zero-point energy of the oscillators in going from ω_p to ω_k will cancel some of the change in the particle kinetic energy. We may conclude, therefore, that the effect of H_{int}, if not completely negligible, is certainly very small. In the interests of simplicity, we shall drop this term also, so that the approximate new Hamiltonian is

$$\begin{aligned}\mathscr{H} = {} & \sum_i \frac{p_i^2}{2m}+\sum_{k<k_c}\tfrac{1}{2}(P_k^\dagger P_k+\omega_p^2 Q_k^\dagger Q_k)\\ & -\sum_{k<k_c}\frac{2\pi e^2 N}{\Omega k^2}+H_{\text{s.r.}},\end{aligned} \tag{4.63}$$

subject to the subsidiary conditions (4.40). This describes the system as a set of plasma oscillations plus a set of particles, or, more correctly, *quasi-particles*, whose interaction is not the full Coulomb interaction but $H_{\text{s.r.}}$. The third term is constant and represents part of the electron self-energy which is not accounted for by the plasma oscillations.

Let us consider $H_{\text{s.r.}}$. From (4.50), we have

$$\begin{aligned}H_{\text{s.r.}} &= \sum_{k>k_c}\tfrac{1}{2}M_k^2(\rho_k^*\rho_k-N)\\ &= \sum_{i,j\neq i}\sum_{k>k_c}\frac{2\pi e^2}{\Omega k^2}\exp[i\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)].\end{aligned} \tag{4.64}$$

An elementary calculation (see WM, §10.5) shows that this may be expressed in the form

$$H_{\text{s.r.}} = \tfrac{1}{2}\sum_{i\neq j}^{N}\sum^{N}\frac{e^2}{r_{ij}}F(k_c r_{ij}), \tag{4.65}$$

where

$$F(k_c r) = 1-\frac{2}{\pi}\int_0^{k_c r}\frac{\sin x}{x}\,\mathrm{d}x. \tag{4.66}$$

The potential energy of the effective interaction between two particles a distance r apart has thus been changed from e^2/r to

$$\frac{e^2}{r}F(k_c r). \tag{4.67}$$

The latter is the potential energy of a screened interaction, which is virtually zero for $k_c r > 2$. In other words, the effective range of the interaction is about $2k_c^{-1}$, or, as we shall see, about 2 or 3 Å.

Of course, the validity of the approximate Hamiltonian (4.63) *depends* upon the value of k_c^{-1} being about 1 Å *or more*, while, if $H_{\text{s.r.}}$ is truly to be a short-range interaction, we must have the value of k_c^{-1} about 1 Å *or less*. It thus appears that an optimum value of k_c^{-1} is about 1 Å, as suggested by the classical theory. We shall now show that this is confirmed by the quantum-mechanical theory.

4.3. The Energy of the Ground State

We have already remarked that the eigenvalues of the Hamiltonian are unchanged by the transformation (otherwise the latter would not be permissible). This is easy to see, since, from the Schrödinger equation

$$H\Psi = E\Psi,$$

we have immediately

$$U^{-1}HUU^{-1}\Psi = EU^{-1}\Psi, \tag{4.68}$$

or

$$\mathscr{H}\Psi_{\text{new}} = E\Psi_{\text{new}}, \tag{4.69}$$

with the same E as before. We shall only be interested in the ground state, so let us take Ψ_{new} and E to be the wave function and energy of the ground state (at the absolute zero of temperature).

With $\mathscr{H}$ given by (4.63) the eq. (4.69) is separable, and Ψ_{new} may be written as a product of a function of the plasma coordinates and a function of the particle coordinates. The plasma oscillations will, of course, be in their ground states, and, since the particle interaction is of such short range, it should be a good first approximation to treat the particles as if they were non-interacting. Hence, approximately, we may write

$$\Psi_{\text{new}} = G(\ldots Q_k \ldots)\Phi_0(\ldots \boldsymbol{x}_i \ldots), \tag{4.70}$$

where G is a product of harmonic oscillator ground-state wave functions and Φ_0 is, as usual, a single determinant of free-electron functions.

The ground-state energy is thus, approximately, the sum of the zero-point energies of the plasma oscillations ($\frac{1}{2}\hbar\omega_{\text{p}}$ for each) and the energy of the particles. If we denote the energy per electron by E_{BP} in this approximation, we have

$$\begin{aligned} E &= NE_{\text{BP}} \\ &= NE_{\text{F}} + \sum_{k<k_{\text{c}}} \left(\frac{\hbar\omega_{\text{p}}}{2} - \frac{2\pi e^2 N}{\Omega k^2}\right) + \langle\Phi_0|H_{\text{s.r.}}|\Phi_0\rangle, \end{aligned} \tag{4.71}$$

the last term being the first-order perturbation energy, or exchange energy, due to $H_{\text{s.r.}}$.

If we replace the sum by an integral, the second term on the right-hand side becomes

$$\frac{\Omega}{8\pi^3}\int_0^{k_{\text{c}}} \left(\frac{\hbar\omega_{\text{p}}}{2} - \frac{2\pi e^2 N}{\Omega k^2}\right) 4\pi k^2\,\mathrm{d}k = \frac{\Omega\hbar\omega_{\text{p}} k_{\text{c}}^3}{12\pi^2} - \frac{Ne^2 k_{\text{c}}}{\pi}. \tag{4.72}$$

Now, from (3.65),

$$k_{\text{F}}^3 = \frac{3\pi^2 N}{\Omega}, \tag{4.73}$$

so that, if we set $\beta = k_{\text{c}}/k_{\text{F}}$ as before, (4.72) becomes

$$\frac{N\hbar\omega_{\text{p}}}{4}\beta^3 - \frac{Ne^2 k_{\text{F}}}{\pi}\beta. \tag{4.74}$$

$H_{\text{s.r.}}$, eq. (4.64), differs from the whole Coulomb interaction H', eq. (3.60), only in that the $\boldsymbol{k}$-sum has the condition $k > k_{\text{c}}$. We may therefore evaluate $\langle\Phi_0|H_{\text{s.r.}}|\Phi_0\rangle$ by the same method, apart from the final integration, as was used to evaluate $\langle\Phi_0|H'|\Phi_0\rangle$ in §3.5. We obtain

$$\langle\Phi_0|H_{\text{s.r.}}|\Phi_0\rangle = -\frac{4\pi e^2}{\Omega} \sum_{\substack{\boldsymbol{p},\boldsymbol{q} \\ (p,q<k_{\text{F}}, \\ |\boldsymbol{p}-\boldsymbol{q}|>k_{\text{c}})}} \frac{1}{|\boldsymbol{p}-\boldsymbol{q}|^2}. \tag{4.75}$$

Replacing the sums by integrals, we find

$$\langle\Phi_0|H_{\text{s.r.}}|\Phi_0\rangle = -\frac{4\pi e^2}{\Omega}\left(\frac{\Omega}{8\pi^3}\right)^2 \iint \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}}{|\boldsymbol{p}-\boldsymbol{q}|^2}, \tag{4.76}$$

the region of integration being given by

$$p, q < k_{\text{F}}, \quad |\boldsymbol{p}-\boldsymbol{q}| > k_{\text{c}}.$$

The evaluation of the integral is straightforward but lengthy (see WM, §10.6), and we merely quote the result

$$\langle\Phi_0|H_{\text{s.r.}}|\Phi_0\rangle = -\frac{3Ne^2 k_{\text{F}}}{4\pi}\left(1 - \frac{4\beta}{3} + \frac{\beta^2}{2} - \frac{\beta^4}{48}\right). \tag{4.77}$$

Substituting (4.77) and (4.74) in (4.71) gives, for the average energy per electron,

$$\begin{aligned} E_{\text{BP}} = E_{\text{F}} &+ \frac{\hbar\omega_{\text{p}}}{4}\beta^3 - \frac{e^2 k_{\text{F}}}{\pi}\beta \\ &- \frac{3e^2 k_{\text{F}}}{4\pi}\left(1 - \frac{4\beta}{3} + \frac{\beta^2}{2} - \frac{\beta^4}{48}\right). \end{aligned} \tag{4.78}$$

Expressing energy in rydbergs and length in Bohr units, and making use of eqs. (3.65), (3.67), (3.68) and (4.10), we have

$$\begin{aligned} E_{\text{BP}} &= \frac{2.21}{r_{\text{s}}^2} + \frac{\sqrt{3}}{2r_{\text{s}}^{3/2}}\beta^3 - \frac{0.916}{r_{\text{s}}}\left(1 + \frac{\beta^2}{2} - \frac{\beta^4}{48}\right)\text{ry} \\ &= \frac{2.21}{r_{\text{s}}^2} - \frac{0.916}{r_{\text{s}}} - \frac{0.458}{r_{\text{s}}}\beta^2 + \frac{0.866}{r_{\text{s}}^{3/2}}\beta^3 + \frac{0.019}{r_{\text{s}}}\beta^4\,\text{ry}. \end{aligned} \tag{4.79}$$

Apart from the fact that we have made certain assumptions about its order of magnitude, β is at present an arbitrary parameter, so that it would seem perfectly legitimate to choose β by minimizing E_{BP}. It must be remembered, however, that β, as well as fixing the number of plasma variables, also determines the number of subsidiary conditions, and one would expect to have to take these into account in any minimization process. Furthermore, we have ignored the subsidiary conditions in constructing our approximate wave function (4.70). Bohm, Huang and Pines† claim to show, however, that the ground-state wave function is independent of the subsidiary conditions, and *faute de mieux* we shall accept this.

---

† Bohm, D., Huang, K. and Pines, D., 1957, Phys. Rev. **107**, 71.

Dropping the last term of E_{BP}, which is of a smaller order than the energy due to H_{int}, we find

$$\frac{\partial E_{\mathrm{BP}}}{\partial \beta}=\frac{2.598}{r_s^{3/2}}\beta^2-\frac{0.916}{r_s}\beta. \tag{4.80}$$

Equating this to zero gives

$$\beta = 0.353 r_s^{1/2} \tag{4.81}$$

or

$$k_c = 0.353 r_s^{1/2} k_F = 0.677 r_s^{-1/2}, \tag{4.82}$$

where r_s is in Bohr units. For sodium, with $r_s = 4$ Bohr units, this gives $\beta = 0.71$ and $k_c^{-1} = 2.95$ Bohr units $= 1.56$ Å. This value is consistent with our previous assumptions about k_c.

Of course, E_{BP} is not the total energy per electron, since it does not include the effect of $H_{\mathrm{s.r.}}$ upon the wave function. In other words, it contains only the *first-order* perturbation energy due to $H_{\mathrm{s.r.}}$. We must now consider the higher-order corrections.

4.4. The Correlation Energy†

The effect of the Coulomb repulsion between electrons is to correlate the electronic motions in such a way as to reduce the probability of two electrons closely approaching each other. Such correlations among the electronic motions may be called *Coulomb correlations*, to distinguish them from another type of correlations, which are due to the Pauli principle. The Hartree–Fock method, for example, takes account of the latter, but takes no account at all of Coulomb correlations.

The *correlation energy* is defined as *the total energy, calculated with proper allowance for Coulomb correlations, minus the Hartree–Fock energy*. We shall denote the average correlation energy per electron by W.

In the work of the previous section we may say that *long-range* Coulomb correlations are accounted for by the plasma oscillations, but *short-range* Coulomb correlations, those due to $H_{\mathrm{s.r.}}$, are neglected. In other words, E_{BP} already includes part of the correlation energy, which we shall call the *long-range correlation energy* and denote by W_l per electron. The remainder of the correlation energy, not included in E_{BP}, we shall call the *short-range correlation energy* and denote by W_s.

---

† See WM, §§6.3, 9.5 and 10.7 for an elementary account.

4.4.1. *The long-range correlation energy*

The Hartree–Fock energy per electron, given in eq. (3.79), is

$$E_{\mathrm{HF}}=\frac{2.21}{r_s^2}-\frac{0.916}{r_s}\,\mathrm{ry}. \tag{4.83}$$

Subtracting this from E_{BP}, we have

$$\begin{aligned} W_l &= E_{\mathrm{BP}} - E_{\mathrm{HF}} \\ &= -\frac{0.458}{r_s}\beta^2+\frac{0.866}{r_s^{3/2}}\beta^3+\frac{0.019}{r_s}\beta^4\,\mathrm{ry}. \end{aligned} \tag{4.84}$$

With the value of β given in (4.81), this becomes

$$W_l = -0.019 + 0.0003 r_s\,\mathrm{ry}. \tag{4.85}$$

This is negative, as expected, since the correlation energy must be negative (i.e. the correct total energy must lie below that of the Hartree–Fock approximation) and W_l is just part of this energy. The value of β has, in fact, been chosen to give the lowest value of E_{BP}, that is, the *largest* value of $-W_l$. This is a reasonable procedure, since it gives the greatest improvement of the total energy on that of the Hartree–Fock theory, without taking into account the more awkward short-range correlations.

4.4.2. *The short-range correlation energy*

It has already been mentioned that E_{BP} contains only the *first-order* perturbation correction due to $H_{\mathrm{s.r.}}$. This is because the particle wave function Φ_0, eq. (4.70), is a single determinant of free-electron functions, which would only be correct if the particles did not interact. The particle wave function should, in fact, be an infinite sum of determinants, with coefficients determined by $H_{\mathrm{s.r.}}$. The method of procedure is that described in §3.6, except that the perturbation is now the short-range interaction instead of the whole Coulomb interaction. We found that the second-order perturbation energy due to the latter diverged, but the new interaction is of such short range that we may expect the second-order perturbation energy no longer to diverge, but in fact to be quite small. In this case, higher-order perturbation corrections might be neglected. We shall anticipate this result and take the short-range correlation energy per electron to be given by

$$W_s = -\frac{1}{N}\sum_{n=1}^{\infty}\frac{|\langle\Phi_n|H_{\mathrm{s.r.}}|\Phi_0\rangle|^2}{E_n-E_0}, \tag{4.86}$$

using the same notation as in eq. (3.80).

Since, as was indicated when obtaining eq. (4.75), $H_{\text{s.r.}}$, eq. (4.64), differs from the whole Coulomb interaction H', eq. (3.60), only in that the $\boldsymbol{k}$-sum is over $k > k_{\text{c}}$, instead of over all $\boldsymbol{k}$, we can adapt the results of §3.6 immediately to the present case merely by imposing the condition $k > k_{\text{c}}$. Thus, for example, in terms of the creation and destruction operators for the free-electron functions, we have, from (3.81),

$$\langle \Phi_n | H_{\text{s.r.}} | \Phi_0 \rangle = \sum_{\substack{\boldsymbol{p},\boldsymbol{q},\boldsymbol{k}(k>k_c)\\ \sigma,\sigma'}} \frac{2\pi e^2}{k^2} \langle \Phi_n | c^{\dagger}_{\boldsymbol{p}+\boldsymbol{k},\sigma} c^{\dagger}_{\boldsymbol{q}-\boldsymbol{k},\sigma'} c_{\boldsymbol{q}\sigma'} c_{\boldsymbol{p}\sigma} | \Phi_0 \rangle . \tag{4.87}$$

Now, in §3.6, since we were only concerned to show that $\Delta E^{(2)}$ diverges, we simply considered the contribution from antiparallel spin correlations, which itself diverges. Here again we intend to consider only antiparallel spin correlations, but this now requires justification, because we hope to obtain a finite numerical value for W_{s}. The justification lies in the effect of those correlations, mentioned previously, which are due, not to the Coulomb interaction, but to the Pauli principle as embodied in the use of determinantal wave functions. These correlations occur only between electrons with parallel spins, and the net effect is that any electron during its motion appears to be surrounded by a spherical hole, called the *exchange or Fermi hole*†, in the distribution of electrons with parallel spins. The radius of the Fermi hole is roughly r_{s}, that is, a distance comparable with the screening distance, about k_{c}^{-1}, of the short-range interaction for actual metallic densities‡. It follows, therefore, that electrons with parallel spins, purely as a result of the Pauli principle, will rarely come close enough to each other for the short-range interaction to have much effect. We shall therefore neglect the contribution of parallel spin interactions to W_{s}, which, from eq. (3.88), is then given by

$$W_{\text{s}} = -\frac{16\pi^2 e^4 m}{N\hbar^2 \Omega^2} \sum_{\boldsymbol{p}\cdot\boldsymbol{q}\cdot\boldsymbol{k}} \frac{1}{k^4 \boldsymbol{k} \cdot (\boldsymbol{p} - \boldsymbol{q} + \boldsymbol{k})}, \tag{4.88}$$

where $p,q < k_{\text{F}}$, $k > k_{\text{c}}$, and $|\boldsymbol{p}+\boldsymbol{k}|, |\boldsymbol{q}-\boldsymbol{k}| > k_{\text{F}}$. This differs from (3.88) only in that here we have $k > k_{\text{c}}$.

---

† A detailed explanation of this is given in WM, §§6.3 and 7.6.

‡ From (4.82), we have, for $r_{\text{s}} = 2$, $k_{\text{c}}^{-1} \approx 2$, and for $r_{\text{s}} = 4$, $k_{\text{c}}^{-1} \approx 3$ (all in Bohr units). It should be noted, however, that, for $r_{\text{s}} = 0.01$, say, $k_{\text{c}}^{-1} \approx 0.15$, and, as $r_{\text{s}} \to 0$, the radius of the Fermi hole becomes negligible compared with the screening distance of the short-range interaction. The approximation used here is not expected, therefore, to be valid in the high density limit.

Transforming this to an integral, as in eq. (3.90), we find

$$W_{\text{s}} = -\frac{3}{16\pi^5} \int \frac{\mathrm{d}\boldsymbol{k}}{k^4} \int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{\boldsymbol{k} \cdot (\boldsymbol{p} - \boldsymbol{q} + \boldsymbol{k})} \text{ry}, \tag{4.89}$$

the region of integration being given by $p,q < 1$, $k > \beta$, and $|\boldsymbol{p}+\boldsymbol{k}|, |\boldsymbol{q}-\boldsymbol{k}| > 1$, since $k_{\text{c}} = \beta$ when measured in units of k_{F}. It is immediately obvious that there is no logarithmic divergence at $k = 0$ in this case, since k is never zero.

We shall not calculate the integral in (4.89) directly, but, by the same method used to show the divergence of (3.90), we shall obtain W_{s} in the approximate form

$$W_{\text{s}} = A \log r_{\text{s}} + B, \tag{4.90}$$

where A and B are constants. We assume that† $\beta \ll 1$, so that the greatest contribution to the integral will come from $k \approx \beta$, for which $p \approx 1, q \approx 1$. From eqs. (3.97) and (3.100) we then obtain

$$\int \mathrm{d}\boldsymbol{q} \int \frac{\mathrm{d}\boldsymbol{p}}{\boldsymbol{k} \cdot (\boldsymbol{p} - \boldsymbol{q} + \boldsymbol{k})} \approx (2\pi)^2 k \int_0^1 y\,\mathrm{d}y \int_0^1 \frac{x\,\mathrm{d}x}{x+y}, \tag{4.91}$$

and, substituting in (4.89),

$$\begin{aligned} W_{\text{s}} &\approx -\frac{3(2\pi)^2}{16\pi^5} \int \frac{\mathrm{d}k}{k^3} \int_0^1 y\,\mathrm{d}y \int_0^1 \frac{x\,\mathrm{d}x}{x+y} \\ &= -\frac{3}{\pi^2} \int_\beta^1 \frac{\mathrm{d}k}{k} \int_0^1 y\,\mathrm{d}y \int_0^1 \frac{x\,\mathrm{d}x}{x+y} + \delta \\ &= \frac{2}{\pi^2}(1 - \log 2) \log \beta + \delta. \end{aligned} \tag{4.92}$$

Here we have arbitrarily cut off k at an upper limit of unity, compensating for this by the addition of a constant δ, which is expected to be small.

If we now assume

$$\beta \propto r_{\text{s}}^{1/2}, \tag{4.93}$$

as in (4.81) (the value of the constant of proportionality is immaterial as far as the logarithmic term is concerned), we obtain

$$\begin{aligned} W_{\text{s}} &\approx \frac{1}{\pi^2}(1 - \log 2) \log r_{\text{s}} + B \\ &= 0.0311 \log r_{\text{s}} + B, \end{aligned} \tag{4.94}$$

---

† For β given by (4.81) and the values of r_{s} occurring in actual metals, this is not strictly true. For example, $\beta = 0.7$ for sodium. However, this is not likely to affect the logarithmic term.

where B is a new constant. The direct evaluation of the integral in (4.89) is rather tedious. It has been carried out by Pines, and we quote the value of B obtained by him, namely,

$$B = -0.091 \text{ ry}. \tag{4.95}$$

Pines also obtained an additional very small term proportional to r_s, which we shall neglect. If we also neglect the equally small term proportional to r_s in W_1, eq. (4.85), we obtain, finally,

$$W = W_1 + W_s = 0.0311 \log r_s - 0.110 \text{ ry}. \tag{4.96}$$

After various refinements, including a more careful treatment of the interaction between the electrons and the plasma oscillations, Nozières and Pines† obtained the slightly amended formula

$$W = 0.0311 \log r_s - 0.115 \text{ ry}. \tag{4.97}$$

The change in the value of the constant is hardly significant, but this is the formula we shall compare with the results obtained by the direct perturbation treatment described in ch. 9. It is certainly incorrect at very high and at very low densities, but in the region of actual metallic densities it yields values of W which are of the size needed in cohesive energy calculations. In the case of sodium, for example, with $r_s = 4$ Bohr units, it gives $W = -0.072$ ry, compared with the value -0.075 ry given by the Wigner formula (see WM, p. 273)

$$W = -\frac{0.88}{r_s + 7.8} \text{ ry}, \tag{4.98}$$

which has been used successfully for many years in calculating cohesive energies.

From a rigorous mathematical point of view the foregoing calculation of the correlation energy is defective in many ways. None the less, it is undoubtedly the simplest available method of obtaining the desired results. It has one great advantage over more mathematically sophisticated calculations, in that it presents a simple and acceptable *physical picture* of the effects of the electronic interactions. Even in a qualitative sense this is valuable, but the plasma theory has, of course, yielded many useful quantitative and semi-quantitative results (particularly in the fields of band theory and electron penetration phenomena) which we do not wish to go into here. Nor do we wish to include accounts of the several attempts to establish the theory on sounder foundations and to extend its range of applicability. These are very complicated, and, as far as the correlation energy, which is our main interest, is concerned, add very little to the results obtained above. Some further remarks about the plasma theory will be made at the end of ch. 9.

---

† Nozières, P. and Pines, D., 1958, Phys. Rev. **111**, 442.

CHAPTER 5

THE SCHRÖDINGER, HEISENBERG AND INTERACTION PICTURES

5.1. Time Dependence and the Schrödinger Picture

The basic problem which concerns us is that of finding the energy levels of a many-electron system, which are the eigenvalues of the Schrödinger equation

$$H\Psi = E\Psi. \tag{5.1}$$

The Hamiltonian H, which was introduced in §1.2, does not contain the time explicitly, so that the time dependence of the wave function Ψ is trivial. The *time-dependent* Schrödinger equation is (see WM, §11.1)

$$H\Psi(t) = \mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(t), \tag{5.2}$$

where we have simply written $\Psi(t)$ for the wave function at time t. In the remainder of the book it will, in fact, be convenient to assume a system of units in which $\hbar = 1$. This will avoid the appearance of a factor $1/\hbar$ in the many exponents which occur. In such a system of units, eq. (5.2) becomes

$$H\Psi(t) = \mathrm{i}\frac{\partial}{\partial t}\Psi(t), \tag{5.3}$$

and this may be integrated formally to give

$$\Psi(t) = \exp(-\mathrm{i}Ht)\Psi(0), \tag{5.4}$$

the operator exp $(-\mathrm{i}Ht)$ being defined by

$$\exp(-\mathrm{i}Ht) = 1-\mathrm{i}Ht+\frac{(\mathrm{i}Ht)^2}{2!}-\frac{(\mathrm{i}Ht)^3}{3!}+\dots \tag{5.5}$$

If $\Psi(0)$, the wave function at time $t = 0$, is an eigenfunction of eq. (5.1), with corresponding eigenvalue E, then†

$$\Psi(t) = \exp(-\mathrm{i}Et)\Psi(0), \tag{5.6}$$

which is the familiar expression for the time-dependent wave function of a stationary state of a conservative system.

This formulation of the problem, in which the operator H and other operators, such as momentum, generally do not depend explicitly upon time, but the wave function is time-dependent, is known as the *Schrödinger picture*.

This is not the only formulation of the problem, however, and others are sometimes more convenient. Any unitary transformation of both operators and wave functions gives an equally valid formulation. This is because, as shown in Appendix II, such a transformation leaves the matrix elements of all operators unchanged, and the physically significant properties of the system depend only upon these matrix elements. Two formulations of particular interest are known as the *Heisenberg* and *interaction pictures*, in the former of which the time dependence is transferred from the wave functions to the operators, and in the latter both operators and wave functions become time-dependent.

---

† In case there is doubt about the validity of this step, it may be helpful to verify it in detail. We have

$$H^2\Psi = H(H\Psi) = HE\Psi = EH\Psi = E^2\Psi.$$

Similarly,

$$H^3\Psi = E^3\Psi,$$

and so on, for all positive integral powers of H. Also, H and t commute, since the differential operators appearing in H act only on the space coordinates. Hence, from (5.5),

$$\begin{aligned}\exp(-\mathrm{i}Ht)\Psi(0) &= \Psi(0)-\mathrm{i}tH\Psi(0)+\frac{(\mathrm{i}t)^2}{2!}H^2\Psi(0)-\dots\\ &= \left[1-\mathrm{i}tE+\frac{(\mathrm{i}t)^2}{2!}E^2-\dots\right]\Psi(0)\\ &= \exp(-\mathrm{i}Et)\Psi(0).\end{aligned}$$

5.2. The Heisenberg Picture

We shall now designate the Schrödinger wave function by $\Psi_{\mathrm{S}}(t)$ to distinguish it from the transformed functions which we shall introduce, so that eq. (5.3) becomes

$$H\Psi_{\mathrm{S}}(t) = \mathrm{i}\frac{\partial}{\partial t}\Psi_{\mathrm{S}}(t), \tag{5.7}$$

and eq. (5.4) becomes

$$\Psi_{\mathrm{S}}(t) = \exp(-\mathrm{i}Ht)\Psi_{\mathrm{S}}(0). \tag{5.8}$$

Now let us consider the operator

$$O = \exp(\mathrm{i}Ht). \tag{5.9}$$

This is a unitary operator, since H is Hermitian and t is real, so that, by the rules of Appendix I,

$$O^{\dagger} = \exp(-\mathrm{i}Ht), \tag{5.10}$$

and

$$OO^{\dagger} = O^{\dagger}O = 1. \tag{5.11}$$

We propose to use this operator O to effect a unitary transformation of the wave functions and operators of the Schrödinger picture. The result is known as the *Heisenberg picture*, and we shall designate a wave function in this picture by Ψ_{H}. Then

$$\Psi_{\mathrm{H}} = O\Psi_{\mathrm{S}}(t) = \exp(\mathrm{i}Ht)\Psi_{\mathrm{S}}(t) = \Psi_{\mathrm{S}}(0), \tag{5.12}$$

from (5.8). The Heisenberg function Ψ_{H} is thus independent of time.

Now let us consider the transformation of a Schrödinger operator A_{S} into a Heisenberg operator A_{H}. We have, from eq. (II.3), Appendix II,

$$A_H = OA_{\mathrm{S}}O^{\dagger} = \exp(\mathrm{i}Ht)A_{\mathrm{S}}\exp(-\mathrm{i}Ht). \tag{5.13}$$

Thus, even if A_{S} does not depend upon the time explicitly, A_{H} generally does so. Assuming that H and A_{S} are independent of t, we have†

$$\begin{aligned}\frac{\partial A_{\mathrm{H}}}{\partial t} &= \mathrm{i}HA_{\mathrm{H}} + \exp(\mathrm{i}Ht)A_{\mathrm{S}}[-\mathrm{i}H\exp(-\mathrm{i}Ht)]\\ &= \mathrm{i}(HA_{\mathrm{H}} - A_{\mathrm{H}}H), \end{aligned}\tag{5.14}$$

since H and O commute.

The equation of motion of the operator A_{H} is thus

$$\mathrm{i}\frac{\partial A_{\mathrm{H}}}{\partial t} = [A_{\mathrm{H}}, H]. \tag{5.15}$$

It should be noted that *the Hamiltonian is the same in both the Schrödinger and the Heisenberg pictures*, since

$$H_{\mathrm{H}} = \exp(\mathrm{i}Ht)H\exp(-\mathrm{i}Ht) = H. \tag{5.16}$$

The equation satisfied by Ψ_{H} is thus the time-independent Schrödinger equation (5.1).

5.3. The Interaction Picture

We have noted that in the Schrödinger picture the wave function depends upon t and the operators generally do not, while in the Heisenberg picture the wave function does not depend upon t and the operators generally do. There is another representation which proves convenient when a system is subject to a time-dependent perturbation. This is called the *interaction picture*, and in it both the wave functions and the operators depend explicitly upon the time.

Let us suppose that the Hamiltonian consists of two parts:

$$H = H_0 + H'(t), \tag{5.17}$$

where H_0 is, as before, independent of t, while H' may depend explicitly upon t. The time-dependent Schrödinger equation is then

$$\mathrm{i}\frac{\partial}{\partial t}\Psi_{\mathrm{S}}(t) = (H_0 + H')\Psi_{\mathrm{S}}(t). \tag{5.18}$$

---

† The derivative of an operator A which depends explicitly upon a parameter t is defined in the same way as for a function, that is,

$$\frac{\partial A}{\partial t} = \lim_{k\to 0}\frac{A(t+k) - A(t)}{k}.$$

It follows that products of such operators may be differentiated by the same rule as for products of functions, provided that, when they do not commute, the order of the operators is preserved, for example,

$$\frac{\partial}{\partial t}(AB) = \frac{\partial A}{\partial t}B + A\frac{\partial B}{\partial t}.$$

Similarly, if Ψ is a function of t,

$$\frac{\partial}{\partial t}(AB\Psi) = \frac{\partial A}{\partial t}B\Psi + A\frac{\partial B}{\partial t}\Psi + AB\frac{\partial \Psi}{\partial t}.$$

We now make a unitary transformation by means of the operator $\exp(iH_0t)$, such that the Schrödinger wave function $\Psi_S(t)$ is transformed into a function $\Psi_I(t)$, where

$$\Psi_I(t) = \exp(iH_0t)\Psi_S(t), \tag{5.19}$$

and a Schrödinger operator A_S is transformed into A_I, where

$$A_I = \exp(iH_0t)A_S\exp(-iH_0t). \tag{5.20}$$

Differentiating (5.19) gives

$$\begin{aligned}\frac{\partial\Psi_I}{\partial t} &= iH_0\Psi_I + \exp(iH_0t)\frac{\partial\Psi_S}{\partial t}\\ &= iH_0\Psi_I - i\exp(iH_0t)(H_0+H')\Psi_S\\ &= -i\exp(iH_0t)H'\exp(-iH_0t)\Psi_I,\end{aligned} \tag{5.21}$$

using (5.18). If we write

$$H_I = \exp(iH_0t)H'\exp(-iH_0t), \tag{5.22}$$

the equation of motion for Ψ_I is thus

$$i\frac{\partial\Psi_I}{\partial t} = H_I\Psi_I. \tag{5.23}$$

This is similar to the time-dependent Schrödinger equation, with H_I replacing H. However, it must be remembered that H_I is not the transformed *total* Hamiltonian in the interaction picture (this, in fact, is H_0+H_I), but only the transformed perturbation or interaction term. It would perhaps be more consistent to designate it by H'_I, but as this symbol will occur many times in complicated expressions it will be convenient to drop the prime.

It may be seen that the time dependence of Ψ_I is determined by both H' and H_0, while, if A_S is time-independent, the time dependence of the transformed operator A_I is determined solely by H_0. We note, in particular, that Ψ_I is independent of t if H' is zero, that is to say, if there is no interaction term – the interaction picture is then the same as the Heisenberg picture (obviously so, since then $H = H_0$).

We now define the operator $U(t, t')$, called the *time-development operator*†, by

$$\Psi_I(t) = U(t, t')\Psi_I(t'). \tag{5.24}$$

---

† This name is often given to an operator defined similarly in the Schrödinger picture. However, in this book we shall only consider time-development operators in the interaction picture.

It follows from this definition that

$$\begin{aligned}\Psi_I(t) &= U(t, t)\Psi_I(t) = U(t, t'')\Psi_I(t'')\\ &= U(t, t')U(t', t'')\Psi_I(t''),\end{aligned} \tag{5.25}$$

and

$$\Psi_I(t') = U^{-1}(t, t')\Psi_I(t) = U(t', t)\Psi_I(t), \tag{5.26}$$

so that we have immediately the three basic properties

$$U(t, t) = 1, \tag{5.27}$$

$$U(t, t'') = U(t, t')U(t', t''), \tag{5.28}$$

and

$$U^{-1}(t, t') = U(t', t). \tag{5.29}$$

The fact that $U(t, t')$ *is a unitary operator* follows from the requirement that the normalization of $\Psi_I(t)$ be independent of t (that is, once normalized, the function must remain so for all time). Thus, from eqs. (5.24) and (I.2), Appendix I,

$$\begin{aligned}\langle\Psi_I(t)|\Psi_I(t)\rangle &= \langle U(t, t')\Psi_I(t')|U(t, t')\Psi_I(t')\rangle\\ &= \langle U^\dagger(t, t')U(t, t')\Psi_I(t')|\Psi_I(t')\rangle\\ &= \langle\Psi_I(t')|\Psi_I(t')\rangle,\end{aligned} \tag{5.30}$$

and, since this is true for all the functions $\Psi_I(t)$, we may deduce that

$$U^\dagger(t, t')U(t, t') = 1. \tag{5.31}$$

Finally, substituting (5.24) in (5.23), we obtain

$$i\frac{\partial}{\partial t}[U(t, t')\Psi_I(t')] = H_I(t)U(t, t')\Psi_I(t')$$

or

$$i\frac{\partial}{\partial t}U(t, t') = H_I(t)U(t, t'), \tag{5.32}$$

which is the equation of motion for $U(t, t')$.

A closed expression can be found for $U(t, t')$ quite simply *when (and only when) H does not depend explicitly upon the time*. In this case, it follows from eq. (5.8) that

$$\Psi_S(0) = \exp(iHt)\Psi_S(t) = \exp(iHt')\Psi_S(t'). \tag{5.33}$$

Hence

$$\Psi_S(t) = \exp\,[-\mathrm{i}H(t-t')]\Psi_S(t'), \tag{5.34}$$

and, using eq. (5.19),

$$\begin{aligned}\Psi_I(t) &= \exp\,(\mathrm{i}H_0t)\exp\,[-\mathrm{i}H(t-t')]\Psi_S(t')\\ &= \exp\,(\mathrm{i}H_0t)\exp\,[-\mathrm{i}H(t-t')]\exp\,(-\mathrm{i}H_0t')\Psi_I(t'),\end{aligned} \tag{5.35}$$

so that, in this case,

$$U(t,t') = \exp\,(\mathrm{i}H_0t)\exp\,[-\mathrm{i}H(t-t')]\exp\,(-\mathrm{i}H_0t'). \tag{5.36}$$

It is a trivial matter to verify that $U(t,t')$ is unitary. By the rules of Appendix I, since H is Hermitian,

$$\begin{aligned}U^\dagger(t,t') &= \exp\,(\mathrm{i}H_0t')\exp\,[\mathrm{i}H(t-t')]\exp\,(-\mathrm{i}H_0t)\\ &= U(t',t) = U^{-1}(t,t').\end{aligned} \tag{5.37}$$

In the following chapters we shall be particularly concerned with an interaction term H' depending explicitly upon t, in which case (5.36) does not apply (and, even when it does, it is not very helpful), so we must now find some convenient method of solving eq. (5.32).

5.4. The Integral Equation for $U(t, t')$

We first replace eq. (5.32) by the integral equation

$$U(t,t') = 1-\mathrm{i}\int_{t'}^{t} H_I(t_1)U(t_1,t')\,\mathrm{d}t_1. \tag{5.38}$$

Differentiation with respect to t gives (5.32) immediately. Also, substitution of t for t' gives the initial condition (5.27).

Eq. (5.38) may now be solved by iteration. Substituting t_1 for t, and changing the variable of integration to t_2, we have

$$U(t_1,t') = 1-\mathrm{i}\int_{t'}^{t_1} H_I(t_2)U(t_2,t')\,\mathrm{d}t_2, \tag{5.39}$$

so that

$$\begin{aligned}U(t,t') &= 1-\mathrm{i}\int_{t'}^{t} H_I(t_1)\left\{1-\mathrm{i}\int_{t'}^{t_1} H_I(t_2)U(t_2,t')\,\mathrm{d}t_2\right\}\mathrm{d}t_1\\ &= 1-\mathrm{i}\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\\ &\quad +(-\mathrm{i})^2\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\int_{t'}^{t_1} H_I(t_2)U(t_2,t')\,\mathrm{d}t_2.\end{aligned} \tag{5.40}$$

Similarly,

$$U(t_2,t') = 1-\mathrm{i}\int_{t'}^{t_2} H_I(t_3)U(t_3,t')\,\mathrm{d}t_3, \tag{5.41}$$

so that

$$\begin{aligned}U(t,t') &= 1-\mathrm{i}\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1+(-\mathrm{i})^2\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\int_{t'}^{t_1} H_I(t_2)\,\mathrm{d}t_2\\ &\quad +(-\mathrm{i})^3\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\int_{t'}^{t_1} H_I(t_2)\,\mathrm{d}t_2\int_{t'}^{t_2} H_I(t_3)U(t_3,t')\,\mathrm{d}t_3.\end{aligned} \tag{5.42}$$

Proceeding in this way, we find

$$U(t,t') = 1+\sum_{n=1}^{\infty}(-\mathrm{i})^n\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\int_{t'}^{t_1} H_I(t_2)\,\mathrm{d}t_2\ldots\int_{t'}^{t_{n-1}} H_I(t_n)\,\mathrm{d}t_n. \tag{5.43}$$

Consideration of the regions of integration for the successive multiple integrals shows that, if $t > t'$, then we must have

$$t \geqslant t_1 \geqslant t_2 \ldots \geqslant t_n \geqslant t'.$$

It must be understood that the operators $H_I(t_1), H_I(t_2), \ldots$, taken at different instants $t_1, t_2, \ldots$ will not generally commute. Thus, from (5.22),

$$\begin{aligned}&[H_I(t_1), H_I(t_2)]\\ &\quad= \exp\,(\mathrm{i}H_0t_1)H'(t_1)\exp\,[-\mathrm{i}H_0(t_1-t_2)]H'(t_2)\exp\,(-\mathrm{i}H_0t_2)\\ &\qquad-\exp\,(\mathrm{i}H_0t_2)H'(t_2)\exp\,[-\mathrm{i}H_0(t_2-t_1)]H'(t_1)\exp\,(-\mathrm{i}H_0t_1).\end{aligned} \tag{5.44}$$

Clearly, if $t_1 \neq t_2$, this will not vanish except in the trivial case where H' commutes with H_0.

In our application of perturbation theory to a many-electron system we shall, in fact, make use of the expression (5.43) for $U(t,t')$ as it stands. However, a more compact expression may be obtained with the aid of an operator, due to Dyson, which we shall call the *chronological operator*, to distinguish it from the similar time-ordering operator, due to Wick, which will be defined in the final chapter of this book. In the following section we shall define and demonstrate the use of the chronological operator, but this account may be omitted by any reader in a hurry.

5.5. The Chronological Operator

This operator, which we will denote by P (not to be confused with the permutation operator P used in the first two chapters), when applied to a product of time-dependent operators $A(t_1)B(t_2)\ldots$, taken at different times $t_1, t_2, \ldots$, rearranges the operators so that the times decrease from left to right. Thus, for the product $A(t_1)B(t_2)$, we have†

$$P[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2), & \text{if} \quad t_1 > t_2, \\ B(t_2)A(t_1), & \text{if} \quad t_2 > t_1. \end{cases} \tag{5.45}$$

We shall first use this operator in order to express the double integral in the term with $n = 2$ of the series in eq. (5.43) in a more symmetrical form. We shall prove, in fact, that

$$\int_{t'}^{t} H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1 \int_{t'}^{t_1} H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2 = \tfrac{1}{2}\int_{t'}^{t}\int_{t'}^{t} P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)]\,\mathrm{d}t_1\,\mathrm{d}t_2. \tag{5.46}$$

Assuming that $t > t'$, the regions of integration are as shown in fig. 5.1. Now,

$$\begin{aligned}
&\int_{t'}^{t}\int_{t'}^{t} P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)]\,\mathrm{d}t_1\,\mathrm{d}t_2 \\
&\quad = \int_{t'}^{t}\mathrm{d}t_1\left\{\int_{t'}^{t_1} P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)]\,\mathrm{d}t_2 + \int_{t_1}^{t} P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)]\,\mathrm{d}t_2\right\} \\
&\quad = \int_{t'}^{t}\mathrm{d}t_1\left\{\int_{t'}^{t_1} H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2 + \int_{t_1}^{t} H_{\mathrm{I}}(t_2)H_{\mathrm{I}}(t_1)\,\mathrm{d}t_2\right\},
\end{aligned} \tag{5.47}$$

since $t_1 > t_2$ in the first term and $t_2 > t_1$ in the second term.

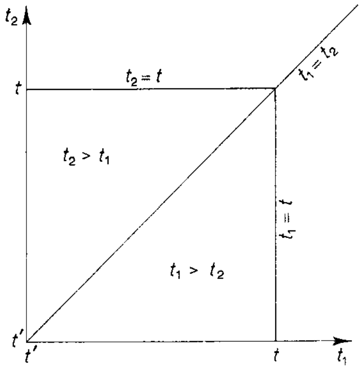


Fig. 5.1. Regions of integration for the double integral in eq. (5.46).

Changing the order of integration in the second term, for which the region of integration is the upper triangle in fig. 5.1, we find

$$\begin{aligned}
\int_{t'}^{t}\mathrm{d}t_1\int_{t_1}^{t} H_{\mathrm{I}}(t_2)H_{\mathrm{I}}(t_1)\,\mathrm{d}t_2 &= \int_{t'}^{t}\mathrm{d}t_2\int_{t'}^{t_2} H_{\mathrm{I}}(t_2)H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1 \\
&= \int_{t'}^{t}\mathrm{d}t_1\int_{t'}^{t_1} H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2,
\end{aligned} \tag{5.48}$$

where in the last step we have interchanged the symbols representing the variables of integration, which is always permissible in a definite integral. Hence

$$\begin{aligned}
\int_{t'}^{t}\int_{t'}^{t} P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)]\,\mathrm{d}t_1\,\mathrm{d}t_2 &= 2\int_{t'}^{t}\mathrm{d}t_1\int_{t'}^{t_1} H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2 \\
&= 2\int_{t'}^{t} H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1\int_{t'}^{t_1} H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2,
\end{aligned} \tag{5.49}$$

as required.

We now make an obvious generalization of this, and write

$$\begin{aligned}
U(t,t') = 1 + \sum_{n=1}^{\infty}\frac{(-\mathrm{i})^n}{n!}\int_{t'}^{t}\mathrm{d}t_1\int_{t'}^{t}\mathrm{d}t_2\ldots\int_{t'}^{t}\mathrm{d}t_n \\
\times P[H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\ldots H_{\mathrm{I}}(t_n)],
\end{aligned} \tag{5.50}$$

which we shall prove by differentiation. Thus

$$\begin{aligned}
\frac{\partial U(t,t')}{\partial t} &= -\mathrm{i}H_{\mathrm{I}}(t) \\
&\quad + \sum_{n=2}^{\infty}\frac{(-\mathrm{i})^n}{n!}\sum_{m=1}^{n}\int_{t'}^{t}\mathrm{d}t_1\ldots\int_{t'}^{t}\mathrm{d}t_{m-1}\int_{t'}^{t}\mathrm{d}t_{m+1}\ldots\int_{t'}^{t}\mathrm{d}t_n \\
&\quad \times P[H_{\mathrm{I}}(t_1)\ldots H_{\mathrm{I}}(t_{m-1})H_{\mathrm{I}}(t)H_{\mathrm{I}}(t_{m+1})\ldots H_{\mathrm{I}}(t_n)] \\
&= -\mathrm{i}H_{\mathrm{I}}(t) \\
&\quad + \sum_{n=2}^{\infty}\frac{(-\mathrm{i})^n}{n!}nH_{\mathrm{I}}(t)\int_{t'}^{t}\mathrm{d}t_1\ldots\int_{t'}^{t}\mathrm{d}t_{n-1}P[H_{\mathrm{I}}(t_1)\ldots H_{\mathrm{I}}(t_{n-1})] \\
&= -\mathrm{i}H_{\mathrm{I}}(t)\left\{1 + \sum_{q=1}^{\infty}\frac{(-\mathrm{i})^q}{q!}\int_{t'}^{t}\mathrm{d}t_1\ldots\int_{t'}^{t}\mathrm{d}t_q\right. \\
&\quad \times P[H_{\mathrm{I}}(t_1)\ldots H_{\mathrm{I}}(t_q)] \\
&= -\mathrm{i}H_{\mathrm{I}}(t)U(t,t'),
\end{aligned} \tag{5.51}$$

---

† The operator P is not defined for $t_1 = t_2$. However, if the operators $A(t_1)$, $B(t_2)$ become the same when $t_1 = t_2$, as is the case with the operators $H_{\mathrm{I}}(t_1)$, $H_{\mathrm{I}}(t_2)$, they then commute, so that for our purposes the symbol $>$ in eq. (5.45) can be replaced by $\geqslant$.

which is correct, according to eq. (5.32). In the second line we have removed $H_I(t)$ to the left of the multiple integral, since this is its correct position given by the chronological operator, and re-labelled the variables of integration from 1 to $n-1$ in each of the n multiple integrals appearing in the general term of the series, each of which has the same value. In the third line we have merely written $q = n-1$, and the sum over q is seen to be the same as the sum over n in (5.50).

Eq. (5.50) may be written compactly in the form

$$U(t, t') = P \exp\left[-\mathrm{i} \int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\right], \tag{5.52}$$

since

$$\left(\int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1\right)^n = \int_{t'}^{t} H_I(t_1)\,\mathrm{d}t_1 \ldots \int_{t'}^{t} H_I(t_n)\,\mathrm{d}t_n,$$

and P is to be applied to each term of the expansion.

The effect of the chronological operator is clearly seen here. The solution of the elementary differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)y, \tag{5.53}$$

where $f(x)$ is just a function of x and not an operator, subject to the boundary condition $y(x') = 1$, is clearly

$$y = \exp\left[\int_{x'}^{x} f(x_1)\,\mathrm{d}x_1\right]. \tag{5.54}$$

One might have been tempted, therefore, to write the formal solution of eq. (5.32), satisfying (5.27), immediately as in eq. (5.52), but without the operator P. This would have been erroneous, owing to the fact that $H_I(t)$ is an *operator* and $H_I(t_1)$, $H_I(t_2)$, with $t_1 \neq t_2$, in general do not commute.

CHAPTER 6

THE ADIABATIC HYPOTHESIS AND THE ENERGY OF THE GROUND STATE

6.1. The Adiabatic Hypothesis

The work of the previous chapter is quite general and may be applied to any system, but clearly it will be most useful in the case of a system subject to a time-dependent perturbation. It would appear to be irrelevant to the problem which concerns us, since the Hamiltonian (1.1) does not depend explicitly upon time and the system is not subject to a time-dependent perturbation. It is our intention to treat the electronic interaction term (1.3) as a perturbation, and this is time-independent. However, in order to employ the method of the previous chapter, together with the highly-developed graphical analysis to be described later, we introduce an artificial time dependence by multiplying the interaction term H' by $\mathrm{e}^{\alpha t}$, where α is a small *positive* constant† ultimately tending to zero. Then, as $t \to -\infty$, the new interaction term tends to zero and the Hamiltonian tends to just H_0, while at $t = 0$ we revert to our original Hamiltonian, $H_0 + H'$. In other words, we assume that, starting at some very large negative time with a system of non-interacting electrons, the interaction is slowly (α small) switched on, until at $t = 0$ we have the full interaction.

According to the *adiabatic hypothesis*, a particular eigenfunction of H_0, or state of the non-interacting system, will, provided the interaction is switched on sufficiently slowly, change continuously into a particular

---

† The exponential factor is frequently written $\mathrm{e}^{-\alpha|t|}$, which tends to zero as $t \to +\infty$ also. However, except in the final chapter, we shall not be concerned with positive values of t, so that $\mathrm{e}^{\alpha t}$ will suffice.

eigenfunction of H, or state of the fully-interacting system, at $t = 0$. The same applies to the corresponding eigenvalues or energy levels. This, in fact, is unlikely to be true for excited states, which are degenerate†, but is *probably* true for the ground state, which is non-degenerate. In what follows we shall be concerned exclusively with the ground state of the system and shall *assume* the truth of the adiabatic hypothesis. A sufficient justification of this will be found in the correspondence between the results and those of the Rayleigh–Schrödinger perturbation theory.

6.2. The Ground State

We denote by Φ_0 the (Heisenberg) ground-state wave function of the non-interacting system, that is, the time-independent eigenfunction of H_0 corresponding to the lowest eigenvalue E_0. Thus

$$H_0\Phi_0 = E_0\Phi_0. \tag{6.1}$$

We have seen in ch. 1 that this equation is separable, so that Φ_0 is a single determinant of one-electron functions, whose orbital factors are eigenfunctions of the equation (cf. eq. (1.8), with $\hbar = 1$)

$$-\frac{1}{2m}\nabla^2\psi + V(\boldsymbol{r})\psi = \epsilon\psi. \tag{6.2}$$

In the following work our creation and destruction operators will refer to the eigenfunctions of this equation, multiplied by the appropriate spin factors. The determinant Φ_0 contains N functions, consisting of $\frac{1}{2}N$ pairs formed from the $\frac{1}{2}N$ lowest eigenfunctions of eq. (6.2) multiplied by the spin functions α and β, respectively.

We denote by Ψ_0 the (Heisenberg) ground-state wave function of the fully-interacting system, that is, the time-independent eigenfunction of H corresponding to its lowest eigenvalue E. Thus

$$H\Psi_0 \equiv (H_0+H')\Psi_0 = E\Psi_0. \tag{6.3}$$

Writing

$$E = E_0 + \Delta E, \tag{6.4}$$

we have immediately, from eq. (1.62),

$$\Delta E = \frac{\langle\Phi_0|H'|\Psi_0\rangle}{\langle\Phi_0|\Psi_0\rangle}. \tag{6.5}$$

---

† A degenerate energy level usually splits into a number of levels under the action of a perturbation (see WM, §3.7).

Now, following the scheme outlined in the previous section, we assume that at some negative time t the time-dependent Hamiltonian is

$$H(t) = H_0 + H'\,e^{\alpha t}. \tag{6.6}$$

In the interaction picture, the wave function $\Psi_\mathrm{I}(t)$ satisfies eq. (5.23), that is,

$$\mathrm{i}\frac{\partial}{\partial t}\Psi_\mathrm{I}(t) = H_\mathrm{I}(t)\Psi_\mathrm{I}(t), \tag{6.7}$$

where

$$\begin{aligned} H_\mathrm{I}(t) &= \exp(\mathrm{i}H_0t)H'\exp(\alpha t)\exp(-\mathrm{i}H_0t) \\ &= \exp(\mathrm{i}H_0t)H'\exp(-\mathrm{i}H_0t)\exp(\alpha t). \end{aligned} \tag{6.8}$$

As $t \to -\infty$, $e^{\alpha t} \to 0$, and hence $H_\mathrm{I} \to 0$, so that according to eq. (5.10) the wave function becomes independent of t, and, for the ground state, is equal to Φ_0.

From eq. (5.24) we have

$$\Psi_\mathrm{I}(t) = U(t,t')\Psi_\mathrm{I}(t'). \tag{6.9}$$

Hence

$$\Psi_\mathrm{I}(t) = U(t,-\infty)\Psi_\mathrm{I}(-\infty) = U(t,-\infty)\Phi_0. \tag{6.10}$$

Also, from eq. (5.19),

$$\Psi_\mathrm{I}(t) = \exp(\mathrm{i}H_0t)\Psi_\mathrm{S}(t), \tag{6.11}$$

so that

$$\Psi_\mathrm{I}(0) = \Psi_\mathrm{S}(0), \tag{6.12}$$

and, for the ground state,

$$\Psi_\mathrm{I}(0) = \Psi_0. \tag{6.13}$$

Therefore, from (6.10),

$$\Psi_0 = U(0,-\infty)\Phi_0. \tag{6.14}$$

Substituting this expression in eq. (6.5), we obtain

$$\Delta E = \frac{\langle\Phi_0|H'U(0,-\infty)|\Phi_0\rangle}{\langle\Phi_0|U(0,-\infty)|\Phi_0\rangle}. \tag{6.15}$$

Although this derivation appears straightforward, it is not, in fact, quite correct. It must be remembered that $U(t,t')$ depends upon α, a fact which we shall recognize by writing $U_\alpha(t,t')$. Clearly the energy correction ΔE cannot depend upon α if the value of this constant is arbitrary.

However, since the adiabatic hypothesis demands that α be very small, an obvious solution is to take the limit as $\alpha \to 0$, so that the interaction is switched on infinitely slowly. We shall see in the following section that agreement with the Rayleigh–Schrödinger perturbation theory is only achieved when this limit is taken. The final expression for ΔE, is, then,

$$\Delta E = \lim_{\alpha \to 0} \frac{\langle \Phi_0 | H' U_\alpha(0, -\infty) | \Phi_0 \rangle}{\langle \Phi_0 | U_\alpha(0, -\infty) | \Phi_0 \rangle}. \tag{6.16}$$

This expression, which, if we use the formula for $U(t, t')$ given in (5.43), is the quotient of two infinite series, can be put in a more convenient form. We note that

$$\frac{\partial}{\partial t} \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle = \left\langle \Phi_0 \left| \frac{\partial}{\partial t} U_\alpha(t, -\infty) \right| \Phi_0 \right\rangle, \tag{6.17}$$

since Φ_0 is independent of t, and also, from eq. (5.32),

$$\mathrm{i} \frac{\partial}{\partial t} U_\alpha(t, -\infty) = H_\mathrm{I}(t) U_\alpha(t, -\infty). \tag{6.18}$$

Therefore

$$\frac{\partial}{\partial t} \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle = -\mathrm{i} \langle \Phi_0 | H_\mathrm{I}(t) U_\alpha(t, -\infty) | \Phi_0 \rangle, \tag{6.19}$$

and hence

$$\frac{\partial}{\partial t} \log \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle = -\frac{\mathrm{i} \langle \Phi_0 | H_\mathrm{I}(t) U_\alpha(t, -\infty) | \Phi_0 \rangle}{\langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle}. \tag{6.20}$$

However, from eq. (6.8), we have

$$H_\mathrm{I}(0) = H', \tag{6.21}$$

so that, putting $t = 0$ in (6.20) and comparing with (6.16), we find

$$\Delta E = \lim_{\alpha \to 0} \mathrm{i} \left[\frac{\partial}{\partial t} \log \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle \right]_{t=0}. \tag{6.22}$$

This is the expression we shall use in the following chapters.

6.3. Correspondence with Rayleigh–Schrödinger Perturbation Theory

In order to increase our confidence in the correctness of eq. (6.22), we shall now show that it leads to the same expressions for the first- and second-order energy corrections as given by the elementary, time-independent, Rayleigh–Schrödinger perturbation theory described in §1.5.

From eq. (5.43) we obtain

$$U_\alpha(t, -\infty) = 1 + \sum_{n=1}^{\infty} U_n, \tag{6.23}$$

where

$$U_n = (-\mathrm{i})^n \int_{-\infty}^{t} H_\mathrm{I}(t_1)\, \mathrm{d}t_1 \int_{-\infty}^{t_1} H_\mathrm{I}(t_2)\, \mathrm{d}t_2 \ldots \int_{-\infty}^{t_{n-1}} H_\mathrm{I}(t_n)\, \mathrm{d}t_n, \tag{6.24}$$

and $H_\mathrm{I}(t)$ is given by (6.8). According to eq. (5.50), this may also be written

$$U_n = \frac{(-\mathrm{i})^n}{n!} \int_{-\infty}^{t} \mathrm{d}t_1 \ldots \int_{-\infty}^{t} \mathrm{d}t_n P[H_\mathrm{I}(t_1) H_\mathrm{I}(t_2) \ldots H_\mathrm{I}(t_n)], \tag{6.25}$$

but (6.24) is more suitable for our present purposes.

Then

$$\begin{aligned} \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle &= \left\langle \Phi_0 \left| 1 + \sum_{n=1}^{\infty} U_n \right| \Phi_0 \right\rangle \\ &= 1 + \sum_{n=1}^{\infty} A_n, \end{aligned} \tag{6.26}$$

say, where

$$A_n = \langle \Phi_0 | U_n | \Phi_0 \rangle. \tag{6.27}$$

Thus[†]

$$\begin{aligned} \log \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle &= \log (1 + \Sigma A_n) \\ &= \Sigma A_n - \tfrac{1}{2}(\Sigma A_n)^2 + \tfrac{1}{3}(\Sigma A_n)^3 - \ldots \\ &= A_1 + A_2 + A_3 + \ldots - \tfrac{1}{2}A_1^2 - A_1 A_2 - \ldots \\ &\quad + \tfrac{1}{3}A_1^3 + \ldots, \end{aligned} \tag{6.28}$$

where the only terms explicitly shown in the last line are those involving the first, second or third powers of H' only.

As in §1.5, we shall write the energy correction, eq. (6.22), in the form

$$\Delta E = \Delta E^{(1)} + \Delta E^{(2)} + \Delta E^{(3)} + \ldots, \tag{6.29}$$

[†] Strictly speaking, this expansion is only permissible if $-1 < \Sigma A_n \leqslant 1$ – that is to say, for a *small* perturbation. However, since our intention is merely to demonstrate the formal correspondence between the present theory and the Rayleigh–Schrödinger theory, we shall not concern ourselves with questions of convergence. Where the given expansion fails we should expect the elementary Rayleigh–Schrödinger theory also to fail.

where $\Delta E^{(n)}$ involves the n-th power of H' only and is called the *n-th-order energy correction.* Eqs. (6.22) and (6.28) give the first-, second- and third-order† energy corrections as

$$\Delta E^{(1)} = \lim_{\alpha\to 0} \mathrm{i}\left[\frac{\partial A_1}{\partial t}\right]_{t=0}, \tag{6.30}$$

$$\Delta E^{(2)} = \lim_{\alpha\to 0} \mathrm{i}\left[\frac{\partial}{\partial t}(A_2-\tfrac{1}{2}A_1^2)\right]_{t=0}, \tag{6.31}$$

$$\Delta E^{(3)} = \lim_{\alpha\to 0} \mathrm{i}\left[\frac{\partial}{\partial t}(A_3-A_1A_2+\tfrac{1}{3}A_1^3)\right]_{t=0}. \tag{6.32}$$

Let us consider the first-order correction. Now

$$\begin{aligned} A_1 &= \langle\Phi_0|U_1|\Phi_0\rangle \\ &= \left\langle\Phi_0\middle|(-\mathrm{i})\int_{-\infty}^{t} H_\mathrm{I}(t_1)\,\mathrm{d}t_1\middle|\Phi_0\right\rangle, \end{aligned} \tag{6.33}$$

so that

$$\frac{\partial A_1}{\partial t} = -\mathrm{i}\langle\Phi_0|H_\mathrm{I}(t)|\Phi_0\rangle \tag{6.34}$$

and

$$\begin{aligned} \left[\frac{\partial A_1}{\partial t}\right]_{t=0} &= -\mathrm{i}\langle\Phi_0|H_\mathrm{I}(0)|\Phi_0\rangle \\ &= -\mathrm{i}\langle\Phi_0|H'|\Phi_0\rangle. \end{aligned} \tag{6.35}$$

Hence, from eq. (6.30),

$$\Delta E^{(1)} = \langle\Phi_0|H'|\Phi_0\rangle, \tag{6.36}$$

which agrees with eq. (1.78) for the first-order energy correction given by the Rayleigh–Schrödinger theory. It may be noted that the limit as α tends to zero is redundant here, since $H_\mathrm{I}(0)$ is independent of α, but we shall see that this does not apply to the higher-order terms.

Now, let us consider the second-order energy correction, which is rather more complicated. First,

$$\begin{aligned} A_2 &= \langle\Phi_0|U_2|\Phi_0\rangle \\ &= (-\mathrm{i})^2\left\langle\Phi_0\middle|\int_{-\infty}^{t} H_\mathrm{I}(t_1)\,\mathrm{d}t_1\int_{-\infty}^{t_1} H_\mathrm{I}(t_2)\,\mathrm{d}t_2\middle|\Phi_0\right\rangle, \end{aligned} \tag{6.37}$$

---

† The order of any term in the expansion (6.28) is simply obtained by summing the subscripts: thus the order of the term $A_3A_5^3A_6^2$ is $3+(3\times 5)+(2\times 6)=30$.

so that

$$\frac{\partial A_2}{\partial t} = -\left\langle\Phi_0\middle|H_\mathrm{I}(t)\int_{-\infty}^{t} H_\mathrm{I}(t_1)\,\mathrm{d}t_1\middle|\Phi_0\right\rangle \tag{6.38}$$

and

$$\begin{aligned} \left[\frac{\partial A_2}{\partial t}\right]_{t=0} &= -\left\langle\Phi_0\middle|H'\int_{-\infty}^{0} H_\mathrm{I}(t)\,\mathrm{d}t\middle|\Phi_0\right\rangle \\ &= -\int_{-\infty}^{0} \langle\Phi_0|H'H_\mathrm{I}(t)|\Phi_0\rangle\,\mathrm{d}t \\ &= -\int_{-\infty}^{0} \langle\Phi_0|H'\exp(\mathrm{i}H_0t)H'\exp(-\mathrm{i}H_0t)|\Phi_0\rangle\exp(\alpha t)\,\mathrm{d}t \\ &= -\int_{-\infty}^{0} \langle\Phi_0|H'\exp(\mathrm{i}H_0t)H'|\Phi_0\rangle\exp(-\mathrm{i}E_0t)\exp(\alpha t)\,\mathrm{d}t. \end{aligned} \tag{6.39}$$

In general, H_0 and H' do not commute, so that we cannot simply substitute $\exp(\mathrm{i}E_0t)$ for $\exp(\mathrm{i}H_0t)$ in this integral, but we proceed by expanding $H'\Phi_0$ in terms of the eigenfunctions of H_0. Let us suppose that Φ_n is an eigenfunction of H_0 corresponding to the eigenvalue E_n, that is,

$$H_0\Phi_n = E_n\Phi_n, \tag{6.40}$$

and that these eigenfunctions are orthogonal and normalized. Then, according to eq. (1.84),

$$H'\Phi_0 = \sum_{n=0}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle\Phi_n. \tag{6.41}$$

Using this, and assuming that H' is Hermitian, we find

$$\begin{aligned} &\langle\Phi_0|H'\exp(\mathrm{i}H_0t)H'|\Phi_0\rangle \\ &\quad= \int \Phi_0^*H'\exp(\mathrm{i}H_0t)H'\Phi_0\,\mathrm{d}\tau' \\ &\quad= \int (H'\Phi_0)^*\exp(\mathrm{i}H_0t)(H'\Phi_0)\,\mathrm{d}\tau' \\ &\quad= \int\left[\sum_{n=0}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle\Phi_n\right]^*\exp(\mathrm{i}H_0t)\left[\sum_{m=0}^{\infty}\langle\Phi_m|H'|\Phi_0\rangle\Phi_m\right]\mathrm{d}\tau' \\ &\quad= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\langle\Phi_n|H'|\Phi_0\rangle^*\langle\Phi_m|H'|\Phi_0\rangle\int\Phi_n^*\exp(\mathrm{i}H_0t)\Phi_m\,\mathrm{d}\tau' \\ &\quad= \sum_{n=0}^{\infty}|\langle\Phi_n|H'|\Phi_0\rangle|^2\exp(\mathrm{i}E_nt). \end{aligned} \tag{6.42}$$

Therefore, from (6.39),

$$\left[\frac{\partial A_2}{\partial t}\right]_{t=0} = -\sum_{n=0}^{\infty} |\langle\Phi_n|H'|\Phi_0\rangle|^2 \int_{-\infty}^{0} \exp\left[-\mathrm{i}(E_0-E_n+\mathrm{i}\alpha)t\right] \mathrm{d}t$$

$$= \sum_{n=0}^{\infty} \frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{\mathrm{i}(E_0-E_n+\mathrm{i}\alpha)}. \tag{6.43}$$

Also,

$$\frac{\partial}{\partial t}A_1^2 = 2A_1\frac{\partial A_1}{\partial t}$$

$$= -2\left\langle\Phi_0\middle|\int_{-\infty}^{t} H_I(t_1)\,\mathrm{d}t_1\middle|\Phi_0\right\rangle\langle\Phi_0|H_I(t)|\Phi_0\rangle, \tag{6.44}$$

and

$$\left[\frac{\partial}{\partial t}A_1^2\right]_{t=0} = -2\left\langle\Phi_0\middle|\int_{-\infty}^{0} H_I(t)\,\mathrm{d}t\middle|\Phi_0\right\rangle\langle\Phi_0|H'|\Phi_0\rangle. \tag{6.45}$$

Now

$$\left\langle\Phi_0\middle|\int_{-\infty}^{0} H_I(t)\,\mathrm{d}t\middle|\Phi_0\right\rangle$$

$$= \left\langle\Phi_0\middle|\int_{-\infty}^{0} \exp(\mathrm{i}H_0t)H'\exp(-\mathrm{i}H_0t)\exp(\alpha t)\,\mathrm{d}t\middle|\Phi_0\right\rangle$$

$$= \int_{-\infty}^{0} \exp(\alpha t)\,\mathrm{d}t \int \Phi_0^* \exp(\mathrm{i}H_0t)H'\exp(-\mathrm{i}H_0t)\Phi_0\,\mathrm{d}\tau'$$

$$= \int_{-\infty}^{0} \exp(\alpha t)\,\mathrm{d}t \int [\exp(-\mathrm{i}H_0t)\Phi_0]^* H'\exp(-\mathrm{i}E_0t)\Phi_0\,\mathrm{d}\tau'$$

$$= \int_{-\infty}^{0} \exp(\alpha t)\,\mathrm{d}t \int \Phi_0^* H'\Phi_0\,\mathrm{d}\tau'$$

$$= \frac{1}{\alpha}\langle\Phi_0|H'|\Phi_0\rangle. \tag{6.46}$$

Therefore

$$\left[\frac{\partial}{\partial t}(\tfrac{1}{2}A_1^2)\right]_{t=0} = -\frac{1}{\alpha}\langle\Phi_0|H'|\Phi_0\rangle^2. \tag{6.47}$$

Substituting (6.43) and (6.47) in (6.31), we obtain

$$\Delta E^{(2)} = \lim_{\alpha\to 0}\left[\sum_{n=0}^{\infty}\frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{E_0-E_n+\mathrm{i}\alpha} + \frac{\mathrm{i}}{\alpha}\langle\Phi_0|H'|\Phi_0\rangle^2\right]$$

$$= \lim_{\alpha\to 0}\sum_{n=1}^{\infty}\frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{E_0-E_n+\mathrm{i}\alpha}$$

$$= \sum_{n=1}^{\infty}\frac{|\langle\Phi_n|H'|\Phi_0\rangle|^2}{E_0-E_n}, \tag{6.48}$$

which is just the formula for the second-order energy correction given by the Rayleigh–Schrödinger perturbation theory, eq. (1.87). Notice that (6.43) and (6.47) separately tend to ∞ as $\alpha \to 0$, but that the awkward terms cancel in (6.48). Unlike the first-order correction, here it is necessary to take the limit as $\alpha \to 0$, owing to the presence of the term $\mathrm{i}\alpha$ in the denominator.

It can be shown that the third-order correction, eq. (6.32), also reduces to the Rayleigh–Schrödinger expression, but this is even more complicated than the second-order correction, and so we leave it as an exercise for the enthusiastic reader. The same applies, of course, to the energy correction of any order.

As stated earlier, the perturbation theory described in this chapter and the foregoing one can, in principle, be applied to any system, not only the many-electron system which is our main concern. That this must be so is obvious from its equivalence to the Rayleigh–Schrödinger theory. One might well ask whether it has any advantages over the latter theory. If, in fact, the first- and second-order terms provided a sufficiently accurate value of the energy correction, as is often the case in simple perturbation problems, there would be no advantage in using the present theory at all. Such advantages as it has only appear when higher-order corrections have to be considered, as in the case of an electron gas when the interaction term is treated as a perturbation. Then we shall see in the following chapters that the time-dependent method, together with the use of Feynman graphs, offers a convenient way of accounting for all the terms of the perturbation series.

CHAPTER 7

FEYNMAN GRAPHS

7.1. Creation and Destruction Operators in the Interaction Picture

The creation and destruction operators defined in ch. 2 become time-dependent in the interaction picture, and, according to eq. (5.20), have the form

$$c_j(t) = \exp(\mathrm{i}H_0t)c_j\exp(-\mathrm{i}H_0t), \tag{7.1}$$

$$c_j^\dagger(t) = \exp(\mathrm{i}H_0t)c_j^\dagger\exp(-\mathrm{i}H_0t), \tag{7.2}$$

the Schrödinger operators c_j, $c_j^\dagger$ being simply $c_j(0)$, $c_j^\dagger(0)$, respectively.

Now, suppose that Φ^N is a determinantal wave function of the non-interacting system, that is to say, a (time-independent) eigenfunction of H_0, corresponding to the eigenvalue E^N, so that

$$H_0\Phi^N = E^N\Phi^N. \tag{7.3}$$

As stated in §6.2, we assume that the creation and destruction operators refer to the one-electron functions whose orbital factors are eigenfunctions of the separated parts of H_0 (eq. (6.2)). Then

$$\begin{aligned} c_j(t)\Phi^N &= \exp(\mathrm{i}H_0t)c_j\exp(-\mathrm{i}H_0t)\Phi^N \\ &= \exp(\mathrm{i}H_0t)c_j\Phi^N\exp(-\mathrm{i}E^Nt) \\ &= \exp(\mathrm{i}H_0t)\Phi^{N-1}_{(j)}\exp(-\mathrm{i}E^Nt), \end{aligned} \tag{7.4}$$

where we have written $\Phi^{N-1}_{(j)}$ for the $(N-1)$-th-order determinant obtained by removing the function ϕ_j from Φ^N.

Now we recall that†, in the occupation number representation,

$$H_0 = \sum_{i,j}\langle i|f|j\rangle c_i^\dagger c_j, \tag{7.5}$$

and, since i and j take on all integral values here, H_0 may operate upon a determinantal function of any order. In the present case, where ϕ_i and ϕ_j are eigenfunctions of f, such that

$$f\phi_i = \epsilon_i\phi_i \tag{7.6}$$

(this is eq. (6.2) in general notation and including the spin factor in ϕ_i), we have

$$\langle i|f|j\rangle = \epsilon_i\delta_{ij}, \tag{7.7}$$

and (7.5) reduces to

$$H_0 = \sum_i \epsilon_i c_i^\dagger c_i. \tag{7.8}$$

Thus

$$\begin{aligned} H_0\Phi^N &= \sum_i \epsilon_i c_i^\dagger c_i\Phi^N \\ &= \sum_i \epsilon_i n_i\Phi^N = E^N\Phi^N, \end{aligned} \tag{7.9}$$

where n_i is the occupation number of the function ϕ_i in Φ^N.

It follows that

$$H_0\Phi^{N-1}_{(j)} = E^{N-1}_{(j)}\Phi^{N-1}_{(j)}, \tag{7.10}$$

where

$$E^{N-1}_{(j)} = \sum_{i\neq j}\epsilon_i n_i = E^N - \epsilon_j. \tag{7.11}$$

Eq. (7.4) therefore becomes

$$\begin{aligned} c_j(t)\Phi^N &= \Phi^{N-1}_{(j)}\exp[\mathrm{i}(E^{N-1}_{(j)} - E^N)t] \\ &= c_j\Phi^N\exp(-\mathrm{i}\epsilon_j t). \end{aligned} \tag{7.12}$$

We have, of course, assumed that Φ^N contains the function ϕ_j, so that the occupation number of ϕ_j is unity. If not, both sides of (7.4) vanish.

We deduce, therefore, that

$$c_j(t) = c_j\exp(-\mathrm{i}\epsilon_j t), \tag{7.13}$$

and, similarly,

$$c_j^\dagger(t) = c_j^\dagger\exp(\mathrm{i}\epsilon_j t). \tag{7.14}$$

---

† H_0 is the same in the Schrödinger and interaction pictures, since

$$H_0 = \exp(\mathrm{i}H_0t)H_0\exp(-\mathrm{i}H_0t).$$

Finally, we note that the product of any number of Schrödinger creation and destruction operators is transformed to the interaction picture by substituting $c_j(t)$ and $c_j^\dagger(t)$ for c_j and $c_j^\dagger$, since, for example,

$$\begin{aligned}\exp(\mathrm{i}H_0t)c_i^\dagger c_j\exp(-\mathrm{i}H_0t) &= \exp(\mathrm{i}H_0t)c_i^\dagger\exp(-\mathrm{i}H_0t)\\ &\quad\times\exp(\mathrm{i}H_0t)c_j\exp(-\mathrm{i}H_0t)\\ &= c_i^\dagger(t)c_j(t).\end{aligned}\tag{7.15}$$

7.2. The interaction term and the energy of the ground state

In the Schrödinger picture the interaction term, eq. (2.68), is

$$H' = \tfrac{1}{2}\sum_{ijkl}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k.\tag{7.16}$$

We wish to use the method described in §6.2, and so must multiply H' by $\exp(\alpha t)$ and transform to the interaction picture. We then obtain for the transformed interaction $H_\mathrm{I}(t)$, eq. (6.11),

$$\begin{aligned}H_\mathrm{I}(t) &= \tfrac{1}{2}\sum_{ijkl}\langle ij|v|kl\rangle c_i^\dagger(t)c_j^\dagger(t)c_l(t)c_k(t)\exp(\alpha t)\\ &= \tfrac{1}{2}\sum_{ijkl}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k\exp[\mathrm{i}(\epsilon_i+\epsilon_j-\epsilon_l-\epsilon_k)t]\exp(\alpha t).\end{aligned}\tag{7.17}$$

The correction to the energy of the ground state due to the interaction is given by eq. (6.22), and to evaluate it we must first calculate $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$. In the notation of eqs. (6.24), (6.26) and (6.27) we have, for the first-order term,

$$\begin{aligned}A_1 &= \langle\Phi_0|U_1|\Phi_0\rangle\\ &= \left\langle\Phi_0\middle|(-\mathrm{i})\int_{-\infty}^{t}H_\mathrm{I}(t_1)\,\mathrm{d}t_1\middle|\Phi_0\right\rangle\\ &= -\tfrac{1}{2}\mathrm{i}\sum_{ijkl}\langle ij|v|kl\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k|\Phi_0\rangle\\ &\quad\times\int_{-\infty}^{t}\exp[\mathrm{i}(\epsilon_i+\epsilon_j-\epsilon_l-\epsilon_k)t_1]\exp(\alpha t_1)\,\mathrm{d}t_1\\ &= -\tfrac{1}{2}\mathrm{i}\sum_{ijkl}\langle ij|v|kl\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k|\Phi_0\rangle\\ &\quad\times\frac{\exp[(\mathrm{i}\Delta_1+\alpha)t]}{\mathrm{i}\Delta_1+\alpha},\end{aligned}\tag{7.18}$$

where

$$\Delta_1 = \epsilon_i+\epsilon_j-\epsilon_l-\epsilon_k.\tag{7.19}$$

For the second-order term we obtain

$$\begin{aligned}A_2 &= \langle\Phi_0|U_2|\Phi_0\rangle\\ &= \left\langle\Phi_0\middle|(-\mathrm{i})^2\int_{-\infty}^{t}H_\mathrm{I}(t_1)\,\mathrm{d}t_1\int_{-\infty}^{t_1}H_\mathrm{I}(t_2)\,\mathrm{d}t_2\middle|\Phi_0\right\rangle\\ &= (-\tfrac{1}{2}\mathrm{i})^2\sum_{ijkl}\sum_{mnpq}\langle ij|v|kl\rangle\langle mn|v|pq\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_m^\dagger c_n^\dagger c_q c_p|\Phi_0\rangle\\ &\quad\times\int_{-\infty}^{t}\exp[\mathrm{i}(\epsilon_i+\epsilon_j-\epsilon_l-\epsilon_k)t_1]\exp(\alpha t_1)\,\mathrm{d}t_1\\ &\quad\times\int_{-\infty}^{t_1}\exp[\mathrm{i}(\epsilon_m+\epsilon_n-\epsilon_q-\epsilon_p)t_2]\exp(\alpha t_2)\,\mathrm{d}t_2\\ &= (-\tfrac{1}{2}\mathrm{i})^2\sum_{ijkl}\sum_{mnpq}\langle ij|v|kl\rangle\langle mn|v|pq\rangle\\ &\quad\times\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_m^\dagger c_n^\dagger c_q c_p|\Phi_0\rangle\frac{\exp\{[\mathrm{i}(\Delta_1+\Delta_2)+2\alpha]t\}}{[\mathrm{i}(\Delta_1+\Delta_2)+2\alpha](\mathrm{i}\Delta_2+\alpha)},\end{aligned}\tag{7.20}$$

where

$$\Delta_2 = \epsilon_m+\epsilon_n-\epsilon_q-\epsilon_p.\tag{7.21}$$

The n-th-order term may easily be deduced from eqs. (7.18) and (7.20). We find

$$\begin{aligned}A_n &= \langle\Phi_0|U_n|\Phi_0\rangle\\ &= (-\tfrac{1}{2}\mathrm{i})^n\sum_{ijkl}\dots\sum_{rsuw}\langle ij|v|kl\rangle\dots\langle rs|v|uw\rangle\\ &\quad\times\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k\dots c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle\\ &\quad\times\int_{-\infty}^{t}\exp[(\mathrm{i}\Delta_1+\alpha)t_1]\,\mathrm{d}t_1\int_{-\infty}^{t_1}\exp[(\mathrm{i}\Delta_2+\alpha)t_2]\,\mathrm{d}t_2\\ &\quad\times\dots\int_{-\infty}^{t_{n-1}}\exp[(\mathrm{i}\Delta_n+\alpha)t_n]\,\mathrm{d}t_n\\ &= (-\tfrac{1}{2}\mathrm{i})^n\sum_{ijkl}\dots\sum_{rsuw}\langle ij|v|kl\rangle\dots\langle rs|v|uw\rangle\\ &\times\frac{\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k\dots c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle\exp\{[\mathrm{i}(\Delta_1+\Delta_2+\dots+\Delta_n)+n\alpha]t\}}{[\mathrm{i}(\Delta_1+\Delta_2+\dots+\Delta_n)+n\alpha][\mathrm{i}(\Delta_2+\dots+\Delta_n)+(n-1)\alpha]\dots(\mathrm{i}\Delta_n+\alpha)},\end{aligned}\tag{7.22}$$

where

$$\Delta_n = \epsilon_r + \epsilon_s - \epsilon_u - \epsilon_w, \tag{7.23}$$

etc.

This is a rather complicated expression. A certain simplification is introduced by the fact that the matrix element

$$\langle \Phi_0 | c_i^\dagger c_j^\dagger c_l c_k \ldots c_r^\dagger c_s^\dagger c_w c_u | \Phi_0 \rangle$$

vanishes if the creation and destruction operators do not balance in pairs: that is to say, if each state destroyed is not created again and *vice versa*. Great complexity still remains, however, and to help to bring some order out of chaos a method of describing the terms of the series in (7.22) by graphs or diagrams has been developed. This will be explained below, but first it will be convenient to re-define the creation and destruction operators in terms of excited particles and holes in the unperturbed ground state Φ_0 – in other words, to use Φ_0 as the 'vacuum state' in describing excited unperturbed states.

7.3. Particle and Hole Operators

In the case of free electrons, as we have already seen in ch. 3, the ground-state wave function Φ_0 of the unperturbed system is a single determinant of functions $\phi_{\boldsymbol{k}}(\boldsymbol{x}_i)$ whose $\boldsymbol{k}$-vectors lie within the Fermi sphere in $\boldsymbol{k}$-space; that is to say, each (one-electron) state with $k \leqslant k_{\mathrm{F}}$ is occupied and no state with $k > k_{\mathrm{F}}$ is occupied. States lying within the Fermi sphere are called *unexcited* states and those outside the Fermi sphere are called *excited* states. Any other eigenfunction Φ of H_0 can be described by stating which excited states are occupied and which unexcited states are unoccupied. An unoccupied, unexcited state is called a *hole*, so that the destruction of an electron in an unexcited state may be regarded as the creation of a hole, and the creation of an electron in an unoccupied, unexcited state is the destruction of a hole.

Although the free-electron case is the only one for which detailed calculations have been made, we wish to point out that what follows is of more general application than this. For example, in a metal the electrons move in the periodic field due to the lattice of ion-cores, so that the potential function $V(\boldsymbol{r}_i)$ appearing in H_0, eq. (1.2), has the period of the lattice. We then construct determinantal eigenfunctions of H_0 from the orthogonal set of one-electron functions whose orbital factors are eigenfunctions of eq. (6.2), and, since $V(\boldsymbol{r})$ is periodic, these are *Bloch functions*, having the form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})u_{\boldsymbol{k}}(\boldsymbol{r}), \tag{7.24}$$

where $u_{\boldsymbol{k}}(\boldsymbol{r})$ has the period of the lattice. The specification of the wave-vector $\boldsymbol{k}$ depends upon the crystal structure. In this case also the one-electron functions comprising the ground state Φ_0 lie within a surface of constant energy in $\boldsymbol{k}$-space, but this surface is not spherical. We call this the *Fermi surface*, and denote the one-electron energy (eigenvalue of eq. (6.2)) at the Fermi surface by ϵ_{F}. An excited state is now one with energy $\epsilon(\boldsymbol{k}) > \epsilon_{\mathrm{F}}$, and an unexcited state one with $\epsilon(\boldsymbol{k}) \leqslant \epsilon_{\mathrm{F}}$. This also applies to the free-electron case, but there it is possible to say $k > k_{\mathrm{F}}$ and $k \leqslant k_{\mathrm{F}}$, which is not possible in general, since k is not constant over a non-spherical Fermi surface. Any determinantal eigenfunction of H_0 can again be described in terms of holes[†], which are unoccupied states lying within the Fermi surface, and particles in excited states lying outside the Fermi surface.

We are now in a position to define the creation and destruction operators for holes and particles (excited states) in terms of our previous operators. Thus:

for $\epsilon_i > \epsilon_{\mathrm{F}}$,

$$\begin{aligned} c_i^\dagger &= a_i^\dagger \quad &&\text{(creates particle),} \\ c_i &= a_i &&\text{(destroys particle);} \end{aligned} \tag{7.25}$$

for $\epsilon_j \leqslant \epsilon_{\mathrm{F}}$,

$$\begin{aligned} c_j^\dagger &= b_j \quad &&\text{(destroys hole),} \\ c_j &= b_j^\dagger &&\text{(creates hole).} \end{aligned} \tag{7.26}$$

The operators $a_i^\dagger$, a_i refer to excited states or particles outside the Fermi surface, while $b_j^\dagger$, b_j refer to holes inside the Fermi surface. We have used different subscripts i and j here merely to emphasize that particle and hole states are distinct, so that we must always have

$$\epsilon_i \neq \epsilon_j, \tag{7.27}$$

or, more specifically,

$$\boldsymbol{k}_i \neq \boldsymbol{k}_j. \tag{7.28}$$

The commutation relations for the particle and hole operators can be found immediately from those for the $c_i^\dagger$, c_j, given in eqs. (2.53), (2.54) and (2.55). Thus

$$\{a_i^\dagger, a_k\} = \delta_{ik}, \tag{7.29}$$

[†] It must be remembered that these holes are not what are usually described as holes in metal theory (see WM, p. 242), which are unoccupied states within a *Brillouin zone*.

$$\{b_j^\dagger, b_l\} = \delta_{jl}, \tag{7.30}$$

but, owing to (7.28), all other anticommutators vanish.

Since the ground state Φ_0 contains neither holes nor excited states, we have

$$a_i\Phi_0 = 0, \tag{7.31}$$

$$b_j\Phi_0 = 0. \tag{7.32}$$

In terms of the particle and hole operators, H_0, eq. (7.8), becomes

$$\begin{aligned} H_0 &= \sum_i \epsilon_i c_i^\dagger c_i \\ &= \sum_{\epsilon_j \le \epsilon_F} \epsilon_j c_j^\dagger c_j + \sum_{\epsilon_i > \epsilon_F} \epsilon_i c_i^\dagger c_i \\ &= \sum_{\epsilon_j \le \epsilon_F} \epsilon_j b_j b_j^\dagger + \sum_{\epsilon_i > \epsilon_F} \epsilon_i a_i^\dagger a_i \\ &= \sum_{\epsilon_j \le \epsilon_F} \epsilon_j - \sum_{\epsilon_j \le \epsilon_F} \epsilon_j b_j^\dagger b_j + \sum_{\epsilon_i > \epsilon_F} \epsilon_i a_i^\dagger a_i, \end{aligned} \tag{7.33}$$

using (7.30). The first term on the right-hand side is just the energy of the state Φ_0.

In the interaction picture, from eqs. (7.13) and (7.14) and the definitions (7.25) and (7.26), we have

$$\left.\begin{aligned} a_i(t) &= a_i \exp(-i\epsilon_i t), \\ a_i^\dagger(t) &= a_i^\dagger \exp(i\epsilon_i t), \\ b_j(t) &= b_j \exp(i\epsilon_j t), \\ b_j^\dagger(t) &= b_j^\dagger \exp(-i\epsilon_j t). \end{aligned}\right\} \tag{7.34}$$

7.4. Feynman Graphs

We have already stated our intention to describe the terms of $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ by means of graphs. These are similar to those introduced by Feynman† in his theory of positrons and are generally called *Feynman graphs* or *Feynman diagrams*.

We recall that U_n, eq. (6.25), contains a chronologically-ordered product of interaction terms $P[H_I(t_1)H_I(t_2)\dots H_I(t_n)]$, with $H_I(t)$ given by eq. (7.17). Let us first see how to represent graphically a single term

$$\tfrac{1}{2}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k \exp[i(\epsilon_i+\epsilon_j-\epsilon_l-\epsilon_k)t] \exp(\alpha t) \tag{7.35}$$

of $H_I(t)$.

---

† Feynman, R. P., 1949, Phys. Rev. **76**, 749. The graphs described here were first used in many-body theory by Goldstone (Goldstone, J., 1957, Proc. Roy. Soc. **A239**, 267) and are sometimes referred to as *Goldstone graphs*. Several authors have developed their own variations of these graphs, which adds to the confusion of the subject.

7.4.1. *First-order graphs*

We draw a horizontal broken line, called an *interaction line*, and from each end, called a *vertex*†, we draw two solid lines bearing arrow heads, one line entering and one leaving each vertex. The left-hand pair of lines represents $c_i^\dagger$, c_k, in (7.35), and the right-hand pair represents $c_j^\dagger$, c_l. The general directions of the lines, that is, whether they go up or down the page (the actual angles they make with the interaction line are immaterial), and the sense of the arrows, depends upon whether the operators are particle or hole, creation or destruction operators, according to the following rules:

(1) An *upward* line leaving a vertex refers to the creation of a *particle* in an excited state, e.g. the line marked $a_i^\dagger$ in fig. 7.1.

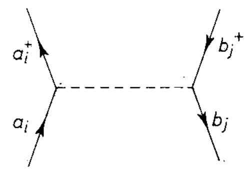


Fig. 7.1. A first-order graph, representing a single term of $H_I(t)$, showing the interaction line (broken) and the two vertices with lines representing hole and particle, creation and destruction operators.

(2) An *upward* line entering a vertex refers to the destruction of a *particle* in an excited state, e.g. the line marked a_i in fig. 7.1.

(3) A *downward* line entering a vertex refers to the creation of a *hole*, e.g. the line marked $b_j^\dagger$ in fig. 7.1.

---

† It should be noted that there are variations in terminology, as indeed there are variations in the type of graphs used, in this part of the subject. Sometimes the interaction line itself is called a vertex. This name is certainly appropriate to the frequently-used graphical scheme in which the interaction line is shrunk to a point. However, the terminology we have chosen seems more suited to Goldstone graphs.

(4) A *downward* line leaving a vertex refers to the destruction of a *hole*, e.g. the line marked b_j in fig. 7.1.

We note that *upward* lines refer to *particles*, and *downward* lines to *holes*. Also, lines *leaving* a vertex refer to the creation of particles or the destruction of holes, that is, the *creation* of excited or unexcited states, respectively. Lines entering a vertex refer to the destruction of particles or the creation of holes, that is, the *destruction* of excited or unexcited states, respectively.

The lines are normally labelled with the subscripts† of the states to which they refer, so that a given graph fixes completely the matrix element, the operators and their order, and consequently also the time factor of a term such as (7.35). Such graphs are called *first-order graphs*. We shall consider the more useful higher-order graphs, representing terms of products $H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\ldots$, later. It might be noted, in passing, that first-order graphs can equally well be used to represent the terms of the time-independent operator H', eq. (2.68).

A convenient way of interpreting a given graph is as follows. Write down on the left-hand side of the matrix element the labels of the lines *leaving* the vertices on the left and right, in order, and on the right-hand side of the matrix element the labels of the lines *entering* the vertices on the left and right, in order. The order of the operators is fixed by the matrix element, for example $\langle ij|v|kl\rangle$ implies $c_i^\dagger c_j^\dagger c_l c_k$. It must be remembered that $c_i^\dagger$ may imply $a_i^\dagger$ or b_i, depending upon whether ϕ_i is an excited or unexcited state, and similarly c_l may imply a_l or $b_l^\dagger$.

We shall now consider a number of examples, in which we will not specify the time factor, since it also is fixed by the matrix element, for example, $\langle ij|v|kl\rangle$ implies $\exp[\mathrm{i}(\epsilon_i+\epsilon_j-\epsilon_k-\epsilon_l)t]$. In addition, all terms have a factor $\frac{1}{2}\exp(\alpha t)$.

(1) $\langle ij|v|kl\rangle a_i^\dagger a_j^\dagger b_l^\dagger b_k^\dagger$

is shown in fig. 7.2. This represents the excitation of electrons from states ϕ_k, ϕ_l within the Fermi surface, leaving holes, to states ϕ_i, ϕ_j, respectively, outside the Fermi surface. As an aid to understanding, a more obvious picture of this process has been drawn alongside the Feynman graph, showing a circle which represents the Fermi surface and two directed lines† representing electrons going from ϕ_k to ϕ_i and from ϕ_l to ϕ_j.

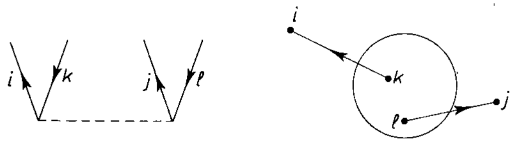


Fig. 7.2.

(2) $\langle ij|v|kl\rangle b_i b_j a_l a_k$

is shown in fig. 7.3. It represents two electrons falling from excited states to unexcited states, destroying holes, as pictured to the right of the graph‡.

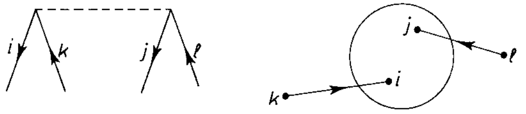


Fig. 7.3.

(3) $\langle ij|v|kl\rangle a_i^\dagger a_j^\dagger b_l^\dagger a_k$

is shown in fig. 7.4. It represents an electron passing from excited state ϕ_k to excited state ϕ_i and another passing from unexcited state ϕ_l to excited state ϕ_j, creating a hole.

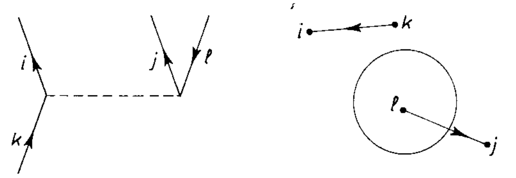


Fig. 7.4.

(4) $\langle ij|v|kl\rangle a_i^\dagger b_j b_l^\dagger a_k$

is shown in fig. 7.5. It represents an electron passing from excited state ϕ_k to excited state ϕ_i and another passing from unexcited state ϕ_l, creating a hole, to unexcited state ϕ_j, destroying a hole.

---

† We shall in future drop the convention, introduced in eqs. (7.25) and (7.26), of reserving the subscript i for a particle state and the subscript j for a hole state. It will be sufficient to remember that the subscript of a particle operator can never be the same as that of a hole operator.

† There is no apparent reason for saying 'from ϕ_k to ϕ_i and from ϕ_l to ϕ_j' rather than 'from ϕ_k to ϕ_j and from ϕ_l to ϕ_i', which would give the same final state. The given interpretation is a natural one, since, in the integrand of $\langle ij|v|kl\rangle$, eq. (1.38), ϕ_i^* and ϕ_k are functions of $\boldsymbol{x}_1$ and ϕ_j^* and ϕ_l are functions of $\boldsymbol{x}_2$. However, the choice is really immaterial, since the expansion of $H_{\mathrm{I}}(t)$ also contains this term with i and j interchanged.

‡ Each figure is to be regarded as a separate illustration of the *type* of term which can arise. Figs. 7.2 and 7.3, if taken literally, cannot apply to the same system, since a given state ϕ_i, for example, cannot be both an excited state and an unexcited state.

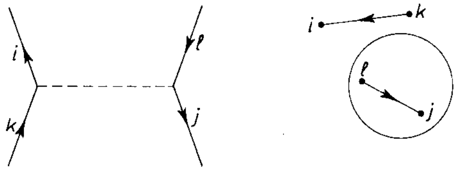
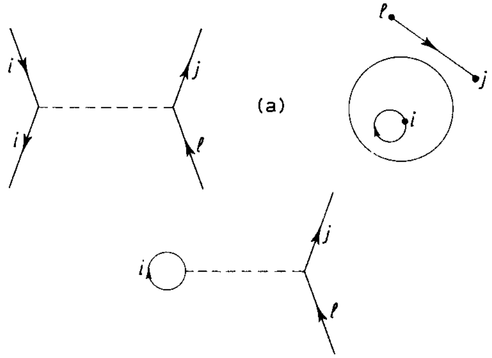


Fig. 7.5.

(5) A pair of operators, one creation and one destruction, may refer to the same state. In this case, the lines representing the two operators are joined to form a loop. For example, the term

$$\langle ij|v|il\rangle b_i a_j^\dagger a_l b_i^\dagger,$$

shown in fig. 7.6(a), is usually drawn as in fig. 7.6(b). It represents an interaction in which an unexcited particle has taken part without changing its state ϕ_i, as pictured to the right of the graph in fig. 7.6(a).



(a)

(b)

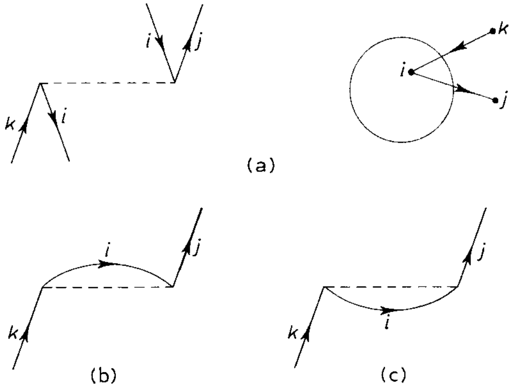
Fig. 7.6.

(6) As another example, consider the term

$$\langle ij|v|ki\rangle b_i a_j^\dagger b_i^\dagger a_k,$$

shown in fig. 7.7(a). This is usually drawn as in figs. 7.7(b) or 7.7(c), the two lines referring to the creation and destruction of the same hole, this time at opposite vertices, again being joined to form a loop. It is customary to draw the loop on one side or the other of the interaction line, as shown, even though fig. 7.7(a) would suggest a sinusoidal curve crossing the interaction line.

It should be noted that interchanging the vertices of a first-order graph



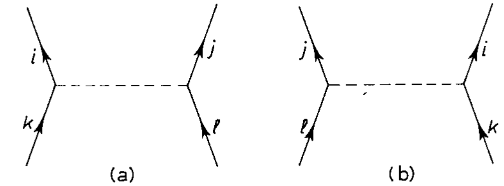
(a)

(b)

(c)

Fig. 7.7.

makes no difference to the term it represents. For example, consider the two graphs shown in fig. 7.8.



(a)

(b)

Fig. 7.8.

Fig. 7.8(a) represents the term

$$\langle ij|v|kl\rangle a_i^\dagger a_j^\dagger a_l a_k, \tag{7.36}$$

while fig. 7.8(b), obtained by interchanging the two vertices in fig. 7.8(a), represents the term

$$\langle ji|v|lk\rangle a_j^\dagger a_i^\dagger a_k a_l. \tag{7.37}$$

Now, by the commutation relations for the operators, we have

$$a_j^\dagger a_i^\dagger a_k a_l = a_i^\dagger a_j^\dagger a_l a_k, \tag{7.38}$$

provided $i \neq j$ and $k \neq l$ (and, if either $i = j$ or $k = l$, the matrix element $\langle \Phi_0 | a_i^\dagger a_j^\dagger a_l a_k | \Phi_0 \rangle$ vanishes and the term is of no interest). Also, from eq. (1.51), we have

$$\langle ji|v|lk\rangle = \langle ij|v|kl\rangle. \tag{7.39}$$

The terms (7.36) and (7.37) are thus one and the same, and the two graphs in fig. 7.8 represent the same term.

On the other hand, the graph shown in fig. 7.9, obtained by interchanging two *lines* at opposite vertices in fig. 7.8(a), represents the term

$$\langle ji|v|kl\rangle a_j^\dagger a_i^\dagger a_l a_k = -\langle ij|v|lk\rangle a_i^\dagger a_j^\dagger a_l a_k, \qquad (7.40)$$

which in general is different from (7.36).

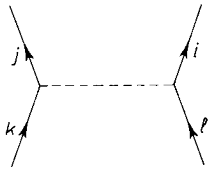


Fig. 7.9.

It must be understood that, although (7.36) and (7.37) are the same, *both* appear in $H_{\mathrm{I}}(t)$, eq. (7.17). When summing the terms of $H_{\mathrm{I}}(t)$ one may sum over all graphs representing *different* terms and then multiply by 2 – in other words, omit the factor $\frac{1}{2}$ in (7.17).

It would seem from the examples given that the representation of terms of $H_{\mathrm{I}}(t)$ by a picture showing the Fermi surface and excited particles, etc., is simpler and more obvious than the use of Feynman graphs. This is undoubtedly true, but the full beauty of the Feynman graphs is only realized when higher orders are considered, that is to say, when we represent graphically the terms of a product of operators $H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\ldots H_{\mathrm{I}}(t_n)$ taken at different times. The use of the Fermi surface representation would in such cases be difficult and confusing.

7.4.2. *Higher-order graphs*

U_n, eq. (6.25), contains a chronologically ordered product of n interaction operators, each of which is a sum of terms as in eq. (7.17). In order to represent graphically a single term of the multiple sum in U_n we draw n interaction lines down the page, one for each $H_{\mathrm{I}}(t)$, starting with t_1 at the top, t_2 below that, and so on. We then draw particle and hole lines at each vertex and label them, as explained above, so that we have a graphical representation of the contribution from each interaction operator, in the correct order, to this particular term of U_n. Fig. 7.10 shows a possible term of U_3.

Since time increases from the bottom to the top of the page (i.e. $t_1 > t_2 > \ldots > t_n$), and upward lines refer to particles, while downward lines refer to holes, it is often said that particles travel forward in time and holes travel backwards in time.

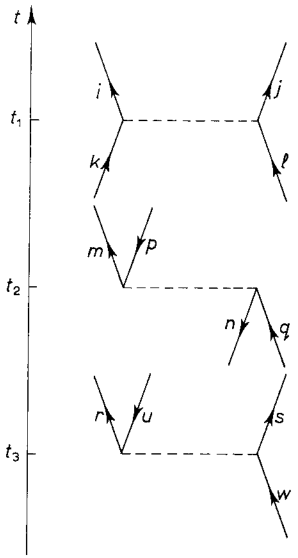


Fig. 7.10. Graphical representation of a possible term of U_3.

The greatest value of the graphical representation lies in classifying those terms which contribute to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$, required in calculating the ground-state energy. The general term A_n is shown in eq. (7.22) and consists of a $4n$-fold infinite sum of terms, each of which can be represented by a graph. The important point is that the matrix element[†]

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k \ldots c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle, \qquad (7.41)$$

as has been pointed out before, vanishes if the destruction and creation operators do not balance in pairs. This means that, if the matrix element is not to vanish, for every particle or hole creation line there must be a corresponding destruction line. If we join up creation and destruction lines for the same particle or hole, whether they occur at the same vertex (forming a loop) or at different vertices, therefore, we may say that *every line must begin and end at a vertex.*

[†] Having defined the Feynman graphs, we shall in future make little explicit use of the particle–hole notation. That is to say, we shall in general use $c_i^\dagger$ without stating explicitly whether this is $a_i^\dagger$ or b_i, and c_i without stating explicitly whether this is a_i or $b_i^\dagger$. In specific examples the direction of the line representing a given state in a Feynman graph will indicate whether it is a particle or a hole state.

Furthermore, since Φ_0 contains neither holes nor excited particles, the matrix element (7.41) will vanish unless the first two operators $c_w c_u$, at the lowest vertices, create holes. Also, at the same vertices, the operators $c_r^\dagger c_s^\dagger$ must either destroy these holes again, or destroy one hole and create one particle, or create two particles. In all cases we either have lines *above* the lowest interaction line or loops. Similarly, the last two operators $c_i^\dagger c_j^\dagger$, at the highest vertices, must destroy holes, and the operators $c_l c_k$ must either create these same holes, or create one of them and destroy a particle, or destroy a pair of particles. In all cases we either have lines *below* the highest interaction line or loops. We deduce, therefore, that there must be no lines below the lowest interaction line or above the highest interaction line, unless they happen to be loops.

All this simplifies considerably the construction of the graphs which contribute to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$. Typical third-order graphs, which *may* contribute to A_3, are shown in fig. 7.11. Here and in future, where we illustrate a *type* of graph, we shall not label the lines.

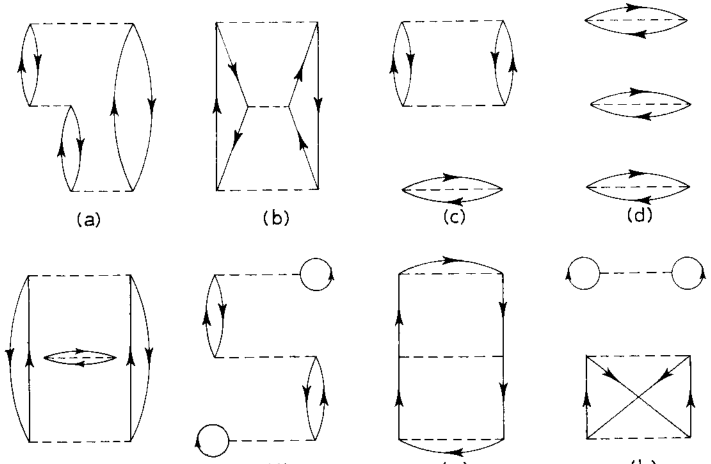


Fig. 7.11. Typical third-order graphs, representing terms which may contribute to A_3.

It should be noted that the vertical ordering of the interaction lines is important, but not their lengths or horizontal positions. Thus, an interaction line may be shortened, lengthened, or moved to right or left, for convenience of drawing, without changing the significance of a graph. For example, fig. 7.11(e) could equally well be drawn as in fig. 7.12.

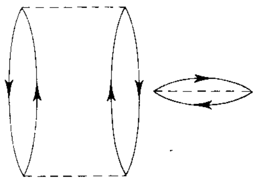


Fig. 7.12. Alternative way of drawing fig. 7.11(e).

It will be shown later that not even all graphs of the type described above contribute to the energy of the ground state. First, however, it might be helpful to give some concrete examples of the evaluation of terms from graphs.

7.5. Examples of the Evaluation of the Contributions of Various Graphs to the Perturbation Series

The general term A_n of the series $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ is shown in eq. (7.22). We shall consider the contributions to A_1, A_2 and A_3 from some typical graphs.

7.5.1. *First-order graphs*

According to the rules of the previous section, the only possible first-order graphs which contribute to A_1 are of the types shown in fig. 7.13.

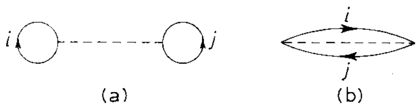


Fig. 7.13. First-order graphs which may contribute to A_1.

The term of A_1, eq. (7.18), corresponding to fig. 7.13(a) is

$$(-\tfrac{1}{2}\mathrm{i})\langle ij|v|ij\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_j c_i|\Phi_0\rangle\frac{e^{\alpha t}}{\alpha}, \tag{7.42}$$

since $\Delta_1 = 0$ here. This vanishes if $i = j$, since

$$c_i c_i \Phi_0 = 0, \tag{7.43}$$

and, if $i \neq j$, the commutation rules for the creation and destruction operators give

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_j c_i|\Phi_0\rangle = \langle\Phi_0|c_i^\dagger c_i c_j^\dagger c_j|\Phi_0\rangle = 1. \tag{7.44}$$

Also, since Φ_0 contains neither holes nor excited particles, the operators c_i, c_j must create holes, and $c_i^\dagger, c_j^\dagger$ destroy them for a non-zero result: that is to say, we must have $\epsilon_i, \epsilon_j \leq \epsilon_F$.

This term is known as a *direct* term (two unexcited particles interact without changing their states – i goes to i and j to j). The total contribution to A_1 from all direct terms is thus

$$(-\tfrac{1}{2}\mathrm{i})\sum_{\substack{i,j\neq i\\(\epsilon_i,\epsilon_j\leq\epsilon_F)}}\langle ij|v|ij\rangle\frac{\mathrm{e}^{\alpha t}}{\alpha}. \tag{7.45}$$

The term of A_1 corresponding to fig. 7.13(b) is

$$(-\tfrac{1}{2}\mathrm{i})\langle ij|v|ji\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_i c_j|\Phi_0\rangle\frac{\mathrm{e}^{\alpha t}}{\alpha}. \tag{7.46}$$

Here again we must have $i \neq j$, and

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_i c_j|\Phi_0\rangle = -\langle\Phi_0|c_i^\dagger c_i c_j^\dagger c_j|\Phi_0\rangle = -1, \tag{7.47}$$

with $\epsilon_i, \epsilon_j \leq \epsilon_F$.

This term is known as an *exchange* term (two unexcited particles 'exchange' their states – i goes to j and j to i). The total contribution to A_1 from all exchange terms is thus

$$-(-\tfrac{1}{2}\mathrm{i})\sum_{\substack{i,j\neq i\\(\epsilon_i,\epsilon_j\leq\epsilon_F)}}\langle ij|v|ji\rangle\frac{\mathrm{e}^{\alpha t}}{\alpha}. \tag{7.48}$$

Adding (7.45) and (7.48), we find

$$A_1 = (-\tfrac{1}{2}\mathrm{i})\sum_{\substack{i,j\\(\epsilon_i,\epsilon_j\leq\epsilon_F)}}[\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]\frac{\mathrm{e}^{\alpha t}}{\alpha}. \tag{7.49}$$

It is unnecessary to specify $i \neq j$ here, since the expression in square brackets vanishes if $i = j$. In other words, *for every direct term which contradicts the Pauli principle*[†] *there is an exchange term which cancels it.* We shall see that this also applies to higher-order graphs.

Inserting (7.49) in eq. (6.30) we have immediately, for the first-order energy correction,

$$\Delta E^{(1)} = \tfrac{1}{2}\sum_{\substack{i,j\\(\epsilon_i,\epsilon_j\leq\epsilon_F)}}[\langle ij|v|ij\rangle - \langle ij|v|ji\rangle]. \tag{7.50}$$

[†] The Pauli principle demands that $i \neq j$, so that no state is destroyed or created twice in succession.

7.5.2. *Second-order graphs*

According to the rules of §7.4.2 the only graphs which may contribute to A_2, eq. (7.20), are of the types shown in fig. 7.14. We shall show that, for a free-electron gas at the absolute zero of temperature at least, the only graphs which contribute to the *ground-state energy* are those in the first row, and, of these, only (a) and (c) need be explicitly considered, since (b) is obtained from (a) and (d) is obtained from (c) by interchanging the vertices at one interaction line. Let us therefore consider the contribution to A_2 from the two graphs shown in fig. 7.15. It is to be understood, in all graphs, that the time axis runs from the bottom to the top of

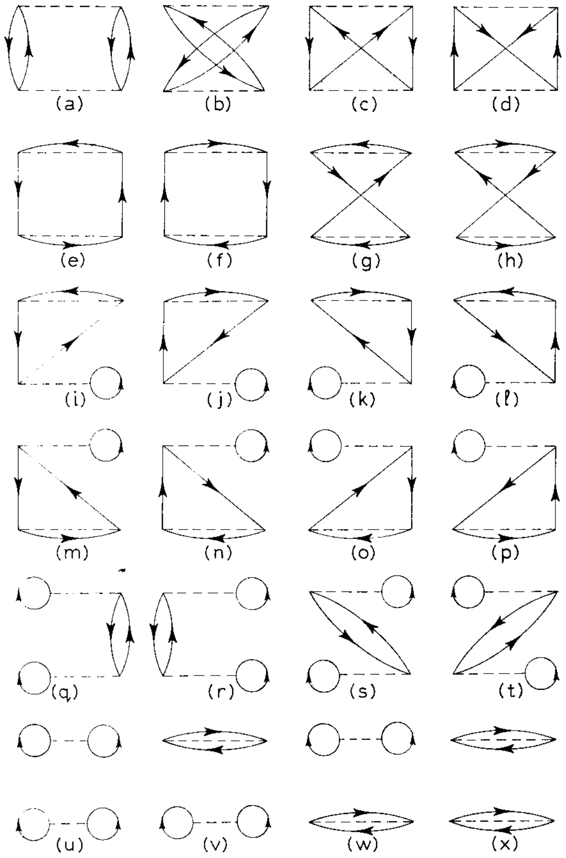


Fig. 7.14. Second-order graphs which may contribute to A_2.

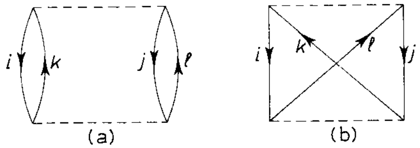


Fig. 7.15.

the page, so that the highest interaction line corresponds to $H_I(t_1)$, the next lower one to $H_I(t_2)$, and so on.

The term of A_2, eq. (7.20), corresponding to fig. 7.15(a) is

$$(-\tfrac{1}{2}\mathrm{i})^2\langle ij|v|kl\rangle\langle kl|v|ij\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j c_i|\Phi_0\rangle\frac{\mathrm{e}^{2\alpha t}}{2\alpha(\mathrm{i}\Delta_2+\alpha)}, \tag{7.51}$$

where

$$\Delta_2=\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j=-\Delta_1. \tag{7.52}$$

For a non-zero result we must have $i\neq j$, $k\neq l$. Furthermore, since the i, j operators are hole operators (*downward* pointing lines) and the k, l operators are particle operators (*upward* pointing lines), so that $\epsilon_i,\epsilon_j\leq\epsilon_F$ and $\epsilon_l,\epsilon_k>\epsilon_F$, it follows that i, j, k, l must all be different. The commutation relations then give†

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j c_i|\Phi_0\rangle=1. \tag{7.53}$$

This term is again called a *direct* term (one particle goes from state i to state k at time t_2 and back again to i at later time t_1, while the other particle goes from j to l at t_2 and back to j at t_1, so that finally the particles are in the same states as they were originally), and fig. 7.15(a) is called a *direct graph*. The total contribution to A_2 from terms with direct graphs of this type is thus

$$(-\tfrac{1}{2}\mathrm{i})^2\sum_{\substack{i,j\neq i\\(\epsilon_i,\epsilon_j\leq\epsilon_F\\\epsilon_k,\epsilon_l>\epsilon_F)}}\sum_{k,l\neq k}\frac{|\langle ij|v|kl\rangle|^2\,\mathrm{e}^{2\alpha t}}{2\alpha[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+\alpha]}, \tag{7.54}$$

since

$$\langle ij|v|kl\rangle\langle kl|v|ij\rangle=|\langle ij|v|kl\rangle|^2.$$

The term of A_2 corresponding to fig. 7.15(b) is

$$(-\tfrac{1}{2}\mathrm{i})^2\langle ij|v|kl\rangle\langle lk|v|ij\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_l^\dagger c_k^\dagger c_j c_i|\Phi_0\rangle\frac{\mathrm{e}^{2\alpha t}}{2\alpha(\mathrm{i}\Delta_2+\alpha)}, \tag{7.55}$$

and

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_l^\dagger c_k^\dagger c_j c_i|\Phi_0\rangle=-1, \tag{7.56}$$

with the same conditions on i, j, k and l as before.

This term is called an *exchange* term (one particle goes from i to l at t_2 and then to j at t_1, while the other particle goes from j to k at t_2 and then to i at t_1, so that the particles have finally 'exchanged' their original states), and fig. 7.15(b) is called an *exchange graph*. The total contribution to A_2 from terms with exchange graphs of this type is thus

$$-(-\tfrac{1}{2}\mathrm{i})^2\sum_{\substack{i,j\neq i\\(\epsilon_i,\epsilon_j\leq\epsilon_F,\\\epsilon_k,\epsilon_l>\epsilon_F)}}\sum_{k,l\neq k}\frac{\langle ij|v|kl\rangle\langle lk|v|ij\rangle\,\mathrm{e}^{2\alpha t}}{2\alpha[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+\alpha]}. \tag{7.57}$$

The terms of A_2 corresponding to the graphs shown in fig. 7.16(a) and fig. 7.16(b) are again just (7.51) and (7.55) respectively, since

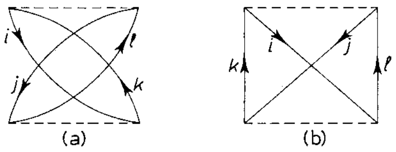


Fig. 7.16.

$$\langle lk|v|ji\rangle=\langle kl|v|ij\rangle \tag{7.58}$$

and

$$\langle kl|v|ji\rangle=\langle lk|v|ij\rangle. \tag{7.59}$$

Hence the total contribution to A_2 from all graphs of the types shown in

---

† An even number of interchanges of adjacent operators leaves the sign of the matrix element unchanged, while an odd number changes the sign. Thus

$$\begin{aligned}\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j c_i|\Phi_0\rangle&=\langle\Phi_0|c_i^\dagger c_i c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j|\Phi_0\rangle &&\text{(6 interchanges)}\\&=\langle\Phi_0|c_i^\dagger c_i c_j^\dagger c_j c_l c_k c_k^\dagger c_l^\dagger|\Phi_0\rangle &&\text{(4 interchanges)}\\&=\langle\Phi_0|c_i^\dagger c_i c_j^\dagger c_j c_l c_l^\dagger c_k c_k^\dagger|\Phi_0\rangle &&\text{(2 interchanges)}\\&=1.\end{aligned}$$

Notice that, since c_k, $c_k^\dagger$ are particle operators,

$$c_k^\dagger c_k\Phi_0=0,$$

but

$$c_k c_k^\dagger\Phi_0=(1-c_k^\dagger c_k)\Phi_0=\Phi_0.$$

the first row of fig. 7.14(a)–(d) is just twice the sum of (7.54) and (7.57), namely

$$(-\tfrac{1}{2}\mathrm{i})^2 \sum_{\substack{ijkl \\ (\epsilon_i,\epsilon_j\leq\epsilon_F, \\ \epsilon_l,\epsilon_k>\epsilon_F)}} \frac{[|\langle ij|v|kl\rangle|^2-\langle ij|v|kl\rangle\langle lk|v|ij\rangle]\,\mathrm{e}^{2\alpha t}}{\alpha[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+\alpha]}. \tag{7.60}$$

Here it is not necessary to specify $i \neq j$, $k \neq l$, since the numerator vanishes when either $i=j$ or $k=l$. In other words, as for the first-order terms, we see that *for every direct term which contradicts the Pauli principle there is an exchange term which cancels it.*

We shall show later that (7.60) gives the total contribution to the second-order energy correction $\Delta E^{(2)}$ for a free-electron gas, the contribution from all other graphs being either disallowed or cancelling one another.

7.5.3. *Third-order graphs*

The number of types of third-order graphs, which may contribute to A_3, is too great for us even to draw them all, although the number contributing to the ground-state energy is again comparatively small. We shall consider only a few third-order graphs, which will serve to illustrate some important points.

From eq. (7.22) we have

$$A_3 = (-\tfrac{1}{2}\mathrm{i})^3 \sum_{ijkl}\sum_{mnpq}\sum_{rsuw} \langle ij|v|kl\rangle\langle mn|v|pq\rangle\langle rs|v|uw\rangle$$

$$\times \langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_m^\dagger c_n^\dagger c_q c_p c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle$$

$$\times \frac{\exp\{[\mathrm{i}(\Delta_1+\Delta_2+\Delta_3)+3\alpha]t\}}{[\mathrm{i}(\Delta_1+\Delta_2+\Delta_3)+3\alpha][\mathrm{i}(\Delta_2+\Delta_3)+2\alpha](\mathrm{i}\Delta_3+\alpha)}, \tag{7.61}$$

where

$$\Delta_1 = \epsilon_i+\epsilon_j-\epsilon_k-\epsilon_l, \tag{7.62}$$

$$\Delta_2 = \epsilon_m+\epsilon_n-\epsilon_p-\epsilon_q, \tag{7.63}$$

$$\Delta_3 = \epsilon_r+\epsilon_s-\epsilon_u-\epsilon_w. \tag{7.64}$$

The term of this sum corresponding to the graph in fig. 7.17 is

$$(-\tfrac{1}{2}\mathrm{i})^3 \frac{\langle ij|v|kl\rangle\langle kn|v|im\rangle\langle ml|v|nj\rangle\,\mathrm{e}^{3\alpha t}}{3\alpha[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+2\alpha][\mathrm{i}(\epsilon_m+\epsilon_l-\epsilon_n-\epsilon_j)+\alpha]}, \tag{7.65}$$

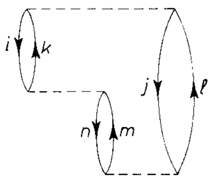


Fig. 7.17. A third-order graph contributing to A_3.

since, in this case,

$$\Delta_1+\Delta_2+\Delta_3 = \epsilon_i+\epsilon_j-\epsilon_k-\epsilon_l+\epsilon_k+\epsilon_n-\epsilon_i-\epsilon_m$$
$$+\epsilon_m+\epsilon_l-\epsilon_n-\epsilon_j = 0, \tag{7.66}$$

and

$$\Delta_2+\Delta_3 = -\Delta_1 = \epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j. \tag{7.67}$$

The energy factor Δ for any interaction is obtained by adding the energies for the outgoing lines and subtracting the energies for the incoming lines. It follows that, as illustrated by eq. (7.66), the sum of *all* the energy factors for *any* allowed graph must be zero, since every line must begin and end at a vertex, so that its energy must occur twice with opposite signs.

Also, for a non-zero result, we must have $i \neq j \neq n$ and $k \neq l \neq m$, and, since i, j, n are hole states while k, l, m are particle states, it follows that i, j, k, l, m, n must all be different. Hence

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_n^\dagger c_m c_i c_m^\dagger c_l^\dagger c_j c_n|\Phi_0\rangle = 1. \tag{7.68}$$

Further examples of third-order graphs will be given in the following section.

7.6. Linked and Unlinked Graphs

A *linked graph* is defined as a graph which can be drawn without removing the pen from the paper (except to dash the interaction lines). An *unlinked graph* is one which cannot be so drawn, but consists of two or more quite separate parts, not connected by lines. For example, in fig. 7.11, the graphs (a), (b), (f) and (g) are linked, all the others being unlinked. In fig. 7.14 all the graphs are linked except those in the last row ((u), (v), (w) and (x)). Clearly, each unlinked graph consists of several lower-order linked graphs, which we will call *unlinked parts*†.

---

† The terminology in this part of the subject is apt to be confusing, because of the different uses of the words *linked* and *unlinked*. An unlinked part, for example, is unlinked from the rest of the graph, but it is *internally* linked, so to speak. For the latter reason it is, unfortunately, often called a *linked* part. The word *connected* is sometimes used instead of *linked*.

We shall show in the next chapter that only linked graphs need be considered in calculating the energy of the ground state. Meanwhile, however, we shall give some examples which will illustrate an important property of unlinked graphs.

Let us consider the three unlinked third-order graphs in fig. 7.18. Notice that each of these graphs consists of the same two unlinked parts – only the *order* of the interactions varies from graph to graph.

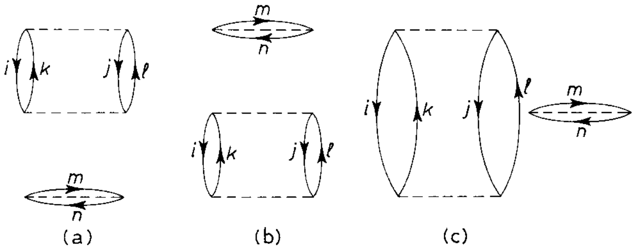


Fig. 7.18.

The contribution to A_3, eq. (7.61), from the graph in fig. 7.18(a) is

$$(-\tfrac{1}{2}\mathrm{i})^3\langle ij|v|kl\rangle\langle kl|v|ij\rangle\langle mn|v|nm\rangle\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j c_i c_m^\dagger c_n^\dagger c_m c_n|\Phi_0\rangle$$
$$\times\frac{\mathrm{e}^{3\alpha t}}{3\alpha^2[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+2\alpha]}. \qquad (7.69)$$

Now i,j are hole states and k,l are particle states, so that $i,j \neq k,l$. For a non-zero contribution we must have $i\neq j$, $k\neq l$, $m\neq n$. We shall also assume that m,n are hole states (otherwise the first two graphs, at least, give zero contribution), and that $i,j\neq m,n$ (otherwise the last graph gives zero contribution). In other words, we shall suppose that i, j, k, l, m, n are all different. Then the commutation relations give

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k c_k^\dagger c_l^\dagger c_j c_i c_m^\dagger c_n^\dagger c_m c_n|\Phi_0\rangle = -1, \qquad (7.70)$$

and the corresponding matrix elements for the other two graphs have the same value. Thus the contribution to A_3 from fig. 7.18(b) is

$$-(-\tfrac{1}{2}\mathrm{i})^3\langle mn|v|nm\rangle\langle ij|v|kl\rangle\langle kl|v|ij\rangle\frac{\mathrm{e}^{3\alpha t}}{6\alpha^2[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+\alpha]}, \qquad (7.71)$$

and the contribution from fig. 7.18(c) is

$$-(-\tfrac{1}{2}\mathrm{i})^3\langle ij|v|kl\rangle\langle mn|v|nm\rangle\langle kl|v|ij\rangle$$
$$\times\frac{\mathrm{e}^{3\alpha t}}{3\alpha[\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+2\alpha][\mathrm{i}(\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j)+\alpha]}. \qquad (7.72)$$

The sum of the contributions to A_3 from the three graphs, (7.69), (7.71) and (7.72), is

$$-(-\tfrac{1}{2}\mathrm{i})^3\langle ij|v|kl\rangle\langle mn|v|nm\rangle\langle kl|v|ij\rangle$$
$$\times\frac{\mathrm{e}^{3\alpha t}}{3\alpha}\left[\frac{1}{\alpha(\mathrm{i}\Delta+2\alpha)}+\frac{1}{2\alpha(\mathrm{i}\Delta+\alpha)}+\frac{1}{(\mathrm{i}\Delta+2\alpha)(\mathrm{i}\Delta+\alpha)}\right]$$
$$=-(-\tfrac{1}{2}\mathrm{i})^3\langle ij|v|kl\rangle\langle mn|v|nm\rangle\langle kl|v|ij\rangle\frac{\mathrm{e}^{3\alpha t}}{2\alpha^2(\mathrm{i}\Delta+\alpha)}, \qquad (7.73)$$

where

$$\Delta=\epsilon_k+\epsilon_l-\epsilon_i-\epsilon_j. \qquad (7.74)$$

Now, let us consider the two unlinked parts of these three graphs separately, that is, the two graphs shown in fig. 7.19.

Fig. 7.19(a) is a first-order graph whose contribution to A_1 is, from eqs. (7.46) and (7.47),

$$-(-\tfrac{1}{2}\mathrm{i})\langle mn|v|nm\rangle\frac{\mathrm{e}^{\alpha t}}{\alpha}, \qquad (7.75)$$

provided m and n are hole states and $m\neq n$.

Fig. 7.19(b) is a second-order graph whose contribution to A_2 is, from eqs. (7.51) and (7.53),

$$(-\tfrac{1}{2}\mathrm{i})^2\langle ij|v|kl\rangle\langle kl|v|ij\rangle\frac{\mathrm{e}^{2\alpha t}}{2\alpha(\mathrm{i}\Delta+\alpha)}, \qquad (7.76)$$

provided $i\neq j$, $k\neq l$.

The product of (7.75) and (7.76) is immediately seen to be just (7.73), which is the sum of the contributions from the three unlinked graphs formed by combining figs. 7.19(a) and (b) with all possible relative orderings of their interactions. We shall now go on to consider the total contribution from *all* graphs of the types shown in fig. 7.18, but first it may be advisable to state more precisely what we mean by the *type* of a graph.

It must be remembered that it is the *topology* of a graph which is important, rather than its shape or size. We shall call two graphs *topo-*

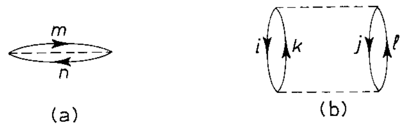


Fig. 7.19.

logically equivalent if, disregarding the labelling of the hole and particle lines, one can be distorted into the other without changing the time-ordering of the interaction lines or interchanging vertices[†]. For example, the two graphs in fig. 7.20 are topologically equivalent.

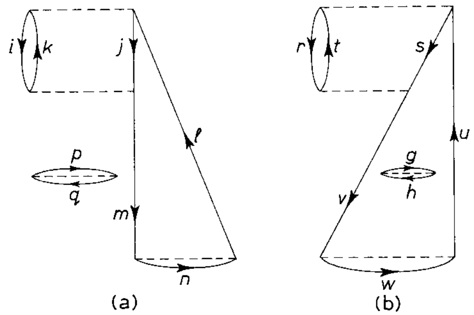


Fig. 7.20. Topologically equivalent fourth-order graphs.

We now state, as a definition, that two graphs are of the same *type* if they are topologically equivalent. Thus the type of a graph is represented simply by the graph itself (or a topologically equivalent one) with arrows but no labels on its hole and particle lines. Two graphs of the same type are *different* if the labels of their hole and particle lines are different (although, owing to the fact mentioned at the end of §7.4.1, regarding the inversion of vertices, they may represent the same term).

We are interested in the total contribution to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$, eq. (6.26), from all different graphs of given types. It should be clear from the above discussion that (provided we ignore certain possible violations of the Pauli principle, as will be explained and justified in §7.8 below) the total contribution from all different graphs of the types shown in fig. 7.18

[†] Sometimes graphs are also said to be topologically equivalent if one may be obtained from the other by interchanging opposite vertices (that is, by inverting the ends of one or more interaction lines). As may be deduced from the remarks at the end of §7.4.1, the only difference this makes is to introduce factors of 2 into the final summing of the contributions from all graphs of a given type.

is just the product[†] of the total contribution from all graphs of the type shown in fig. 7.19(a) and the total contribution from all graphs of the type shown in fig. 7.19(b). We may represent this result symbolically as in fig. 7.21, where an unlabelled graph stands for the sum of the contributions to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ from *all* different graphs of that type.

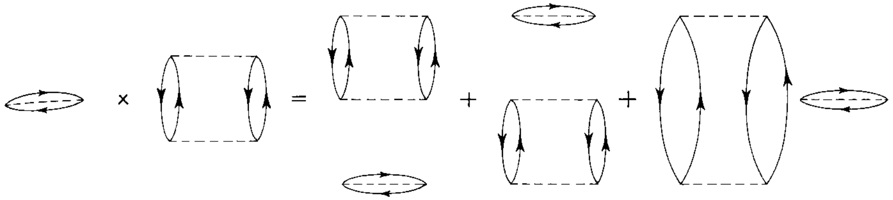


Fig. 7.21.

We shall generalize this result in the following chapter and show that it applies to any set of unlinked graphs containing the same unlinked parts, all possible relative orderings of the latter being included, *provided that these unlinked parts are all of different types.*

A slight modification is necessary if several unlinked parts of a graph are of the same type. For example, consider the graph shown in fig. 7.22,

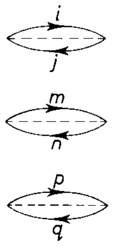


Fig. 7.22.

which has three unlinked parts, each of the same type. The contribution to A_3 from all such graphs is, from eq. (7.61),

$$-(-\tfrac{1}{2}\mathrm{i})^3 \sum_{ijklmn} \langle ij|v|ji\rangle\langle kl|v|lk\rangle\langle mn|v|nm\rangle\frac{\mathrm{e}^{3\alpha t}}{6\alpha^3}. \tag{7.77}$$

However, the cube of the contribution to A_1 from all graphs of the type shown in fig. 7.19(a), which is the type of each unlinked part of fig. 7.22,

---

[†] This follows from the fact that

$$\sum_{ijklmn} \langle ij|v|kl\rangle\langle mn|v|nm\rangle\langle kl|v|ij\rangle = \sum_{ijkl} \langle ij|v|kl\rangle\langle kl|v|ij\rangle \sum_{mn} \langle mn|v|nm\rangle.$$

is, from eq. (7.75),

$$-(-\tfrac{1}{2}\mathrm{i})^3\left[\sum_{ij}\langle ij|v|ji\rangle\right]^3\frac{\mathrm{e}^{3\alpha t}}{\alpha^3}, \tag{7.78}$$

which is 6 or 3! times the expression (7.77). We may represent this result symbolically as in fig. 7.23, an unlabelled graph again standing for the sum of the contributions to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ from *all* different graphs of that type.

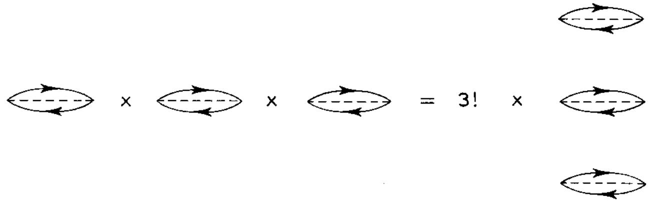


Fig. 7.23.

The reason for this result is simply that permuting unlinked parts of the same type does not change the type of an unlinked graph, and there are 3! permutations of the unlinked parts of the graph in fig. 7.22. The generalization of this is fairly obvious and will be discussed in the following chapter.

7.7. The Vacuum–Vacuum Matrix Element

We shall in future refer to the matrix element

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k \ldots c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle$$

appearing in any term of A_n, eq. (7.22), whatever the value of n, as the *vacuum–vacuum matrix element* for that term, since Φ_0 is the unperturbed ground state, which, as explained in §7.2, plays the part of the vacuum state in the present work. We shall also denote this matrix element symbolically by $\langle\ \ \rangle$.

The possible values of $\langle\ \ \rangle$ are 0 and ±1. It vanishes, in fact, only in cases where the Pauli principle is violated, that is, where the same one-electron state is created or destroyed twice in succession. Apart from such cases, which we shall ignore (for reasons to be discussed in the following section), the value of $\langle\ \ \rangle$ can be found very simply from the graph of the term in which it appears. The rule is as follows:

If, for a given graph, λ *is the number of closed loops formed by the hole and particle lines, and* μ *is the number of hole lines, then the value of the vacuum–vacuum matrix element for this graph is* $(-1)^{\lambda+\mu}$.

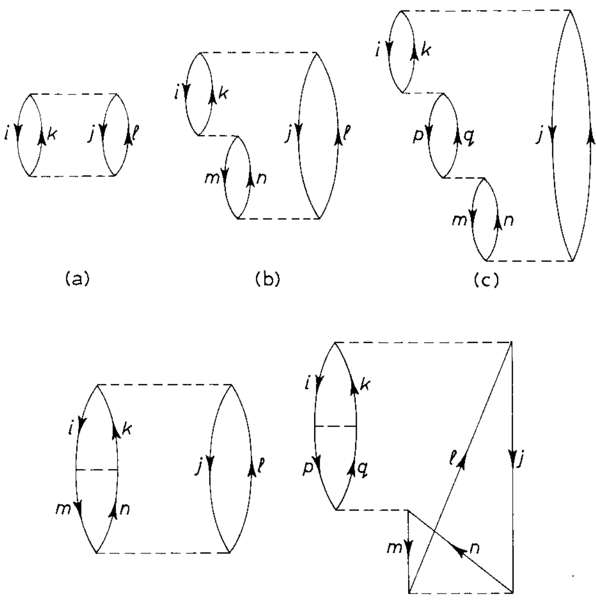
We shall not prove this rule rigorously[†], but shall make it plausible. First, we note that (again ignoring certain possible violations of the Pauli principle) the value of the vacuum–vacuum matrix element for an unlinked graph is just the product of the values for the unlinked parts. This is obvious in the case of unlinked parts which are completely separate in time, as in figs. 7.18(a) and (b). When there is overlap, as in fig. 7.18(c), it is only necessary to realize that the interchange of two interactions involves the interchange of two blocks of four operators in $\langle\ \ \rangle$. This entails an *even* number of interchanges of neighbouring operators, and hence does not change the sign of $\langle\ \ \rangle$. It follows that the value of $\langle\ \ \rangle$ is the same for all unlinked graphs obtained by permuting the same set of unlinked parts. It also follows that, in verifying the rule for evaluating $\langle\ \ \rangle$, we need only consider linked graphs.

Secondly, it may be verified immediately that the rule applies to all the graphs whose contributions we have evaluated, namely, figs. 7.13, 7.15 and 7.17. In fig. 7.13(a) there are two loops and two hole lines[‡], giving $\langle\ \ \rangle = 1$, while in fig. 7.13(b) there is one loop and two hole lines, giving $\langle\ \ \rangle = -1$. Similarly, the graph in fig. 7.15(a) has two loops and two hole lines, giving $\langle\ \ \rangle = 1$, while that in fig. 7.15(b) has one loop (which intersects itself) and two hole lines, giving $\langle\ \ \rangle = -1$. Finally, the graph in fig. 7.17 has three loops and three hole lines, giving $\langle\ \ \rangle = 1$.

Thirdly, we note that when two lines at opposite vertices (that is, at the two ends of an interaction line) are interchanged, as in the case of the i, j lines in going from fig. 7.15(a) to fig. 7.15(b), the sign of $\langle\ \ \rangle$ is changed. This is because the two lines interchanged must either both leave or both enter their respective vertices, so that in the block of four operators, $c_i^\dagger c_j^\dagger c_l c_k$ say, coming from this interaction, either $c_i^\dagger, c_j^\dagger$ must be interchanged or c_l, c_k. Interchanging two lines in this manner always adds or subtracts a closed loop, as seen in fig. 7.15, but does not affect the number of hole lines. This strongly suggests that adding or subtracting a closed loop, leaving the number of hole lines unchanged, changes the sign of the vacuum–vacuum matrix element, as implied by the rule.

[†] The rigorous proof entails the use of Wick's theorem, stated in Appendix VII. However, this requires tedious and complicated algebra which it is the principal purpose of the graphical method to avoid. We therefore prefer not to make explicit use of Wick's theorem, even at the occasional cost of a little rigour.

[‡] Loops of the type shown in fig. 7.13 always refer to *holes*, as we have seen in §7.5. In order to conform with the rule, circular loops of the type of fig. 7.13(a), when they occur singly, must be assumed to include one hole line, so that such loops do not contribute to the sign of $\langle\ \ \rangle$.



(a) (b) (c)

(d) (e)

Fig. 7.24.

In order to obtain the dependence of the sign of $\langle\ \ \rangle$ upon the number of hole lines, let us consider the graphs in fig. 7.24. Those in the upper row, (a) to (c), are all 'direct' graphs, as explained in §7.5. They are also known as *ring graphs*, for an obvious reason, and we shall find them to be very important in calculating the correlation energy of a free-electron gas (see ch. 9). Going from one ring graph to that of next higher order adds a loop and a hole line. Each time this is done a similar block of four operators is inserted into $\langle\ \ \rangle$, so that an additional *even* number of interchanges of adjacent operators is required to bring those relating to the same state into juxtaposition. In fact, $\langle\ \ \rangle = 1$ for all ring graphs, so that adding a loop *and* a hole line does not change the sign of $\langle\ \ \rangle$.

Now consider the third-order 'exchange' graph shown in fig. 7.24(d). This is obtained from (b) by interchanging the k,m lines at the t_2 interaction, the sign of $\langle\ \ \rangle$ thus becoming negative. We therefore find that, in going from (a) to (d), the number of loops remains the same, but the number of hole lines increases by one, and the sign of $\langle\ \ \rangle$ is changed. Similarly, (e) is obtained from (c) by interchanging the k,p lines at the t_2 interaction and the n,l lines at the t_4 interaction. Since there are two interchanges, the sign of $\langle\ \ \rangle$ remains the same in (c) and (e). Thus $\langle\ \ \rangle = -1$ for (d) and $\langle\ \ \rangle = 1$ for (e), but (d) and (e) both have two loops and differ only in that (e) has an additional hole line. We deduce that the sign of the matrix element depends upon the number of hole lines, as well as upon the number of loops, as stated in the rule.

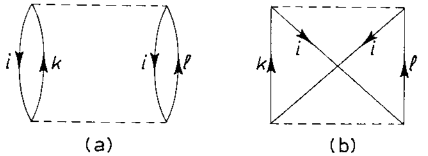
Although not completely rigorous, the foregoing verification should be sufficiently convincing. We have gone into the question at some length because it receives scant attention in most texts. Generally, even when the rule is correctly stated, no proof is attempted or it is left implicit in a mass of algebra.

7.8. Graphs which Violate the Pauli Principle

In the foregoing sections we stated in several places that we were ignoring certain possible violations of the Pauli principle. We shall now explain the meaning of this statement and justify the procedure.

A graph is said to violate the Pauli principle if it represents a term in which the same one-electron state is created or destroyed twice in succession. The vacuum–vacuum matrix element $\langle\ \ \rangle$ for such a term must vanish. However, it would be convenient not to have to take such exceptional cases into account when summing over graphs, but to assume that all graphs of the same type, whatever the labelling of their hole and particle lines, have the same value of $\langle\ \ \rangle$. This can, in fact, be done.

We have seen in §7.5 that, in summing over first-order and second-order graphs, the Pauli principle can be ignored, because any direct graph which violates it is automatically cancelled by an exchange graph which represents a term of the same magnitude but opposite sign. For example, the two graphs in fig. 7.25 cancel each other. The vacuum–vacuum matrix element should really vanish for these two graphs, because the hole state i is created twice in succession. However, if this fact is ignored, and $\langle\ \ \rangle$ is evaluated according to the rule given in the previous section, its value for fig. 7.25(a) is 1, and its value for fig. 7.25(b) is -1. Apart from this the terms represented by the two graphs are the same, since corresponding

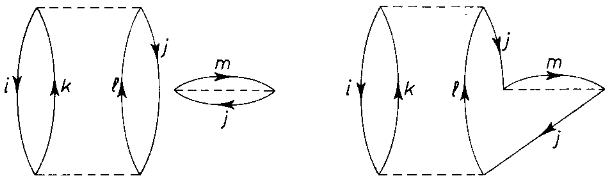


(a) (b)

Fig. 7.25. Two second-order graphs which violate the Pauli principle, but which cancel each other.

interactions are the same in both graphs and only the method of joining up the i lines is different.

It is not difficult to see that, whenever the same hole or particle line appears twice in a graph, it is possible to join up the relevant creation and destruction lines in two alternative ways, one of which has an extra loop. Both graphs have the same number of hole lines, so that the sign of $\langle \quad \rangle$ is different for the two graphs and their contributions therefore cancel.



(a) (b)

Fig. 7.26. Two third-order graphs which violate the Pauli principle, but which cancel each other.

The two graphs shown in fig. 7.26 illustrate a further important point. Both graphs violate the Pauli principle, but, if this fact is ignored, the value of $\langle \quad \rangle$ is found to be 1 for (a) and -1 for (b). Clearly, the two graphs differ only in the way the j hole lines are connected up, so that, if both are included in the sum over graphs, they cancel each other. Now, fig. 7.26(a) is of the same type as fig. 7.18(c), with $n = j$. If, indeed, we take $n = j$ in all the graphs of figs. 7.18 and 7.19, and follow through the same calculation as in §7.6, we find that the sum of the contributions from the three graphs in fig. 7.18 is equal to the product of the contributions from 7.19(a) and 7.19(b) *only if the value of the vacuum–vacuum matrix element is assumed to be the same for all the graphs in fig. 7.18*. In fact, the value of this matrix element should be zero for fig. 7.18(c), with $n = j$, since this graph violates the Pauli principle. Thus the situation depicted in fig. 7.21 is true *only if violations of the Pauli principle are ignored.* We shall see in the following chapter that, in order for the linked graph theorem to hold, it is necessary that the situation depicted in fig. 7.21, and similar situations, be true. In other words, the linked graph theorem holds only if violations of the Pauli principle are ignored.

To summarize: violations of the Pauli principle can be ignored in summing the contributions from all graphs, since graphs which violate the Pauli principle always occur in pairs which cancel each other. In applications of the linked graph theorem violations of the Pauli principle *must* be ignored.

CHAPTER 8

THE LINKED GRAPH THEOREM

8.1. Statement and General Discussion of the Theorem

We have shown in ch. 6 that, if the energy of the ground state is

$$E = E_0 + \Delta E, \tag{8.1}$$

where E_0 is the energy of the unperturbed ground state, then

$$\Delta E = \lim_{\alpha\to 0} \mathrm{i}\left[\frac{\partial}{\partial t}\log\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle\right]_{t=0}. \tag{8.2}$$

Here

$$\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle = 1 + \sum_{n=1}^{\infty} A_n, \tag{8.3}$$

and A_n is defined in eq. (6.27). In the present chapter we shall discuss an important theorem which greatly simplifies the evaluation of ΔE. This is known as the *linked graph theorem*†, and may be stated as follows:

$$\log\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle = \langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle_{\mathrm{L}}, \tag{8.4}$$

the subscript L on the right-hand side denoting that only the contributions from linked graphs are to be included in the evaluation of that expression. Thus, only linked graphs contribute to the energy of the ground state.

Before proving this theorem, let us see what it implies. As shown in eq. (6.28), we have

$$\log\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle = \sum A_n - \tfrac{1}{2}\Big(\sum A_n\Big)^2 + \tfrac{1}{3}\Big(\sum A_n\Big)^3 - \ldots, \tag{8.5}$$

---

† It is also known as the *linked cluster theorem.*

where A_n includes the contributions from *all* graphs of order n, both linked and unlinked. Now, it should be clear from the work of §7.6 that the terms of $(\Sigma A_n)^2$, $(\Sigma A_n)^3, \ldots$, can all be represented by unlinked graphs, which implies that, if the linked graph theorem is true, all the terms after the first on the right-hand side of (8.5) must simply cancel the contributions of all unlinked graphs to the first term, ΣA_n.

We can, in fact, easily see that this is so for the lower-order contributions. First, in order to reduce the number of graphs we need consider†, we shall neglect all graphs which do not conserve momentum. The meaning of this will be made clear in the following chapter, where it will be shown that terms represented by such graphs do not occur, for example, in the energy of a free-electron gas at the absolute zero of temperature. Essentially, the implication is that graphs containing loops of the types shown in fig. 8.1 must be excluded, so that only exchange graphs of the type of fig. 7.13(b) are to be included in first order, and only graphs of the types shown in figs. 7.14(a), (b), (c), (d) and (x) are to be included in second order.

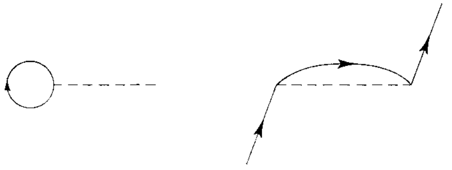


Fig. 8.1.

Also, again to reduce the number of graphs, we shall, in this section only, represent by a single graph, without arrowheads, all graphs of a given type *as well as those of the types obtained by interchanging the vertices at any of the interaction lines.* For example, in the second-order case we represent all graphs of the types shown in figs. 7.14(a) and (b) by just fig. 7.14(a) without the arrowheads, and all graphs of the types shown in figs. 7.14(c) and (d) by just 7.14(c) without the arrowheads. As we have seen in §7.5.2, the contributions from the types in each pair are the same, so that the inclusion of both in a single graph merely requires the introduction of a factor of 2. The same thing must be done in the case of unlinked third-order graphs with these second-order graphs as unlinked

---

† The theorem remains true if all graphs are taken into account, as we shall prove, but the demonstration becomes more cumbersome.

parts, so that the factor of 2 may be neglected when cancelling unlinked graphs, as we shall do below. We shall not consider terms of order higher than the third, and so may restrict ourselves to the graphs shown in fig. 8.2.

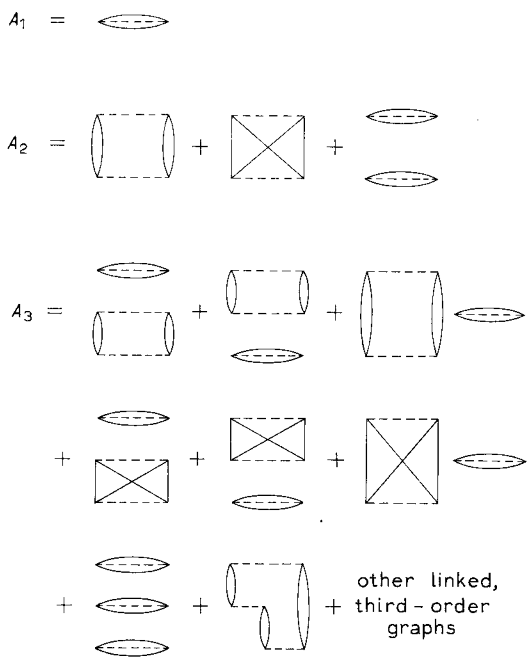


Fig. 8.2.

Expanding the right-hand side of eq. (8.5), we have

$$\log \langle \Phi_0 | U_\alpha(t, -\infty) | \Phi_0 \rangle = A_1 + A_2 + A_3 + \ldots$$
$$- \tfrac{1}{2} A_1^2 - A_1 A_2 - \ldots + \tfrac{1}{3} A_1^3 + \ldots, \qquad (8.6)$$

the terms shown explicitly being *all* those of the first, second and third orders. We shall now make use of the results expressed in figs. 7.21 and 7.23, together with the similar and easily verified results expressed in fig. 8.3, to represent graphically the final three terms of (8.6). These terms, shown in fig. 8.4, clearly cancel the unlinked graphs of $A_1 + A_2 + A_3$, shown in fig. 8.2.

This verifies the linked graph theorem up to and including terms of the third order. It is a useful demonstration in that it clearly shows how the unlinked graphs cancel, but it does not lead directly to a general proof of

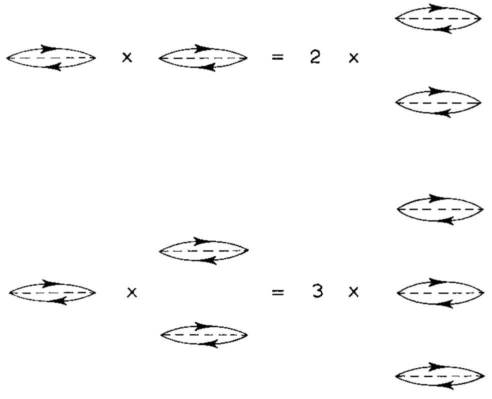


Fig. 8.3.

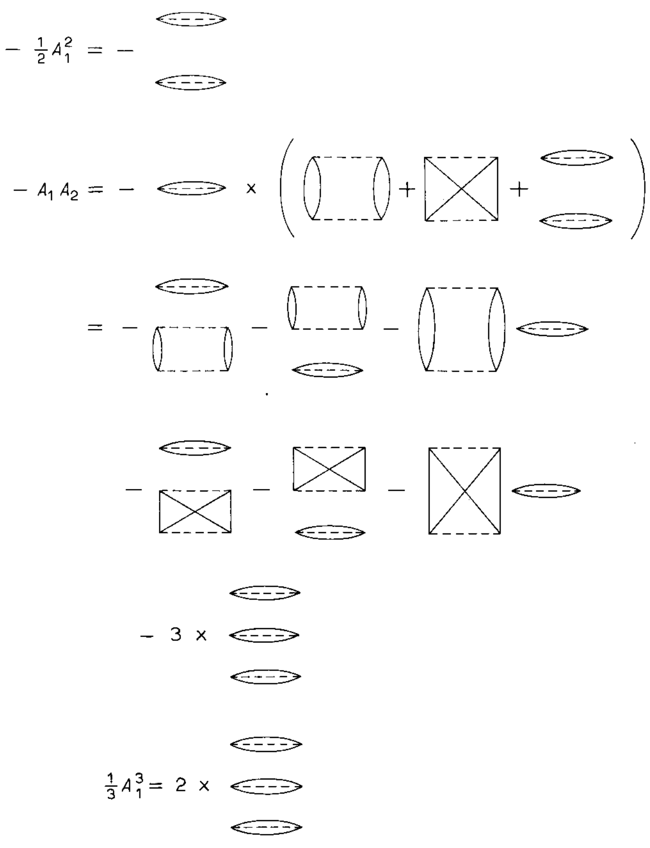


Fig. 8.4.

the theorem. However, once the result discussed in the following section is firmly established, we shall see that the general proof of the theorem is quite simple.

8.2. Unlinked Graphs with the Same Unlinked Parts

We shall now prove in general a result already demonstrated in certain particular cases – for example, figs. 7.21 and 7.23. From eq. (7.22), we have

$$A_n = (-\tfrac{1}{2}\mathrm{i})^n \sum_{ijkl} \dots \sum_{rsuw} \langle ij|v|kl\rangle \dots \langle rs|v|uw\rangle$$
$$\times (-1)^{\lambda+\mu} \int_{-\infty}^{t} \exp[(\mathrm{i}\Delta_1+\alpha)t_1]\,\mathrm{d}t_1 \int_{-\infty}^{t_1} \exp[(\mathrm{i}\Delta_2+\alpha)t_2]\,\mathrm{d}t_2 \dots$$
$$\times \int_{-\infty}^{t_{n-1}} \exp[(\mathrm{i}\Delta_n+\alpha)t_n]\,\mathrm{d}t_n, \qquad (8.7)$$

where $\lambda+\mu$, explained in §7.7, may vary from term to term of the sum. We first wish to consider the contributions to A_n from unlinked graphs which are constructed from the same two unlinked parts, with different *relative* orderings of their interactions, the ordering of the interactions in each unlinked part being kept fixed. Let us suppose that the orders of the two unlinked graphs are m and $n-m$, respectively. We shall show that the contribution from *all* such graphs to A_n is just the product of the contribution from the first unlinked part to A_m and the contribution of the second unlinked part to A_{n-m}.

It is clear that the matrix elements $\langle ij|v|kl\rangle$, etc., and the factor $(-1)^{\lambda+\mu}$ are the same† for all such graphs, as well as for the product of the contributions from the unlinked parts. It is only necessary, therefore, to consider the time factor in (8.7), which may be written

$$\int_{t>t_1>t_2>\dots>t_n} \dots \int \exp[(\mathrm{i}\Delta_1+\alpha)t_1] \dots \exp[(\mathrm{i}\Delta_n+\alpha)t_n]\,\mathrm{d}t_1 \dots \mathrm{d}t_n, \qquad (8.8)$$

the integration being over all the values of the variables subject to the given condition.

Let us denote the times of the interactions in the two unlinked parts by

---

† The result we wish to establish is, in fact, true only if *all* such graphs are included, with the value of the vacuum–vacuum matrix element being taken to be $(-1)^{\lambda+\mu}$ for each. As explained in §7.8, this means that possible violations of the Pauli principle must be ignored.

$t_1, t_2, \ldots, t_m$ and $t'_1, t'_2, \ldots, t'_{n-m}$, respectively. Also, let us denote by $\Delta_1, \Delta_2, \ldots, \Delta_m$ and $\Delta'_1, \Delta'_2, \ldots, \Delta'_{n-m}$ the corresponding energy factors for the two parts. We wish to find the contribution to A_n from all† such graphs, in which the t_i and t'_i are arranged in every possible order subject to the conditions

$$t > t_1 > t_2 > \ldots t_m \tag{8.9}$$

and

$$t > t'_1 > t'_2 > \ldots t'_{n-m}. \tag{8.10}$$

Since all *relative* orderings of the interactions in the two unlinked parts are included, the sum of the time factors like (8.8) for all such graphs is just

$$\int_{t>t_1>t_2>\ldots>t_m} \!\!\ldots\! \int \int_{t>t'_1>t'_2>\ldots>t'_{n-m}} \!\!\ldots\! \int \exp\left[(\mathrm{i}\Delta_1+\alpha)t_1\right] \ldots \exp\left[(\mathrm{i}\Delta'_1+\alpha)t'_1\right] \ldots$$

$$\times \, \mathrm{d}t_1 \ldots \mathrm{d}t'_{n-m}$$

$$= \int_{t>t_1>t_2>\ldots>t_m} \!\!\ldots\! \int \exp\left[(\mathrm{i}\Delta_1+\alpha)t_1\right] \ldots \exp\left[(\mathrm{i}\Delta_m+\alpha)t_m\right] \mathrm{d}t_1 \ldots \mathrm{d}t_m$$

$$\times \int_{t>t'_1>t'_2>\ldots>t'_{n-m}} \!\!\ldots\! \int \exp\left[(\mathrm{i}\Delta'_1+\alpha)t'_1\right] \ldots \exp\left[(\mathrm{i}\Delta'_{n-m}+\alpha)t'_{n-m}\right] \mathrm{d}t'_1 \ldots \mathrm{d}t'_{n-m}, \tag{8.11}$$

which is the product of the time factors for the two unlinked parts taken separately.

This completes the proof of our first point. Before proceeding, however, it may be found helpful to have a concrete example of the foregoing. Let us, therefore, reconsider the graphs in fig. 7.18. We shall denote the times of the interactions in the second-order unlinked part by t_1, t_2, with $t_1 > t_2$, and the time of the interaction in the other part by t'_1. The sum of the time factors for the three graphs is, then,

$$\iiint_{t>t_1>t_2>t'_1} (\quad)\, \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t'_1 + \iiint_{t>t'_1>t_1>t_2} (\quad)\, \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t'_1$$

$$+ \iiint_{t>t_1>t'_1>t_2} (\quad)\, \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t'_1, \tag{8.12}$$

---

† There are $n!$ ways of ordering n different objects, and, in a fraction $1/m!(n-m)!$ of these, m chosen objects have a particular order among themselves and the remaining $n-m$ objects also have a particular order among themselves. The number of such graphs is thus $n!/m!(n-m)!$. For example, in fig. 7.18, the number is $3!/2!1! = 3$.

the integrand ( ) being

$$\exp\left[(\mathrm{i}\Delta_1+\alpha)t_1\right] \exp\left[(\mathrm{i}\Delta_2+\alpha)t_2\right] \exp\left[(\mathrm{i}\Delta'_1+\alpha)t'_1\right]$$

in each case. Since all ranges of integration of t'_1 relative to t_1 and t_2 are included, (8.12) can be written as

$$\int_{t>t'_1} \exp\left[(\mathrm{i}\Delta'_1+\alpha)t'_1\right] \mathrm{d}t'_1 \iint_{t>t_1>t_2} \exp\left[(\mathrm{i}\Delta_1+\alpha)t_1\right] \exp\left[(\mathrm{i}\Delta_2+\alpha)t_2\right] \mathrm{d}t_1 \mathrm{d}t_2, \tag{8.13}$$

which is the product of the time factors of the two unlinked parts taken separately. This was proved in §7.6 by direct calculation of the time factors.

The generalization of the result proved above is immediate. Thus the sum of the contributions to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ from all graphs constructed from the same set of unlinked parts, by arranging the relative ordering of the interactions in all possible ways (retaining the ordering of the interactions in each unlinked part separately, however), is equal to the product of the contributions from the unlinked parts considered separately.

When we come to sum over i, j, k, l, etc., that is, to find the contribution from all graphs of the same types, constructed as described above from the same unlinked parts, a little care must be exercised. If all the unlinked parts are of different types, then the total contribution is just the product of the total contributions from all graphs of the types of the unlinked parts (for example, the result expressed in fig. 7.21), but this is not so if several of the unlinked parts are of the same type. We have already had an example of this, which led to the result expressed in fig. 7.23, but the point is so important that a further example might be worth while. Let us consider the six graphs shown in fig. 8.5, which are constructed by arranging in all possible ways the relative ordering of the interactions in two unlinked parts of the same type. The sum of the contributions to A_4 from these six graphs is just the product of the contributions to A_2 from the two unlinked parts. However, if we sum over i, j, k, l, m, n, p, q, to obtain the total contribution from all graphs of these types, we see that the same graphs are included *twice*, since the graphs (a) and (d) are of the same type, as are the pair (b), (e) and the pair (c), (f). The square of the total contribution to A_2 from all graphs of the type of the unlinked parts has therefore to be *divided by 2* in order to obtain the total contribution to A_4 from all *different* graphs of the types shown in fig. 8.5.

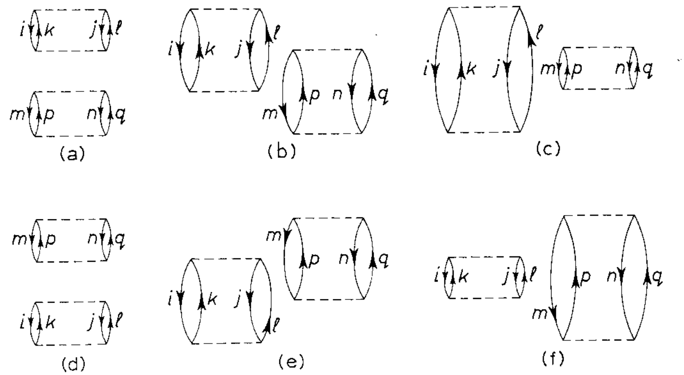


Fig. 8.5.

It follows that, for graphs constructed from N unlinked parts of the same type, the product of the contributions from the unlinked parts must be divided by $N!$, since this is the number of ways of ordering the unlinked parts without changing the type of the graph. The final generalization of this result, to graphs constructed from several different sets of unlinked parts, the graphs of each set being of the same type, is obvious, but we shall leave it until the next section, where we shall introduce a convenient notation.

8.3. Proof of the Linked Graph Theorem

Let us denote the *types* of all *linked* graphs by† $G_1, G_2, G_3, \ldots$, as shown in fig. 8.6, for example. The most general type of graph will thus consist of $n_1+n_2+\ldots$ unlinked parts, n_1 of type G_1, n_2 of type G_2, and so on, where $n_1, n_2, \ldots = 0, 1, 2, \ldots$.

We shall denote by $\langle\Phi_0|U_{G_i}|\Phi_0\rangle$ the contribution to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ from *all* graphs of the type G_i. In general, a large number of different types of graphs can be constructed from n_1 unlinked parts of type G_1, n_2 of type G_2, etc., by arranging the relative ordering of the interactions in all possible ways. It follows directly from the work of the previous section that the total contribution to $\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle$ from all graphs of all these different types is

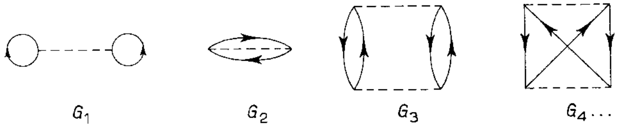


Fig. 8.6.

$$\frac{1}{n_1!n_2!\ldots}\langle\Phi_0|U_{G_1}|\Phi_0\rangle^{n_1}\langle\Phi_0|U_{G_2}|\Phi_0\rangle^{n_2}\ldots,$$

since there are $n_1!$ ways of arranging the unlinked parts of type G_1 among themselves, $n_2!$ ways of arranging those of type G_2, etc., without changing the type of the whole graph.

The total contribution from all graphs of all types is obtained by summing over all values of $n_1, n_2, \ldots$. Thus

$$\begin{aligned}\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle\\ &= \sum_{n_1,n_2,\ldots}\frac{1}{n_1!n_2!\ldots}\langle\Phi_0|U_{G_1}|\Phi_0\rangle^{n_1}\langle\Phi_0|U_{G_2}|\Phi_0\rangle^{n_2}\ldots\\ &= \left[\sum_{n_1=0}^{\infty}\frac{1}{n_1!}\langle\Phi_0|U_{G_1}|\Phi_0\rangle^{n_1}\right]\left[\sum_{n_2=0}^{\infty}\frac{1}{n_2!}\langle\Phi_0|U_{G_2}|\Phi_0\rangle^{n_2}\right]\ldots\\ &= \exp\left[\langle\Phi_0|U_{G_1}|\Phi_0\rangle+\langle\Phi_0|U_{G_2}|\Phi_0\rangle+\ldots\right]\\ &= \exp\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle_{\mathrm{L}},\end{aligned}\tag{8.14}$$

where L indicates the contribution from all linked graphs only. The linked graph theorem, in the form of eq. (8.4), follows immediately.

8.4. The Energy of the Ground State

We shall now consider in greater detail the energy of the ground state of the perturbed system. From eqs. (8.2), (8.3) and (8.4), we have

$$\begin{aligned}\Delta E &= \lim_{\alpha\to 0}\mathrm{i}\left[\frac{\partial}{\partial t}\langle\Phi_0|U_\alpha(t,-\infty)|\Phi_0\rangle_{\mathrm{L}}\right]_{t=0}\\ &= \lim_{\alpha\to 0}\mathrm{i}\sum_{n=1}^{\infty}\left[\left(\frac{\partial A_n}{\partial t}\right)_{\mathrm{L}}\right]_{t=0}.\end{aligned}\tag{8.15}$$

---

† The subscripts here have no significance apart from the enumeration of the different types. In particular, they do not indicate the *order* of a graph. There are many different types of linked graphs of any given order – for example, figs. 7.14(a)–(t) show the *twenty* different types of second-order linked graphs. We are here using the definition of *type* given in §7.6, and we are not restricting the graphs in any way, as we did in §8.1.

Also, from eq. (7.22),

$$A_n = (-\tfrac{1}{2}\mathrm{i})^n \sum_{ijkl} \dots \sum_{rsuw} \langle ij|v|kl\rangle \dots \langle rs|v|uw\rangle$$
$$\times \frac{(-1)^{\lambda+\mu} \exp\{[\mathrm{i}(\Delta_1+\Delta_2+\dots+\Delta_n)+n\alpha]t\}}{[\mathrm{i}(\Delta_1+\Delta_2+\dots+\Delta_n)+n\alpha][\mathrm{i}(\Delta_2+\Delta_3+\dots+\Delta_n)+(n-1)\alpha]\dots[\mathrm{i}\Delta_n+\alpha]}, \tag{8.16}$$

where

$$\Delta_n = \epsilon_r+\epsilon_s-\epsilon_u-\epsilon_w,$$

etc., and we have substituted $(-1)^{\lambda+\mu}$ for the vacuum–vacuum matrix element, as explained in §7.7, so that

$$\lim_{\alpha\to 0} \mathrm{i}\left(\frac{\partial A_n}{\partial t}\right)_{t=0}$$
$$= -(-\tfrac{1}{2})^n \sum_{ijkl} \dots \sum_{rsuw} \frac{\langle ij|v|kl\rangle \dots \langle rs|v|uw\rangle (-1)^{\lambda+\mu}}{(\Delta_2+\Delta_3+\dots+\Delta_n)(\Delta_3+\dots+\Delta_n)\dots\Delta_n}. \tag{8.17}$$

The value of $(-1)^{\lambda+\mu}$ depends, of course, upon the type of graph representing a given term, and $\Delta_2, \Delta_3, \dots$, will vary from term to term.

If we write

$$\Delta E = \sum_{n=1}^{\infty} \Delta E^{(n)}, \tag{8.18}$$

as in eq. (6.29), then

$$\Delta E^{(n)} = -(-\tfrac{1}{2})^n \underset{\text{(linked graphs)}}{\sum_{ijkl} \dots \sum_{rsuw}} \frac{\langle ij|v|kl\rangle \dots \langle rs|v|uw\rangle (-1)^{\lambda+\mu}}{(\Delta_2+\Delta_3+\dots+\Delta_n)(\Delta_3+\dots+\Delta_n)\dots\Delta_n}. \tag{8.19}$$

This is the *n-th-order energy correction*. The sum is over all terms represented by n-th-order linked graphs. It should be noted that, although the time factors and the creation and destruction operators have vanished from (8.19), the graph representing any term can be constructed from the matrix elements $\langle ij|v|kl\rangle, \dots$, taken in the correct order. Conversely, if an n-th-order linked graph is drawn at random, the corresponding term of the sum in (8.19) can immediately be written down. Time dependence, which was used as scaffolding in order to arrive at the desired result (8.19), can now be forgotten.

8.5. Intermediate States

The graphs we have been considering, which contribute to the energy of the ground state, are often called *vacuum–vacuum graphs*, because they represent a series of interactions in which the initial state Φ_0 ultimately becomes Φ_0 again (with a possible change of sign), and Φ_0 takes the place of the vacuum state in this work. We note that each unlinked part of an unlinked graph is itself a vacuum–vacuum graph.

The states which occur during the time sequence of interactions represented by a graph are called *intermediate states*. Consider, for example, the vacuum–vacuum matrix element

$$\langle\Phi_0|c_i^\dagger c_j^\dagger c_l c_k \dots c_r^\dagger c_s^\dagger c_w c_u|\Phi_0\rangle$$

occurring in the general term of A_n, eq. (7.22). We may write

$$c_r^\dagger c_s^\dagger c_w c_u \Phi_0 = \Phi_1, \tag{8.20}$$

where Φ_1 is the first of the intermediate states. The last interaction, represented by the operators $c_i^\dagger c_j^\dagger c_l c_k$, must transform the final intermediate state back to Φ_0 (or $-\Phi_0$) for a vacuum–vacuum graph.

Two things about intermediate states should be emphasized. First, intermediate states are not real physical states of the system, but are sometimes called 'virtual' states. This becomes obvious when we remember that, since we are neglecting certain possible violations of the Pauli principle, *intermediate states need not obey the Pauli principle.* They *may* be states of the unperturbed system, but the unperturbed system itself is generally a mathematical fiction. Secondly, for a linked graph, Φ_0 *can never be an intermediate state.* This is because, if Φ_0 were an intermediate state at some point, the graph would separate at this point into two unlinked parts, each being a vacuum–vacuum graph.

The energy factors $\Delta_2+\Delta_3+\dots+\Delta_n$, etc., may be expressed in terms of the energies of the intermediate states. Thus, if E_1 is the energy of the intermediate state Φ_1, given in eq. (8.20), then†

$$\Delta_n = \epsilon_r+\epsilon_s-\epsilon_w-\epsilon_u = E_1-E_0. \tag{8.21}$$

Similarly, if E_2 is the energy of the intermediate state formed by operating upon Φ_1 with the next four operators, then

$$\Delta_{n-1} = E_2-E_1. \tag{8.22}$$

---

† It should be noted that E_1 is still given by eq. (8.21) even when $w=u$, for example, so that Φ_1 is not a state of the unperturbed system.

Thus, the second last factor in the denominator of the general term of $\Delta E^{(n)}$, in eq. (8.19), becomes

$$\Delta_{n-1}+\Delta_n = E_2-E_0. \tag{8.23}$$

Every factor can be treated similarly, and, as there are $n-1$ intermediate states, the denominator becomes

$$(E_{n-1}-E_0)(E_{n-2}-E_0)\dots(E_1-E_0).$$

It must be remembered, of course, that the energies of the intermediate states will vary from term to term – that is to say, E_1, for example, will not be the same for every term of $\Delta E^{(n)}$.

In a notation which requires a little care in its interpretation, therefore, eq. (8.19) may be written

$$\Delta E^{(n)} = (\tfrac{1}{2})^n \sum_{\substack{ijkl\\ \text{(linked graphs)}}}\dots\sum_{rsuw} \frac{\langle ij|v|kl\rangle\dots\langle rs|v|uw\rangle(-1)^{\lambda+\mu}}{(E_0-E_{n-1})(E_0-E_{n-2})\dots(E_0-E_1)}. \tag{8.24}$$

Since only linked graphs are allowed, Φ_0 cannot be an intermediate state. Also, we assumed in ch. 6 that the ground state Φ_0 is non-degenerate, so that no intermediate state can have energy E_0. It follows that no factor in the denominator of (8.24) can ever be zero.

8.6. Alternative Formulation of the Perturbation Series

We shall now derive the perturbation series for the ground-state energy correction in a different form, which occurs frequently in the literature and which can be compared with the Rayleigh–Schrödinger series in eq. (1.76).

From eq. (8.15) we have

$$\Delta E = \lim_{\alpha\to 0} \mathrm{i}\sum_{n=1}^{\infty}\left[\left(\frac{\partial A_n}{\partial t}\right)_{\mathrm{L}}\right]_{t=0}, \tag{8.25}$$

and eqs. (6.24) and (6.27) give

$$A_n = \langle\Phi_0|U_n|\Phi_0\rangle, \tag{8.26}$$

where

$$U_n = (-\mathrm{i})^n\int_{-\infty}^{t} H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1\int_{-\infty}^{t_1} H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2\dots\int_{-\infty}^{t_{n-1}} H_{\mathrm{I}}(t_n)\,\mathrm{d}t_n. \tag{8.27}$$

Thus

$$\left(\frac{\partial A_n}{\partial t}\right)_{\mathrm{L}} = \left\langle\Phi_0\middle|(-\mathrm{i})^n H_{\mathrm{I}}(t)\int_{-\infty}^{t} H_{\mathrm{I}}(t_2)\,\mathrm{d}t_2\int_{-\infty}^{t_2} H_{\mathrm{I}}(t_3)\,\mathrm{d}t_3\dots \right.$$
$$\left.\times\dots\int_{-\infty}^{t_{n-1}} H_{\mathrm{I}}(t_n)\,\mathrm{d}t_n\middle|\Phi_0\right\rangle_{\mathrm{L}}. \tag{8.28}$$

Now, from eq. (6.8),

$$H_{\mathrm{I}}(t) = \exp(\mathrm{i}H_0t)H'\exp(-\mathrm{i}H_0t)\exp(\alpha t), \tag{8.29}$$

so that

$$\int_{-\infty}^{t_{n-1}} H_{\mathrm{I}}(t_n)\,\mathrm{d}t_n\Phi_0$$
$$= \int_{-\infty}^{t_{n-1}} \exp(\mathrm{i}H_0t_n)H'\exp(-\mathrm{i}H_0t_n)\exp(\alpha t_n)\,\mathrm{d}t_n\Phi_0$$
$$= \int_{-\infty}^{t_{n-1}} \exp(\mathrm{i}H_0t_n)H'\exp(-\mathrm{i}E_0t_n)\exp(\alpha t_n)\,\mathrm{d}t_n\Phi_0$$
$$= \int_{-\infty}^{t_{n-1}} \exp[\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)t_n]\,\mathrm{d}t_nH'\Phi_0$$
$$= \frac{\exp[\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)t_{n-1}]}{\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)}H'\Phi_0, \tag{8.30}$$

since

$$H\Phi_0 = E_0\Phi_0, \tag{8.31}$$

and H', which does not commute with H_0, is time-independent.

It follows that

$$\int_{-\infty}^{t_{n-2}} H_{\mathrm{I}}(t_{n-1})\,\mathrm{d}t_{n-1}\int_{-\infty}^{t_{n-1}} H_{\mathrm{I}}(t_n)\,\mathrm{d}t_n\Phi_0$$
$$= \int_{-\infty}^{t_{n-2}} \exp(\mathrm{i}H_0t_{n-1})H'\exp(-\mathrm{i}H_0t_{n-1})\exp(\alpha t_{n-1})$$
$$\times\frac{\exp[\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)t_{n-1}]}{\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)}\mathrm{d}t_{n-1}H'\Phi_0$$
$$= \int_{-\infty}^{t_{n-2}} \exp[\mathrm{i}(H_0-E_0-2\mathrm{i}\alpha)t_{n-1}]H'\frac{1}{\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)}\mathrm{d}t_{n-1}H'\Phi_0$$
$$= \frac{\exp[\mathrm{i}(H_0-E_0-2\mathrm{i}\alpha)t_{n-2}]}{\mathrm{i}(H_0-E_0-2\mathrm{i}\alpha)}H'\frac{1}{\mathrm{i}(H_0-E_0-\mathrm{i}\alpha)}H'\Phi_0. \tag{8.32}$$

Continuing this process, we finally obtain, from eq. (8.28),

$$\begin{aligned}\left(\frac{\partial A_n}{\partial t}\right)_{\rm L} &= \Big\langle\Phi_0\Big|(-{\rm i})^n \exp[{\rm i}(H_0-E_0-n{\rm i}\alpha)t]H'\\ &\quad\times\frac{1}{{\rm i}[H_0-E_0-(n-1){\rm i}\alpha]}H'\dots H'\frac{1}{{\rm i}(H_0-E_0-{\rm i}\alpha)}H'\Big|\Phi_0\Big\rangle_{\rm L}\\ &= \Big\langle\Phi_0\Big|(-{\rm i})\exp[{\rm i}(H_0-E_0-n{\rm i}\alpha)t]H'\\ &\quad\times\frac{1}{E_0-H_0+(n-1){\rm i}\alpha}H'\dots H'\frac{1}{E_0-H_0+{\rm i}\alpha}H'\Big|\Phi_0\Big\rangle_{\rm L}.\end{aligned}\tag{8.33}$$

Thus, from (8.25),

$$\begin{aligned}\Delta E &= \lim_{\alpha\to 0}\sum_{n=1}^{\infty}\Big\langle\Phi_0\Big|H'\frac{1}{E_0-H_0+(n-1){\rm i}\alpha}H'\dots H'\frac{1}{E_0-H_0+{\rm i}\alpha}H'\Big|\Phi_0\Big\rangle_{\rm L}\\ &= \sum_{n=1}^{\infty}\Big\langle\Phi_0\Big|H'\Big(\frac{1}{E_0-H_0}H'\Big)^{n-1}\Big|\Phi_0\Big\rangle_{\rm L}.\end{aligned}\tag{8.34}$$

It is permissible to put $\alpha = 0$ immediately because, as was pointed out in the previous section, Φ_0 cannot occur as an intermediate state in a linked graph. Also, we have assumed that Φ_0 is non-degenerate, so that all intermediate states have energies greater than E_0, which means that $E_0 - H_0$ will never operate upon an intermediate state to give zero.

Eq. (8.34) gives the linked graph, or linked cluster, expansion as it is usually encountered in the literature. It is similar to the Rayleigh–Schrödinger perturbation series in eq. (1.76), which may be written in the form

$$\Delta E = \sum_{n=1}^{\infty}\Big\langle\Phi_0\Big|H'\Big[\frac{R}{E_0-H_0}(E_0-E+H')\Big]^{n-1}\Big|\Phi_0\Big\rangle.\tag{8.35}$$

Clearly, the projection operator R, which excludes the Φ_0 component of any function it operates upon, has the same effect as the restriction to linked graphs only in (8.34). This will become more obvious if we compare the first- and second-order terms of (8.34) and (8.35).

The first-order correction ($n = 1$) given by eq. (8.34) is

$$\Delta E^{(1)} = \langle\Phi_0|H'|\Phi_0\rangle_{\rm L},\tag{8.36}$$

which agrees with that from (8.35), given in eq. (1.78), since *all* first-order graphs are linked, so that the L is superfluous.

The second-order correction given by eq. (8.34) is

$$\Delta E^{(2)} = \Big\langle\Phi_0\Big|H'\frac{1}{E_0-H_0}H'\Big|\Phi_0\Big\rangle_{\rm L},\tag{8.37}$$

which is to be compared with that from (8.35), given in eq. (1.81), namely,

$$\Delta E^{(2)} = \Big\langle\Phi_0\Big|H'\frac{R}{E_0-H_0}H'\Big|\Phi_0\Big\rangle.\tag{8.38}$$

The restriction to linked graphs only in (8.37) removes the Φ_0 component from $H'\Phi_0$, just as the operator R does in (8.38), since it disallows Φ_0 as an intermediate state. More explicitly, we have, from eq. (2.68),

$$H'\Phi_0 = \tfrac{1}{2}\sum_{ijkl}\langle ij|v|kl\rangle c_i^\dagger c_j^\dagger c_l c_k\Phi_0,\tag{8.39}$$

and the restriction to linked graphs only means that we must omit every term for which

$$c_i^\dagger c_j^\dagger c_l c_k\Phi_0 = \pm\Phi_0\tag{8.40}$$

(i.e. those terms for which $i,j = k,l$), since these would give rise to unlinked graphs of the types shown in fig. 7.14(u), (v), (w) and (x). Comparison of higher-order corrections becomes more complicated, but the principle remains the same.

CHAPTER 9

THE CORRELATION ENERGY OF A FREE-ELECTRON GAS

9.1. The First-order Perturbation Energy

We have already considered, in ch. 3, the first- and second-order perturbation corrections to the energy of the ground state of a free-electron gas, treating the whole Coulomb interaction as a perturbation. We shall now, as a prelude to our discussion of the correlation energy, show how to obtain these terms using the graphical method described in the previous chapter.

The interaction term H' was found to be (eq. (3.59))

$$\begin{aligned} H' &= \tfrac{1}{2}\sum_{\substack{\mathbf{k},\mathbf{p},\mathbf{q}\\ \sigma,\sigma'}} \langle \mathbf{p}+\mathbf{k},\sigma; \mathbf{q}-\mathbf{k},\sigma'|v|\mathbf{p}\sigma,\mathbf{q}\sigma'\rangle\, c^{\dagger}_{\mathbf{p}+\mathbf{k},\sigma} c^{\dagger}_{\mathbf{q}-\mathbf{k},\sigma'} c_{\mathbf{q}\sigma'} c_{\mathbf{p}\sigma} \\ &= \sum_{\substack{\mathbf{k},\mathbf{p},\mathbf{q}\\ \sigma,\sigma'}} \frac{2\pi e^2}{\Omega k^2}\, c^{\dagger}_{\mathbf{p}+\mathbf{k},\sigma} c^{\dagger}_{\mathbf{q}-\mathbf{k},\sigma'} c_{\mathbf{q}\sigma'} c_{\mathbf{p}\sigma}, \end{aligned} \tag{9.1}$$

where $\mathbf{k} = \mathbf{0}$ is to be omitted from the sum.

According to eq. (8.19), the first-order term is given by

$$\Delta E^{(1)} = \tfrac{1}{2} \sum_{\substack{ijkl\\ \text{(linked graphs)}}} \langle ij|v|kl\rangle (-1)^{\lambda+\mu}, \tag{9.2}$$

and the first-order graphs are of the types shown in fig. 9.1. There are, of course, no unlinked first-order graphs.

Fig. 9.1.

The only non-zero matrix elements in (9.2) must be of the type shown in (9.1), namely,

$$\langle \mathbf{p}+\mathbf{k},\sigma; \mathbf{q}-\mathbf{k},\sigma'|v|\mathbf{p}\sigma,\mathbf{q}\sigma'\rangle = \frac{4\pi e^2}{\Omega k^2}. \tag{9.3}$$

Now, a graph of the type of fig. 9.1(a) would represent a term of (9.2) with $i = k, j = l$. According to (9.3) this would imply

$$\mathbf{p}+\mathbf{k} = \mathbf{p} \quad \text{and} \quad \mathbf{q}-\mathbf{k} = \mathbf{q},$$

which is impossible, since $\mathbf{k} \neq \mathbf{0}$. Graphs of this type, therefore, do not contribute to $\Delta E^{(1)}$.

A graph of the type of fig. 9.1(b) represents a term of (9.2) with $i = l$, $j = k$, and this implies

$$\mathbf{p}+\mathbf{k} = \mathbf{q}, \quad \sigma = \sigma', \qquad \text{or} \qquad \mathbf{k} = \mathbf{q}-\mathbf{p}, \quad \sigma = \sigma'. \tag{9.4}$$

There is a single closed loop and two hole lines (c_k, c_l must create holes or they will operate upon Φ_0 to give zero), so that $\lambda + \mu = 3$.

We thus find

$$\Delta E^{(1)} = -\frac{4\pi e^2}{\Omega} \sum_{\substack{\mathbf{p},\mathbf{q}\neq\mathbf{p}\\ (p,q<k_{\mathrm{F}})}} \frac{1}{|\mathbf{p}-\mathbf{q}|^2}, \tag{9.5}$$

in agreement with eq. (3.75). The factor $\frac{1}{2}$ in (9.2) has been cancelled by the sum over spins, which, since $\sigma = \sigma'$, merely introduces a factor of 2. Also $p, q < k_{\mathrm{F}}$ since the loops refer to holes, and $\mathbf{p} \neq \mathbf{q}$ because $\mathbf{k} \neq \mathbf{0}$.

$\Delta E^{(1)}$ is the exchange energy, or Hartree–Fock term in the ground-state energy, and is given as a function of r_s in eq. (3.77).

9.2. The Second-order Perturbation Energy

From eq. (8.19) we have

$$\Delta E^{(2)} = -\tfrac{1}{4} \sum_{\substack{ijkl\\ \text{(Linked graphs)}}} \sum_{rsuw} \langle ij|v|kl\rangle \langle rs|v|uw\rangle (-1)^{\lambda+\mu} \frac{1}{\Delta_2}, \tag{9.6}$$

where

$$\Delta_2 = \epsilon_r + \epsilon_s - \epsilon_u - \epsilon_w. \tag{9.7}$$

The second-order linked graphs are of the types shown in fig. 7.14, with the exception of (u), (v), (w) and (x). In fact, however, only (a), (b), (c) and (d) contribute to the sum in (9.6). All the other graphs give rise to matrix elements which, according to (9.3), must vanish, as we shall now show.

Let us consider fig. 7.14(e), which contains loops of the type shown in fig. 9.2(a). Such a loop implies a matrix element $\langle ij|v|kl\rangle$ in which $i = l$, but $j \neq k$ (otherwise the graph would be unlinked, as in fig. 7.14(x)). According to (9.3), this means that $\sigma = \sigma'$ and

$$\boldsymbol{p}+\boldsymbol{k} = \boldsymbol{q}, \qquad \boldsymbol{q}-\boldsymbol{k} \neq \boldsymbol{p}, \tag{9.8}$$

which is impossible. It follows that there is no contribution to $\Delta E^{(2)}$ from any graph containing an 'exchange' loop of this type.

Fig. 9.2.

Also, let us consider fig. 7.14(q), which contains two loops of the type shown in fig. 9.2(b). Such a loop implies a matrix element $\langle ij|v|kl\rangle$ in which $i = k$, but $j \neq l$ (otherwise the graph would be unlinked, as in fig. 7.14(u)). According to (9.3), this means that

$$\boldsymbol{p}+\boldsymbol{k} = \boldsymbol{p}, \tag{9.9}$$

which is impossible, since $\boldsymbol{k} \neq \boldsymbol{0}$. It follows that there is no contribution to $\Delta E^{(2)}$ from any graph containing a loop of this type either.

It should be noticed that the foregoing arguments depend only upon the situation at a single interaction line and not upon the order of the graph (provided it is not of first order). We deduce, therefore, that *there is no contribution to the ground-state energy of a free-electron gas from graphs of any order (above the first) containing loops of the types shown in fig. 9.2.* First-order graphs containing loops of the type of 9.2(a) do contribute, of course, as was shown above.

The disallowed graphs may be said to be so because they contain interactions which *do not conserve momentum.* The matrix element (9.3) describes an interaction in which two electrons in states $\boldsymbol{p},\sigma$ and $\boldsymbol{q},\sigma'$ are excited to states $\boldsymbol{p}+\boldsymbol{k},\sigma$ and $\boldsymbol{q}-\boldsymbol{k},\sigma'$, respectively. Owing to this interaction the first electron has acquired additional momentum† $\boldsymbol{k}$, and the second electron has lost momentum $\boldsymbol{k}$. The total momentum remains unchanged, and $\boldsymbol{k}$ is referred to as the *momentum transfer* due to the interaction. Since the only non-zero matrix elements are of the type (9.3), it follows that any interaction which does not conserve momentum is automatically excluded.

It will be useful to note also that, according to eq. (9.3), *the value of every non-vanishing matrix element is* $4\pi e^2/\Omega$ *divided by the square of the appropriate momentum transfer.*

The only permitted second-order graphs are thus of the types shown in fig. 9.3 (and the two which are obtained from them by exchanging the vertices at *one* interaction, which we shall account for by multiplying the sum in eq. (9.6) by 2). The lines have been labelled in the '$\boldsymbol{k},\sigma$' notation and the momentum transfer has been indicated at each interaction. For both graphs we have

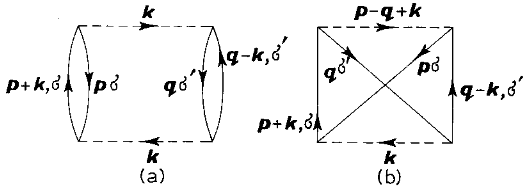


Fig. 9.3.

$$\begin{aligned}\Delta_2 &= \epsilon_{p+k}+\epsilon_{q-k}-\epsilon_p-\epsilon_q \\ &= [(\boldsymbol{p}+\boldsymbol{k})^2+(\boldsymbol{q}-\boldsymbol{k})^2-p^2-q^2]\frac{1}{2m} \\ &= \boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})\frac{1}{m}.\end{aligned} \tag{9.10}$$

Let us first consider the graph in fig. 9.3(a). There are two closed loops and two hole lines, so that $\lambda+\mu = 4$. The momentum transfer is $\boldsymbol{k}$ at both interactions, so that both matrix elements have value $(4\pi e^2/\Omega k^2)$. Hence the contribution to (9.6) from all graphs of this type is

$$\begin{aligned}\Delta E_{\mathrm{a}}^{(2)} &= -\tfrac{1}{2}\sum_{\substack{pqk\\ \sigma\sigma'}} \frac{m(4\pi e^2/\Omega k^2)^2}{\boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})} \\ &= -\frac{32\pi^2 e^4 m}{\Omega^2}\sum_{pqk}\frac{1}{k^4\boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})},\end{aligned} \tag{9.11}$$

---

† We retain the system of units in which $\hbar = 1$. In ordinary units the momentum of a free electron is $\hbar\boldsymbol{k}$.

where

$$p, q < k_F \quad \text{and} \quad |p+k|, |q-k| > k_F.$$

The sum over spins merely introduces a factor 4 – parallel spin ($\sigma = \sigma'$) and antiparallel spin ($\sigma \neq \sigma'$) interactions contribute equally to $\Delta E_a^{(2)}$.

For the graph in fig. 9.3(b) the matrix element at the lower interaction is the same as before, but that at the upper interaction is

$$\langle q\sigma', p\sigma | v | p+k, \sigma; q-k, \sigma' \rangle, \tag{9.12}$$

which vanishes unless $\sigma = \sigma'$, when its value is

$$\frac{4\pi e^2}{\Omega (p-q+k)^2}, \tag{9.13}$$

since the momentum transfer is $p-q+k$ (e.g. $q-k \to p$). This graph has one closed loop and two hole lines, so that $\lambda + \mu = 3$. Hence the contribution to (9.6) from all graphs of this type is

$$\Delta E_b^{(2)} = \frac{16\pi^2 e^4 m}{\Omega^2} \sum_{pqk} \frac{1}{k^2 (p-q+k)^2 k \cdot (p-q+k)}, \tag{9.14}$$

again with

$$p, q < k_F \quad \text{and} \quad |p+k|, |q-k| > k_F.$$

The sum over spins has only introduced a factor 2 here, since we must have $\sigma = \sigma'$ – that is to say, only parallel spin interactions contribute to $\Delta E_b^{(2)}$.

These results may be compared with those of §3.6. We see that $\Delta E_a^{(2)}$ is just twice $\Delta E_{\uparrow\downarrow}^{(2)}$, given in eq. (3.88) (with $\hbar = 1$). This is because, as remarked above, $\Delta E_a^{(2)}$ contains equal contributions from parallel spin and antiparallel spin interactions. We may thus write

$$\Delta E^{(2)} = \Delta E_a^{(2)} + \Delta E_b^{(2)}, \tag{9.15}$$

or

$$\Delta E^{(2)} = \Delta E_{\uparrow\downarrow}^{(2)} + \Delta E_{\uparrow\uparrow}^{(2)}, \tag{9.16}$$

where

$$\Delta E_{\uparrow\downarrow}^{(2)} = \tfrac{1}{2} \Delta E_a^{(2)} \tag{9.17}$$

and

$$\Delta E_{\uparrow\uparrow}^{(2)} = \tfrac{1}{2} \Delta E_a^{(2)} + \Delta E_b^{(2)}. \tag{9.18}$$

$\Delta E_{\uparrow\uparrow}^{(2)}$ is, of course, the contribution from parallel spin interactions, which was not explicitly calculated in §3.6.

If, as in §3.6, we replace the sums by integrals, measure energy in rydbergs and k, p and q in units of k_F, we find

$$\Delta E_a^{(2)} = -\frac{3N}{8\pi^5} \int \frac{dk}{k^4} \int dq \int \frac{dp}{k \cdot (p-q+k)} \text{ ry} \tag{9.19}$$

and

$$\Delta E_b^{(2)} = \frac{3N}{16\pi^5} \int \frac{dk}{k^2} \int dq \int \frac{dp}{(p-q+k)^2 k \cdot (p-q+k)} \text{ ry}, \tag{9.20}$$

the region of integration being defined by

$$p, q < 1 \quad \text{and} \quad |p+k|, |q-k| > 1$$

in both cases.

We note for future reference that, owing to the symmetrical distribution of the vectors p and q, we may equally well write $\Delta E_a^{(2)}$, for example, in the form

$$\Delta E_a^{(2)} = -\frac{3N}{8\pi^5} \int \frac{dk}{k^4} \int dp_1 \int \frac{dp_2}{k \cdot (p_1 + p_2 + k)} \text{ ry}, \tag{9.21}$$

with

$$p_1, p_2 < 1 \quad \text{and} \quad |p_1 + k|, |p_2 + k| > 1.$$

$\Delta E_a^{(2)}$ diverges logarithmically at $k = 0$, since it is just twice $\Delta E_{\uparrow\downarrow}^{(2)}$, which was shown to diverge in §3.6. On the other hand, $\Delta E_b^{(2)}$ does not diverge, since the integrand contains the factor k^{-2} instead of k^{-4}. It is difficult to evaluate, however, and Gell-Mann and Brueckner† quote the value

$$\Delta E_b^{(2)} = 0.046N \text{ ry}, \tag{9.22}$$

obtained by a numerical method, which we shall use in the following work.

Now, it is fairly obvious that the higher-order perturbation corrections will also diverge‡ at $k = 0$, since there will be one integral at least containing in its integrand the factor $4\pi e^2/\Omega^2 k^2$ for each interaction. Gell-Mann and Brueckner showed that, in the high-density limit at least, the perturbation series can be summed in such a way that the divergences cancel out to give an exact expression for the correlation energy. We shall now consider their method.

---

† Gell-Mann, M. and Brueckner, K., 1957, Phys. Rev. **106**, 364.

‡ There is never a divergence as $k \to \infty$, because it may easily be seen that, for very large k, the integrand of every integral appearing in $\Delta E^{(n)}$ becomes approximately k^{2-4n}.

9.3. The Correlation Energy of a Free-electron Gas at High Densities, according to Gell-Mann and Brueckner

9.3.1. *The perturbation series*

We shall first show that $\Delta E^{(n)}$ is proportional to r_s^{n-2}, where, as before,

$$\frac{4\pi}{3} r_s^3 = \frac{\Omega}{N}.$$

It has already been found that the first-order correction $\Delta E^{(1)}$, eq. (3.77), is proportional to r_s^{-1}, and the above calculation shows that $\Delta E^{(2)}$ is independent of r_s.

Now, in going from one perturbation correction to that of next higher order, there is an extra integration over $\boldsymbol{k}$-space and four extra linear factors containing $\boldsymbol{k}$ are introduced into the denominators of the integrands. When the integrals are made dimensionless, therefore, as in (9.19) and (9.20), an extra factor k_F^{-1} appears in the multiplying constant, and, from eq. (3.65),

$$k_F^{-1} = \left(\frac{\Omega}{3\pi^2 N}\right)^{1/3} = \alpha r_s, \tag{9.23}$$

where

$$\alpha = \left(\frac{4}{9\pi}\right)^{1/3}. \tag{9.24}$$

Since $\Delta E^{(1)}$ is proportional to r_s^{-1}, it thus follows that $\Delta E^{(n)}$ is proportional to r_s^{n-2}.

The perturbation series after $\Delta E^{(1)}$ is therefore a power series in r_s. The coefficient of each power of r_s will be a sum of several integrals, corresponding to different types of graphs, and some of these will have different orders of divergence. Extending the work of §3.6 in an obvious way, the situation may be represented schematically as follows:

$$\Delta E^{(2)} = C_1^{(2)} \int_0 \frac{dk}{k} + C_2^{(2)} \int_0 k\,dk, \tag{9.25}$$

$$\Delta E^{(3)} = \left(C_1^{(3)} \int_0 \frac{dk}{k^3} + C_2^{(3)} \int_0 \frac{dk}{k} + C_3^{(3)} \int_0 k\,dk\right) r_s, \tag{9.26}$$

$$\Delta E^{(4)} = \left(C_1^{(4)} \int_0 \frac{dk}{k^5} + C_2^{(4)} \int_0 \frac{dk}{k^3} + C_3^{(4)} \int_0 \frac{dk}{k} + C_4^{(4)} \int_0 k\,dk\right) r_s^2, \tag{9.27}$$

etc., where the $C_i^{(j)}$ are constants, independent of r_s, and we have only shown the integration which is most important at $k = 0$.

Thus

$$\begin{aligned} \Delta E^{(2)} + \Delta E^{(3)} + \Delta E^{(4)} + \ldots = {} & (C_2^{(2)} + C_3^{(3)} r_s + C_4^{(4)} r_s^2 + \ldots) \int_0 k\,dk \\ & + (C_1^{(2)} + C_2^{(3)} r_s + C_3^{(4)} r_s^2 + \ldots) \int_0 \frac{dk}{k} \\ & + (C_1^{(3)} r_s + C_2^{(4)} r_s^2 + C_3^{(5)} r_s^3 + \ldots) \int_0 \frac{dk}{k^3} \\ & + (C_1^{(4)} r_s^2 + C_2^{(5)} r_s^3 + C_3^{(6)} r_s^4 + \ldots) \int_0 \frac{dk}{k^5} \\ & + \ldots \end{aligned} \tag{9.28}$$

The divergence at $k = 0$ is due to the long range of the Coulomb interaction. If, as in ch. 4, the interaction were screened, so that the lower limit of the k-integration were $\beta (> 0)$, then at this limit we should have

$$\int_\beta \frac{dk}{k} \propto \log \beta, \tag{9.29}$$

$$\int_\beta \frac{dk}{k^3} \propto \beta^{-2}, \tag{9.30}$$

$$\int_\beta \frac{dk}{k^5} \propto \beta^{-4}, \tag{9.31}$$

and so on. If also, as in eq. (4.81), we take

$$\beta \propto r_s^{1/2}, \tag{9.32}$$

multiplication by the appropriate integral will make the leading term in each bracket of (9.28), after the second, independent of r_s. All the following terms will still contain positive powers of r_s, however. This suggests that *in the high density limit* ($r_s \to 0$) it is sufficient to sum only the most divergent integrals in each order of perturbation theory (i.e. those with coefficients $C_1^{(2)}$, $C_1^{(3)}$, $C_1^{(4)}, \ldots$). The second-order exchange term $\Delta E_b^{(2)}$ must also be included, since it is a constant, independent of r_s.

Now, in each order of perturbation theory, the most divergent integrals are those coming from graphs in which the same momentum transfer $\boldsymbol{k}$ occurs at *every* interaction, because this will give the highest power of k in the denominator of the integrand, as we have already seen in the second-order case. For these graphs, each pair of lines starting at a given vertex must end together at another vertex. They are called *ring graphs*, and some types of various orders are shown in fig. 9.4.

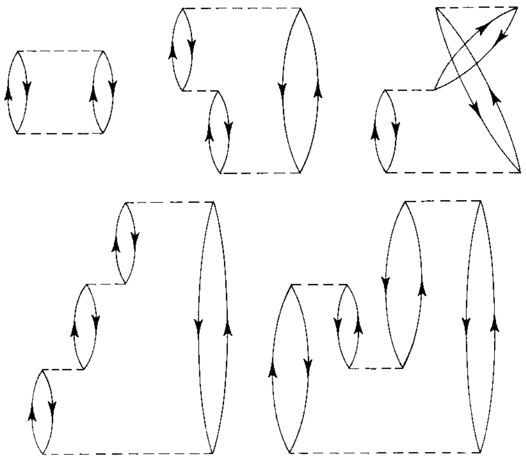


Fig. 9.4. Some types of ring graphs.

The initial simplification introduced by Gell-Mann and Brueckner (*loc. cit.*) was thus to take the total correlation energy† at high densities to be

$$NW = \Delta E^{(2)} + (\Delta E^{(3)} + \Delta E^{(4)} + \ldots)_{\mathrm{ring}}$$

$$= \Delta E_{\mathrm{b}}^{(2)} + (\Delta E_{\mathrm{a}}^{(2)} + \Delta E^{(3)} + \ldots)_{\mathrm{ring}}, \tag{9.33}$$

since $\Delta E_{\mathrm{a}}^{(2)}$ is due to ring graphs only. The success of the method depends upon the sum in brackets yielding a finite result, which we shall show to be so provided the integrals are summed before the final k-integration is performed.

We shall now obtain as an integral the contribution from ring graphs to $\Delta E^{(3)}$. There are eight types of third-order graphs. One is shown in fig. 9.5, the others being obtained from it by reversing the vertices at one or more interactions. The contribution from each of these graphs is the same, the value of the matrix element at each interaction being $4\pi e^2/\Omega k^2$, since the momentum transfer is $\boldsymbol{k}$.

---

† The correlation energy, W per electron, was defined in §4.4 as the exact ground-state energy minus the Hartree–Fock energy. In the present case, where the whole Coulomb interaction is treated as a perturbation, it is just the sum of the perturbation corrections of all orders after the first.

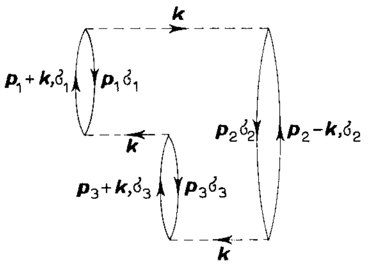


Fig. 9.5. A third-order ring graph.

The energy factors at the lower interactions are

$$\Delta_2 = [(\boldsymbol{p}_1+\boldsymbol{k})^2 + p_3^2 - (\boldsymbol{p}_3+\boldsymbol{k})^2 - p_1^2]\frac{1}{2m}$$

$$= \boldsymbol{k}\cdot(\boldsymbol{p}_1-\boldsymbol{p}_3)\frac{1}{m}, \tag{9.34}$$

$$\Delta_3 = \boldsymbol{k}\cdot(\boldsymbol{p}_3-\boldsymbol{p}_2+\boldsymbol{k})\frac{1}{m}, \tag{9.35}$$

so that

$$\Delta_2+\Delta_3 = \boldsymbol{k}\cdot(\boldsymbol{p}_1-\boldsymbol{p}_2+\boldsymbol{k})\frac{1}{m}. \tag{9.36}$$

The number of hole lines is always the same as the number of loops in a ring graph, so that

$$(-1)^{\lambda+\mu} = 1 \tag{9.37}$$

for all ring graphs. Thus, according to eq. (8.19), with $n=3$, the total contribution from all ring graphs to $\Delta E^{(3)}$ is

$$\Delta E_{\mathrm{ring}}^{(3)} = -8(-\tfrac{1}{2})^3 \sum_{\substack{\boldsymbol{p}_1\boldsymbol{p}_2\boldsymbol{p}_3\boldsymbol{k}\\ \sigma_1\sigma_2\sigma_3}} \frac{m^2(4\pi e^2/\Omega k^2)^3}{[\boldsymbol{k}\cdot(\boldsymbol{p}_1-\boldsymbol{p}_2+\boldsymbol{k})][\boldsymbol{k}\cdot(\boldsymbol{p}_3-\boldsymbol{p}_2+\boldsymbol{k})]}, \tag{9.38}$$

where

$$p_1, p_2, p_3 < k_{\mathrm{F}} \qquad \text{and} \qquad |\boldsymbol{p}_1+\boldsymbol{k}|, |\boldsymbol{p}_2-\boldsymbol{k}|, |\boldsymbol{p}_3+\boldsymbol{k}| > k_{\mathrm{F}}.$$

The sum over spins merely contributes a factor 8. Also, owing to the symmetrical distribution of the vector $\boldsymbol{p}_2$, we may change its sign, and (9.38) becomes

$$\Delta E_{\mathrm{ring}}^{(3)} = \sum_{\boldsymbol{p}_1\boldsymbol{p}_2\boldsymbol{p}_3\boldsymbol{k}} \frac{8m^2(4\pi e^2/\Omega)^3}{k^6[\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})][\boldsymbol{k}\cdot(\boldsymbol{p}_2+\boldsymbol{p}_3+\boldsymbol{k})]}, \tag{9.39}$$

where

$$p_1, p_2, p_3 < k_{\mathrm{F}} \qquad \text{and} \qquad |\boldsymbol{p}_1+\boldsymbol{k}|, |\boldsymbol{p}_2+\boldsymbol{k}|, |\boldsymbol{p}_3+\boldsymbol{k}| > k_{\mathrm{F}}.$$

Expressing this as an integral, with p_1, p_2, p_3 and k measured in units of k_F, energy in rydbergs and length in Bohr units, we find†

$$\Delta E^{(3)}_{\text{ring}} = \frac{3N}{4\pi^5}\left(\frac{\alpha r_\mathrm{s}}{\pi^2}\right) \int \frac{\mathrm{d}\boldsymbol{k}}{k^6} \int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2$$

$$\times \int \frac{\mathrm{d}\boldsymbol{p}_3}{[\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})][\boldsymbol{k}\cdot(\boldsymbol{p}_2+\boldsymbol{p}_3+\boldsymbol{k})]}\,\text{ry}, \tag{9.40}$$

the region of integration being given by

$$p_1, p_2, p_3 < 1 \qquad \text{and} \qquad |\boldsymbol{p}_1+\boldsymbol{k}|, |\boldsymbol{p}_2+\boldsymbol{k}|, |\boldsymbol{p}_3+\boldsymbol{k}| > 1.$$

We shall not consider in detail the contribution of ring graphs to $\Delta E^{(4)}$, but a brief mention may be useful.

There are three *fundamentally* different types of fourth-order ring graphs, which give different contributions to the energy. These are shown in fig. 9.6. Corresponding to each of these types there are fifteen others, obtained by reversing the vertices at one or more interactions, each of which gives the same contribution to the energy as the fundamental type.

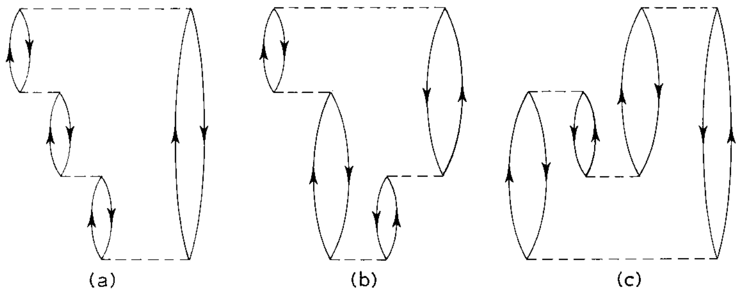

(a) (b) (c)

Fig. 9.6. Types of fourth-order ring graphs giving different contributions to the energy.

---

† Since, with $\hbar = 1$, we have 1 rydberg $= \frac{1}{2}me^4$, 1 Bohr unit $= 1/me^2$, the constant multiplying the integral is

$$8m^2\left(\frac{4\pi e^2}{\Omega}\right)^3\left(\frac{\Omega}{8\pi^3}\right)^4 k_\mathrm{F}^2\left(\frac{2}{me^4}\right) = \frac{3Nme^2}{4\pi^7}\left(\frac{\Omega}{3\pi^2 N}\right)k_\mathrm{F}^2 = \frac{3Nme^2}{4\pi^7}\frac{1}{k_\mathrm{F}}$$

$$= \frac{3Nme^2}{4\pi^7}\frac{\alpha r_\mathrm{s}}{me^2} = \frac{3N}{4\pi^5}\left(\frac{\alpha r_\mathrm{s}}{\pi^2}\right),$$

where r_s is measured in Bohr units.

Applying the same method as before, it is found† that

$$\Delta E^{(4)}_{\text{ring}} = -\frac{3N}{4\pi^5}\left(\frac{\alpha r_\mathrm{s}}{\pi^2}\right)^2 \int \frac{\mathrm{d}\boldsymbol{k}}{k^8} \int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3$$

$$\times \int \Bigg[\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_3+\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_4+\boldsymbol{k})}$$

$$+\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_3+\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_3+\boldsymbol{p}_4+\boldsymbol{k})}$$

$$+\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{p}_3+\boldsymbol{p}_4+2\boldsymbol{k})}\,\frac{1}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_3+\boldsymbol{k})}\Bigg]\mathrm{d}\boldsymbol{p}_4, \tag{9.41}$$

the region of integration being given by

$$p_1, p_2, p_3, p_4 < 1 \quad \text{and} \quad |\boldsymbol{p}_1+\boldsymbol{k}|, |\boldsymbol{p}_2+\boldsymbol{k}|, |\boldsymbol{p}_3+\boldsymbol{k}|, |\boldsymbol{p}_4+\boldsymbol{k}| > 1.$$

9.3.2. *Summing the contributions from ring graphs*

First, following Gell-Mann and Brueckner, we wish to show that

$$\Delta E^{(n)}_{\text{ring}} = (-1)^{n+1}\frac{3N}{8\pi^5}\left(\frac{\alpha r_\mathrm{s}}{\pi^2}\right)^{n-2} \int I_n k^{-2n}\,\mathrm{d}\boldsymbol{k}, \tag{9.42}$$

where

$$I_n = \frac{1}{n}\int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{\infty}\mathrm{d}t_2 \ldots \int_{-\infty}^{\infty} F_k(t_1)F_k(t_2)\ldots F_k(t_n)$$

$$\times\,\delta(t_1+t_2+\ldots+t_n)\,\mathrm{d}t_n \tag{9.43}$$

and

$$F_k(t) = \int \exp\left[-|t|\left(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}\right)\right]\mathrm{d}\boldsymbol{p}. \tag{9.44}$$

We shall not prove this rigorously, which would be tedious, but merely show that it gives correct results for $n = 2$ and $n = 3$.

Let us consider $E^{(2)}_\mathrm{a}$. We have‡

$$I_2 = \tfrac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}t_1 \int_{-\infty}^{\infty} F_k(t_1)F_k(t_2)\delta(t_1+t_2)\,\mathrm{d}t_2$$

$$= \tfrac{1}{2}\int_{-\infty}^{\infty} F_k(t_1)F_k(-t_1)\,\mathrm{d}t_1 = \int_0^{\infty} F_k(t)F_k(-t)\,\mathrm{d}t$$

---

† See Gell-Mann and Brueckner (*loc. cit.*).

‡ From the conditions $p_1, p_2 < 1$ and $|\boldsymbol{p}_1+\boldsymbol{k}|, |\boldsymbol{p}_2+\boldsymbol{k}| > 1$ it follows that $\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})$ is positive, because even in the most extreme case, when $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are both in the opposite sense to $\boldsymbol{k}$, $\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k}$ still has the same sense as $\boldsymbol{k}$.

$$= \int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int_0^\infty \exp\left[-t(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_1)\right] \exp\left[-t(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_2)\right] \mathrm{d}t$$

$$= \int \mathrm{d}\boldsymbol{p}_1 \int \left[-\frac{\exp\left[-t\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})\right]}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})}\right]_0^\infty \mathrm{d}\boldsymbol{p}_2$$

$$= \int \mathrm{d}\boldsymbol{p}_1 \int \frac{\mathrm{d}\boldsymbol{p}_2}{\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})}, \tag{9.45}$$

which, when substituted in (9.42), with $n = 2$, yields the expression (9.21) for $\Delta E_{\mathrm{a}}^{(2)}$.

Now, let us consider $\Delta E_{\mathrm{ring}}^{(3)}$, which is slightly more complicated. We have

$$I_3 = \tfrac{1}{3}\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \mathrm{d}t_2 \int_{-\infty}^{\infty} F_k(t_1)F_k(t_2)F_k(t_3)\delta(t_1+t_2+t_3)\,\mathrm{d}t_3$$

$$= \tfrac{1}{3}\int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} F_k(t_1)F_k(t_2)F_k(-t_1-t_2)\,\mathrm{d}t_2$$

$$= \tfrac{1}{3}\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3 \int_{-\infty}^{\infty} \mathrm{d}t_1 \int_{-\infty}^{\infty} \exp\left[-|t_1|(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_1)\right.$$
$$\left. -|t_2|(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_2) - |t_1+t_2|(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_3)\right]\mathrm{d}t_2. \tag{9.46}$$

To eliminate the moduli signs in the integrand we must divide the region of the t_1, t_2 integrations into six parts, as follows:

$$\text{(A)}\quad \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_2 \quad [|t_1| = t_1, \quad |t_2| = t_2, \quad |t_1+t_2| = t_1+t_2],$$

$$\text{(B)}\quad \int_0^\infty \mathrm{d}t_1 \int_{-t_1}^{0} \mathrm{d}t_2 \quad [|t_1| = t_1, \quad |t_2| = -t_2, \; |t_1+t_2| = t_1+t_2],$$

$$\text{(C)}\quad \int_0^\infty \mathrm{d}t_1 \int_{-\infty}^{-t_1} \mathrm{d}t_2 \; [|t_1| = t_1, \quad |t_2| = -t_2, \; |t_1+t_2| = -t_1-t_2],$$

$$\text{(D)}\quad \int_{-\infty}^{0} \mathrm{d}t_1 \int_{-\infty}^{0} \mathrm{d}t_2 \; [|t_1| = -t_1, \; |t_2| = -t_2, \; |t_1+t_2| = -t_1-t_2],$$

$$\text{(E)}\quad \int_{-\infty}^{0} \mathrm{d}t_1 \int_{0}^{t_1} \mathrm{d}t_2 \; [|t_1| = -t_1, \; |t_2| = t_2, \quad |t_1+t_2| = -t_1-t_2],$$

$$\text{(F)}\quad \int_{-\infty}^{0} \mathrm{d}t_1 \int_{t_1}^{\infty} \mathrm{d}t_2 \; [|t_1| = -t_1, \; |t_2| = t_2, \quad |t_1+t_2| = t_1+t_2].$$

It may be seen immediately, by changing the signs of t_1 and t_2, that (A) = (D), (B) = (E) and (C) = (F). In fact, if the variables are transformed linearly (with inversion of the order of integration where necessary), so that the limits of integration are 0 and ∞ for both t_1 and t_2, it will be found that the integral with respect to $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ of each of these integrals is the same (see Problem 9.2). We may thus write

$$I_3 = 2\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3 \int_0^\infty \mathrm{d}t_1 \int_0^\infty \exp\left[-t_1(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_1)\right.$$
$$\left. - t_2(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_2) - (t_1+t_2)(\tfrac{1}{2}k^2+\boldsymbol{k}\cdot\boldsymbol{p}_3)\right]\mathrm{d}t_2$$

$$= 2\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3 \int_0^\infty \exp\left[-t_1\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})\right]\mathrm{d}t_1$$
$$\times \int_0^\infty \exp\left[-t_2\boldsymbol{k}\cdot(\boldsymbol{p}_2+\boldsymbol{p}_3+\boldsymbol{k})\right]\mathrm{d}t_2$$

$$= 2\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \frac{\mathrm{d}\boldsymbol{p}_3}{[\boldsymbol{k}\cdot(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{k})][\boldsymbol{k}\cdot(\boldsymbol{p}_2+\boldsymbol{p}_3+\boldsymbol{k})]}, \tag{9.47}$$

which, when substituted in (9.42), with $n = 3$, yields the expression (9.40) for $\Delta E_{\mathrm{ring}}^{(3)}$.

We shall assume without further calculation that the agreement extends to all orders. Now, as shown in Appendix V, we may write

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(iux)\,\mathrm{d}u, \tag{9.48}$$

or, equally well, since $\mathrm{d}(ku) = k\,\mathrm{d}u$,

$$\delta(x) = \frac{k}{2\pi}\int_{-\infty}^{\infty} \exp(ikux)\,\mathrm{d}u. \tag{9.49}$$

We may thus get rid of the δ-function in I_n by writing

$$\delta(t_1+t_2+\ldots+t_n) = \frac{k}{2\pi}\int_{-\infty}^{\infty} \exp\left[iku(t_1+\ldots+t_n)\right]\mathrm{d}u. \tag{9.50}$$

This gives

$$I_n = \frac{k}{2\pi n}\int_{-\infty}^{\infty} \mathrm{d}t_1 \ldots \int_{-\infty}^{\infty} \mathrm{d}t_n \int_{-\infty}^{\infty} F_k(t_1)\ldots F_k(t_n)$$
$$\times \exp\left[iku(t_1+\ldots+t_n)\right]\mathrm{d}u$$

$$= \frac{k}{2\pi n}\int_{-\infty}^{\infty} [Q_k(u)]^n\,\mathrm{d}u, \tag{9.51}$$

where

$$Q_k(u) = \int_{-\infty}^{\infty} \exp(\mathrm{i}kut) F_k(t)\,\mathrm{d}t$$

$$= \int \mathrm{d}\boldsymbol{p} \int_{-\infty}^{\infty} \exp(\mathrm{i}kut) \exp\left[-|t|\left(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}\right)\right] \mathrm{d}t. \tag{9.52}$$

Substituting in (9.42), we obtain

$$\Delta E_{\text{ring}}^{(n)} = (-1)^{n+1} \frac{3N}{8\pi^5}\left(\frac{\alpha r_{\mathrm{s}}}{\pi^2}\right)^{n-2} \frac{1}{2\pi n} \int \frac{\mathrm{d}\boldsymbol{k}}{k^{2n-1}} \int_{-\infty}^{\infty} [Q_k(u)]^n \,\mathrm{d}u. \tag{9.53}$$

Eqs. (9.33), (9.22) and (9.53) now yield the following expression for the correlation energy per electron:

$$W - 0.046 = \frac{1}{N}\sum_{n=2}^{\infty} \Delta E_{\text{ring}}^{(n)}$$

$$= -\frac{3}{16\pi^6}\left(\frac{\pi^2}{\alpha r_{\mathrm{s}}}\right)^2 \int k\,\mathrm{d}k \int_{-\infty}^{\infty} \sum_{n=2}^{\infty} \frac{(-1)^n}{n}\left[\frac{Q_k(u)\alpha r_{\mathrm{s}}}{\pi^2 k^2}\right]^n \mathrm{d}u$$

$$= \frac{3}{16\pi^6}\left(\frac{\pi^2}{\alpha r_{\mathrm{s}}}\right)^2 \int k\,\mathrm{d}k \int_{-\infty}^{\infty} \left[\log\left\{1 + \frac{Q_k(u)\alpha r_{\mathrm{s}}}{\pi^2 k^2}\right\} - \frac{Q_k(u)\alpha r_{\mathrm{s}}}{\pi^2 k^2}\right] \mathrm{d}u. \tag{9.54}$$

We note that the series in the integrand converges only if

$$-1 < \frac{Q_k(u)\alpha r_{\mathrm{s}}}{\pi^2 k^2} \leqslant 1. \tag{9.55}$$

Now, the divergences occur at $k = 0$, so that it will be sufficient in evaluating $Q_k(u)$ to assume that $k \ll 1$ (the difficulty to which this leads in respect of the convergence of the series in (9.54) will be discussed later). As in §3.6, we write

$$\boldsymbol{k}\cdot\boldsymbol{p} = kpx, \tag{9.56}$$

and, since $\boldsymbol{k}$ and $\boldsymbol{p}$ obey the conditions $p < 1$, $|\boldsymbol{p}+\boldsymbol{k}| > 1$, we have, from eq. (3.95),

$$1 > p > 1 - kx, \tag{9.57}$$

which implies that

$$1 > x > 0. \tag{9.58}$$

Since $k^2 \approx 0$ and $p \approx 1$, we may, by the same argument used in obtaining eq. (3.99), write

$$Q_k(u) = \int \mathrm{d}\boldsymbol{p} \int_{-\infty}^{\infty} \exp(\mathrm{i}kut) \exp\left[-|t|\left(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}\right)\right] \mathrm{d}t$$

$$\approx \int \mathrm{d}\boldsymbol{p} \int_{-\infty}^{\infty} \exp(\mathrm{i}kut) \exp(-|t|kx)\,\mathrm{d}t$$

$$\approx 2\pi k \int_0^1 x\,\mathrm{d}x \int_{-\infty}^{\infty} \exp(\mathrm{i}kut) \exp(-|t|kx)\,\mathrm{d}t$$

$$= 2\pi \int_0^1 x\,\mathrm{d}x \int_{-\infty}^{\infty} \exp(\mathrm{i}su) \exp(-|s|x)\,\mathrm{d}s$$

$$= 2\pi \int_0^1 x\left\{\int_{-\infty}^{0} \exp[s(\mathrm{i}u + x)]\,\mathrm{d}s + \int_0^{\infty} \exp[s(\mathrm{i}u - x)]\,\mathrm{d}s\right\}\mathrm{d}x$$

$$= 4\pi R(u), \tag{9.59}$$

where

$$R(u) = \int_0^1 \frac{x^2}{x^2 + u^2}\,\mathrm{d}x = 1 - u\tan^{-1}\left(\frac{1}{u}\right), \tag{9.60}$$

independent of k, provided $k \ll 1$.

Since the most important contribution comes from very small values of k, Gell-Mann and Brueckner used the function given in (9.59) and arbitrarily cut off the k-integration in (9.54) at the convenient upper limit of unity – in other words, they *approximated* to $Q_k(u)$ as follows:

$$Q_k(u) = \begin{cases} 4\pi R(u), & 0 \leqslant k \leqslant 1, \\ 0, & k > 1. \end{cases} \tag{9.61}$$

To be precise, Gell-Mann and Brueckner used this approximation only for terms with $n > 2$ in the perturbation series. They used the *correct* expression (9.21) for $\Delta E_{\mathrm{a}}^{(2)}$, and the difference between this and the value given by (9.61) provides a small additional constant in the correlation energy.

Inverting the order of integration in (9.54), and using (9.61), we thus have

$$W - 0.046 = \frac{3}{16\pi^6}\left(\frac{\pi^2}{\alpha r_{\mathrm{s}}}\right)^2 \int_{-\infty}^{\infty} \mathrm{d}u \int_0^1 4\pi k^2\left[\log\left(1 + \frac{X}{k^2}\right) - \frac{X}{k^2}\right] \mathrm{d}k + \delta, \tag{9.62}$$

where

$$X = 4\alpha r_s R(u)/\pi, \tag{9.63}$$

and δ is the small constant resulting from the correction to $\Delta E_a^{(2)}$ mentioned above. X is independent of k, so that the k-integration can be performed without much difficulty. This yields

$$W - 0.046 - \delta = \frac{3}{16\pi^5}\left(\frac{\pi^2}{\alpha r_s}\right)^2 \int_{-\infty}^{\infty} [X^2 \log X + (1-X^2)\log(1+X) - X]\,\mathrm{d}u. \tag{9.64}$$

Now, X is proportional to r_s, so that, as $r_s \to 0$,

$$(1-X^2)\log(1+X) - X \approx \tfrac{1}{2}X^2, \tag{9.65}$$

all higher powers of X, when multiplied by the factor outside the integral sign, tending to zero with r_s. We may therefore write†

$$\begin{aligned} W - 0.046 - \delta &= \frac{3}{16\pi^5}\left(\frac{\pi^2}{\alpha r_s}\right)^2 \int_{-\infty}^{\infty} X^2(\log X - \tfrac{1}{2})\,\mathrm{d}u + \mathrm{O}(r_s) \\ &= \frac{3}{\pi^3}\int_{-\infty}^{\infty} R^2\left[\log\left(\frac{4\alpha r_s}{\pi}\right) + \log R - \tfrac{1}{2}\right]\mathrm{d}u + \mathrm{O}(r_s) \\ &= \frac{2}{\pi^2}(1-\log 2)\left[\log\left(\frac{4\alpha r_s}{\pi}\right) - \tfrac{1}{2}\right] \\ &\quad + \frac{3}{\pi^3}\int_{-\infty}^{\infty} R^2 \log R\,\mathrm{d}u + \mathrm{O}(r_s), \end{aligned} \tag{9.66}$$

where $\mathrm{O}(r_s) \to 0$ as $r_s \to 0$.

---

† The evaluation of $\int_{-\infty}^{\infty} R^2\,\mathrm{d}u$ may be done as follows, using (9.60):

$$\begin{aligned} \int_{-\infty}^{\infty} R^2\,\mathrm{d}u &= \int_{-\infty}^{\infty}\left[\int_0^1 \frac{x^2}{x^2+u^2}\mathrm{d}x\right]^2 \mathrm{d}u \\ &= \int_{-\infty}^{\infty}\mathrm{d}u \int_0^1 \frac{x^2\,\mathrm{d}x}{x^2+u^2}\int_0^1 \frac{y^2\,\mathrm{d}y}{y^2+u^2} \\ &= \int_0^1 x^2\,\mathrm{d}x \int_0^1 y^2\,\mathrm{d}y \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{(x^2+u^2)(y^2+u^2)} \\ &= \pi\int_0^1 x\,\mathrm{d}x \int_0^1 \frac{y\,\mathrm{d}y}{x+y} \\ &= \pi\int_0^1 [x - x^2\log(x+1) + x^2\log x]\,\mathrm{d}x \\ &= \frac{2\pi}{3}(1-\log 2). \end{aligned}$$

The immediate result we obtain from this, without further calculation, is that, for $r_s \ll 1$, the correlation energy per electron is given by

$$W = \frac{2}{\pi^2}(1-\log 2)\log r_s + C + \mathrm{O}(r_s), \tag{9.67}$$

where C is a constant. The remaining integral in (9.64) has to be evaluated numerically, as also has the constant δ. Using the values quoted by Gell-Mann and Brueckner, we obtain, finally,

$$W = 0.0622 \log r_s - 0.096 + \mathrm{O}(r_s)\ \text{ry}, \tag{9.68}$$

with r_s measured in Bohr units. Although, strictly speaking, only accurate for $r_s \ll 1$ Bohr unit, this formula yields, for $r_s = 1$ Bohr unit and $\mathrm{O}(r_s)$ being neglected, $W = -0.096$ ry, which is closer than might have been expected to the value -0.113 ry given by the Wigner formula, eq. (4.98).

9.4. Comparison with the Results of the Plasma Theory

According to the theory of plasma oscillations presented in ch. 4, the correlation energy per electron is, from eq. (4.97),

$$W = 0.0311 \log r_s - 0.115\ \text{ry}. \tag{9.69}$$

It must be emphasized, however, that the regions of validity of (9.68) and (9.69) are not the same. None the less, there is a discrepancy between the two logarithmic terms which, at first sight, appears surprising. The reason for it is simply the neglect of the *parallel spin* part of the short-range correlation energy in the plasma theory. This is an approximation, but, for the reason given in §4.4.2, it is expected to be a good one in the region of actual metallic densities. If the parallel spin part is explicitly calculated, it yields a further logarithmic term, the same as in (9.69), and a term which is approximately constant, but not the same as in (9.69). That this must be so is sufficiently obvious from the argument leading to eqs. (9.17) and (9.18), which there applies to the perturbation treatment of the whole Coulomb interaction, but applies equally well to the perturbation treatment of the short-range interaction. The antiparallel spin part (originally taken to be the whole of W_s, and given in eqs. (4.94) and (4.95)), is simply one-half the integral in (9.19) with the additional condition $k > \beta$, while the parallel spin part is the same thing plus the integral in (9.20), again subject to $k > \beta$.

Thus, according to the plasma theory also, we may express the correlation energy in the form

$$W = 0.0622 \log r_s + \text{constant}, \tag{9.70}$$

neglecting as before a small term proportional to r_s. The new constant can be evaluated directly, but this involves the evaluation of a further complicated integral by numerical methods. We shall, therefore, simply estimate the constant by equating (9.69) and (9.70) at $r_s = 4$ (i.e. we assume that (9.69) gives a reasonably accurate value of W for $r_s = 4$). In this way, we obtain, for the correlation energy per electron as given by the plasma theory,

$$W = 0.0622 \log r_s - 0.158 \text{ ry}. \tag{9.71}$$

Ignoring for a moment the different values of the constants in (9.68) and (9.71), we note that, as $r_s \to 0$, the only term of importance is the logarithmic term, which is the same in both expressions. This implies that *summing over the ring graphs only*, as done by Gell-Mann and Brueckner, has precisely the same effect as the *random phase approximation* introduced by Bohm and Pines, and, furthermore, that the latter is exact at very high densities.

The values of the constants in (9.68) and (9.71) are very different. For the electron densities occurring in actual metals, corresponding roughly to $2 < r_s < 6$ Bohr units, the values given by the Gell-Mann and Brueckner formula (9.68), on the assumption that the term $O(r_s)$ is negligible, are much too small, whereas those given by the Bohm and Pines formula (9.71) are of the magnitude required by cohesive energy calculations. However, little is known† about the term $O(r_s)$, except that it tends to zero with r_s, and the basic approximation of Gell-Mann and Brueckner is only valid for very small r_s, so that speculation about why the formula fails for actual metallic densities is rather pointless. A term proportional to r_s has also been omitted from the Bohm and Pines formula (9.71), but this is certainly small at actual metallic densities. We conclude that, although the Gell-Mann and Brueckner formula is undoubtedly correct at high densities (possibly up to about $r_s = 1$ Bohr unit), the Bohm and Pines formula (9.71) (or (9.69)) is, perhaps fortuitously, more correct at actual metallic densities. Neither formula is correct at low densities, when the kinetic energy of the electrons becomes negligible compared

---

† See Du Bois, D. F., 1959, Ann. Phys. **7**, 174, for a discussion of some contributions to $O(r_s)$.

with their potential energy and the electrons tend to form a stable lattice (see WM, p. 271).

A difficulty with the Gell-Mann and Brueckner treatment, which we indicated previously, is the convergence of the series in (9.54). According to the condition (9.55), the series will not converge for very small values of k, which we have seen to be the most important values. None the less, Gell-Mann and Brueckner *assumed* that the logarithmic representation of the series could be used for *all* values of k, with the result obtained above. To this extent their method is not completely rigorous. However, their results have been verified by other methods – for example, the Sawada method†, in which the Hamiltonian is modified by the exclusion of all terms which do not give rise to ring graphs. We may therefore take the results of Gell-Mann and Brueckner to be correct at high densities.

---

† Sawada, K., 1957, Phys. Rev. **106**, 372. Sawada, K., Brueckner, K., Fukuda, N. and Brout, R., 1957, Phys. Rev. **108**, 507.

CHAPTER 10

GREEN FUNCTIONS AND THE ONE-ELECTRON SCHRÖDINGER EQUATION

10.1. The Time-independent Schrödinger Equation

The use of Green functions is now widespread in solid state theory and the theory of many-particle systems in general. In the present chapter, as a preliminary to an account of the application of Green functions to many-electron systems, we consider the simpler problem of the solution of the one-electron Schrödinger equation by Green functions.

Let us consider a single electron moving in an electrostatic field with potential energy $V(\boldsymbol{r})$. The time-independent Schrödinger equation is then (with $\hbar = 1$)

$$-\frac{1}{2m}\nabla^2\psi(\boldsymbol{r}) + V(\boldsymbol{r})\psi(\boldsymbol{r}) = \epsilon\psi(\boldsymbol{r}), \tag{10.1}$$

which we shall write in the form

$$(H_0-\epsilon)\psi(\boldsymbol{r}) = -V(\boldsymbol{r})\psi(\boldsymbol{r}), \tag{10.2}$$

where

$$H_0 = -\frac{1}{2m}\nabla^2. \tag{10.3}$$

Since we shall be mainly concerned with applications to the theory of metals, let us apply periodic boundary conditions (see WM, p. 161) over a large cube, say, of side L and volume $\Omega = L^3$.

According to the method described in Appendix IV, a solution of eq. (10.2) is given by the integral equation

$$\psi(\boldsymbol{r}) = \int G(\boldsymbol{r},\boldsymbol{r}';\epsilon)V(\boldsymbol{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \tag{10.4}$$

where $G(\boldsymbol{r},\boldsymbol{r}';\epsilon)$ is the *Green function* for the problem and satisfies the equation

$$(H_0-\epsilon)G(\boldsymbol{r},\boldsymbol{r}';\epsilon) = -\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{10.5}$$

subject to the same periodic boundary conditions as $\psi(\boldsymbol{r})$. The integral in eq. (10.4) is taken throughout the volume Ω.

That (10.4) is indeed a solution of eq. (10.2) can be verified immediately by substitution and the use of eq. (10.5). We have

$$\begin{aligned}(H_0-\epsilon)\psi(\boldsymbol{r}) &= (H_0-\epsilon)\int G(\boldsymbol{r},\boldsymbol{r}';\epsilon)V(\boldsymbol{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'\\ &= \int (H_0-\epsilon)G(\boldsymbol{r},\boldsymbol{r}';\epsilon)V(\boldsymbol{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'\\ &= -\int \delta(\boldsymbol{r}-\boldsymbol{r}')V(\boldsymbol{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'\\ &= -V(\boldsymbol{r})\psi(\boldsymbol{r}).\end{aligned} \tag{10.6}$$

Suppose that $\psi_k(\boldsymbol{r})$ is an eigenfunction, and ϵ_k the corresponding eigenvalue, of eq. (10.2) with the right-hand side put equal to zero, that is,

$$(H_0-\epsilon_k)\psi_k(\boldsymbol{r}) = 0, \tag{10.7}$$

subject to the same periodic boundary conditions as before. According to the elementary theory of an electron gas described in §3.4, the normalized eigenfunctions are given by

$$\psi_k(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega}}\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}), \tag{10.8}$$

with eigenvalues

$$\epsilon_k = k^2/2m, \tag{10.9}$$

the wave vector $\boldsymbol{k}$ being

$$\boldsymbol{k} = \frac{2\pi}{L}(n_1\boldsymbol{e}_1 + n_2\boldsymbol{e}_2 + n_3\boldsymbol{e}_3), \tag{10.10}$$

with $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ unit vectors along the cube edges, and n_1, n_2, n_3 integers, positive, negative or zero.

In terms of the $\psi_k(\boldsymbol{r})$ the Green function is

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon) = \sum_k \frac{\psi_k(\boldsymbol{r})\psi_k^*(\boldsymbol{r}')}{\epsilon-\epsilon_k} \tag{10.11}$$

(cf. eq. (IV.10), Appendix IV). This again may be verified by substitution in eq. (10.5), since

$$(H_0-\epsilon)G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=\sum_{k}\frac{\psi_k^*(\boldsymbol{r}')}{\epsilon-\epsilon_k}(H_0-\epsilon)\psi_k(\boldsymbol{r})$$

$$=-\sum_{k}\psi_k^*(\boldsymbol{r}')\psi_k(\boldsymbol{r})=-\delta(\boldsymbol{r}-\boldsymbol{r}'), \qquad (10.12)$$

by the closure or completeness property of the eigenfunctions (see eq. (1.15)).

Using eq. (10.8), we find

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=\frac{1}{\Omega}\sum_{k}\frac{\exp\left[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')\right]}{\epsilon-\epsilon_k}, \qquad (10.13)$$

and, assuming that the eigenvalues form a quasi-continuum, we may substitute an integral for the sum, according to the usual prescription

$$\sum_{k} f(\boldsymbol{k}) \rightarrow \frac{\Omega}{8\pi^3}\int f(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}, \qquad (10.14)$$

so that

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=\frac{1}{8\pi^3}\int\frac{\exp\left[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')\right]}{\epsilon-\epsilon_k}\mathrm{d}\boldsymbol{k}, \qquad (10.15)$$

the integral being taken throughout the whole of ***k***-space.

We note that the integrand depends upon ***r*** and ***r***′ only through ***r***−***r***′, so that we may write

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=G(\boldsymbol{r}-\boldsymbol{r}';\epsilon). \qquad (10.16)$$

For simplicity, therefore, we may consider the function $G(\boldsymbol{r};\epsilon)$, which satisfies the equation

$$(H_0-\epsilon)G(\boldsymbol{r};\epsilon)=-\delta(\boldsymbol{r}). \qquad (10.17)$$

If θ is the angle between ***k*** and ***r***, we have

$$G(\boldsymbol{r};\epsilon)=\frac{1}{8\pi^3}\int\frac{\exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right)}{\epsilon-\epsilon_k}\mathrm{d}\boldsymbol{k}$$

$$=\frac{1}{4\pi^2}\int_0^\infty k^2\,\mathrm{d}k\int_{-1}^{1}\frac{\exp\left(\mathrm{i}kr\cos\theta\right)}{\epsilon-\epsilon_k}\mathrm{d}(\cos\theta)$$

$$=\frac{1}{2\pi^2 r}\int_0^\infty\frac{\sin kr}{\epsilon-\epsilon_k}k\,\mathrm{d}k$$

$$=-\frac{1}{2\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\int_0^\infty\frac{\cos kr}{\epsilon-\epsilon_k}\mathrm{d}k$$

$$=-\frac{1}{4\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\int_{-\infty}^\infty\frac{\exp\left(\mathrm{i}kr\right)}{\epsilon-\epsilon_k}\mathrm{d}k, \qquad (10.18)$$

since ϵ_k is an even function. If we write $\epsilon=k'^2/2m$, this becomes

$$G(r;\epsilon)=-\frac{m}{2\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\int_{-\infty}^\infty\frac{\exp\left(\mathrm{i}kr\right)}{k'^2-k^2}\mathrm{d}k. \qquad (10.19)$$

The integral, unfortunately, does not exist, since the integrand has infinities at $k=\pm k'$, at which the integral does not converge. This difficulty is overcome by using, instead of the integral

$$\int_{-\infty}^\infty\frac{\exp\left(\mathrm{i}kr\right)}{\epsilon-\epsilon_k}\mathrm{d}k, \qquad (10.20)$$

the integral

$$\int_{-\infty}^\infty\frac{\exp\left(\mathrm{i}kr\right)}{\epsilon-\epsilon_k+\mathrm{i}\eta}\mathrm{d}k \qquad (10.21)$$

where η is a small *positive* constant. This integral does exist, since the integrand has no infinities, and it tends to (10.20) in the limit as $\eta\rightarrow 0$. The procedure, then, is to evaluate the function

$$G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)=\frac{1}{8\pi^3}\int\frac{\exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right)}{\epsilon-\epsilon_k+\mathrm{i}\eta}\mathrm{d}\boldsymbol{k}, \qquad (10.22)$$

which satisfies the equation

$$(H_0-\epsilon-\mathrm{i}\eta)G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)=-\delta(\boldsymbol{r}), \qquad (10.23)$$

and then take the limit as $\eta\rightarrow 0$.

The integral (10.21), which is

$$2m\int_{-\infty}^\infty\frac{\exp\left(\mathrm{i}kr\right)}{k'^2-k^2+2\mathrm{i}\eta' k'}\mathrm{d}k, \qquad (10.24)$$

where $\eta' = m\eta/k'$, with $k' > 0$, is most easily evaluated by means of a contour integration, k being treated as a complex variable. We may take the contour to be that part of the real k-axis from $-R$ to R, closed by a semicircle of radius R in the upper half-plane, and let $R \to \infty$. Now we have

$$k'^2 - k^2 + 2\mathrm{i}\eta' k' \approx (k' - k + \mathrm{i}\eta')(k' + k + \mathrm{i}\eta'),$$

neglecting $-\eta'^2$, since η' is small, so that the integrand of (10.24) has simple poles at $k = k' + \mathrm{i}\eta'$ and $k = -k' - \mathrm{i}\eta'$, of which only the former lies inside the chosen contour, as shown in fig. 10.1. The integrand is, in fact,

$$\frac{1}{2}\left(\frac{1}{k + k' + \mathrm{i}\eta'} - \frac{1}{k - k' - \mathrm{i}\eta'}\right)\frac{\exp(\mathrm{i}kr)}{k' + \mathrm{i}\eta'}, \tag{10.25}$$

and the residue (see Appendix VI) at the pole $k = k' + \mathrm{i}\eta'$ is

$$-\frac{\exp[\mathrm{i}(k' + \mathrm{i}\eta')r]}{2(k' + \mathrm{i}\eta')}. \tag{10.26}$$

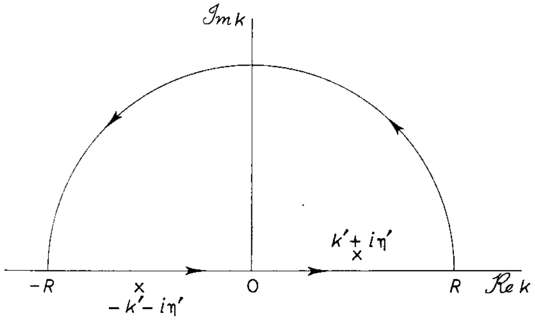


Fig. 10.1.

The contribution to the contour integral from the semicircle is zero when $R \to \infty$, as may be seen by setting $k = R\mathrm{e}^{\mathrm{i}\theta}$, so that the value of the integral in (10.24) is just $2\pi\mathrm{i}$ times the residue (10.26). Hence, from eq. (10.19),

$$\begin{aligned} G(r; \epsilon + \mathrm{i}\eta) &= -\frac{m}{2\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\int_{-\infty}^{\infty} \frac{\exp(\mathrm{i}kr)}{k'^2 - k^2 + 2\mathrm{i}\eta' k'}\mathrm{d}k \\ &= \frac{m}{2\pi^2 r}\frac{\mathrm{d}}{\mathrm{d}r}\left\{\frac{\pi\mathrm{i}\exp[\mathrm{i}(k' + \mathrm{i}\eta')r]}{k' + \mathrm{i}\eta'}\right\} \\ &= -\frac{m}{2\pi}\frac{\exp[\mathrm{i}(k' + \mathrm{i}\eta')r]}{r}. \end{aligned} \tag{10.27}$$

Taking the limit as $\eta \to 0$ (i.e. as $\eta' \to 0$), we have

$$G(r; \epsilon) = -\frac{m \exp(\mathrm{i}k'r)}{2\pi r}, \tag{10.28}$$

and

$$G(\boldsymbol{r}, \boldsymbol{r}'; \epsilon) = -\frac{m \exp(\mathrm{i}k'|\boldsymbol{r} - \boldsymbol{r}'|)}{2\pi|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{10.29}$$

The justification of the foregoing procedure, which is used in all applications of Green functions, lies in the fact that the function $G(r; \epsilon)$ given in eq. (10.28) is indeed a solution of eq. (10.17), as we shall now verify. For $r \neq 0$, we have

$$\nabla^2 G = \frac{\mathrm{d}^2 G}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}G}{\mathrm{d}r} = -k'^2 G, \tag{10.30}$$

so that

$$(H_0 - \epsilon)G = -\frac{1}{2m}(\nabla^2 + k'^2)G = 0. \tag{10.31}$$

The delta function on the right-hand side of eq. (10.17) is due to the infinity in $G(r; \epsilon)$ at $r = 0$. Let us integrate the left-hand side of the equation through a sphere, centred at $\boldsymbol{r} = \boldsymbol{0}$, with very small radius ρ. Within this sphere we may take $\exp(\mathrm{i}k'r) = 1$, so that, using the divergence theorem of Gauss,

$$\begin{aligned} \int (H_0 - \epsilon)G\,\mathrm{d}\boldsymbol{r} &= \frac{1}{4\pi}\int (\nabla^2 + k'^2)\frac{1}{r}\mathrm{d}\boldsymbol{r} \\ &= \frac{1}{4\pi}\int \nabla\left(\frac{1}{r}\right)\cdot \mathrm{d}\boldsymbol{S} + \frac{k'^2}{4\pi}\int \frac{\mathrm{d}\boldsymbol{r}}{r} \\ &= -\frac{1}{4\pi}\int \frac{1}{r^3}\boldsymbol{r}\cdot \mathrm{d}\boldsymbol{S} + k'^2\int_0^{\rho} r\,\mathrm{d}r \\ &= -\frac{1}{4\pi\rho^2}(4\pi\rho^2) + \tfrac{1}{2}k'^2\rho^2 \\ &= -1 = -\int \delta(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}, \end{aligned} \tag{10.32}$$

in the limit as $\rho \to 0$.

Although this justifies the method of obtaining the function $G(\boldsymbol{r}, \boldsymbol{r}'; \epsilon)$ in eq. (10.29), a small difficulty still remains, for this function does not satisfy the same periodic boundary conditions as the function given in eq. (10.13). This is due to the replacement of the sum in eq. (10.13) by an integral, which is only strictly true in the limit as the cube edge $L \to \infty$. Clearly, if L is very large, the functions $G(\boldsymbol{r},\boldsymbol{r}';\epsilon)$ given in eq. (10.29) will approximately satisfy the periodic boundary conditions (provided $\boldsymbol{r}'$ is not too close to a cube face) owing to the fact that it will practically vanish over the faces of the cube.

The function $G(\boldsymbol{r};\epsilon)$ in eq. (10.28) is not the only possible Green function for the problem. It is sometimes denoted by $G_+(\boldsymbol{r};\epsilon)$, since it was obtained by using a *positive* constant η. Had η been negative, the significant pole would have been that at $k = -k' - \mathrm{i}\eta'$ and we should have obtained the complex conjugate of $G_+(\boldsymbol{r};\epsilon)$, denoted by $G_-(\boldsymbol{r};\epsilon)$. These Green functions are found in the theory of the scattering of a free particle by a central field, where they give rise to outgoing and incoming scattered waves, respectively. There are other possibilities, but we shall not pursue the matter further. Our principal purpose in calculating the Green function for this problem has been to justify the introduction of the constant $\mathrm{i}\eta$ and ultimately to take the limit as $\eta \to 0$. This will always be done in the following work.

It should be noted that eq. (10.4) gives only a *particular* solution of eq. (10.2). To obtain the complete solution we must add the appropriate solution of the homogeneous eq. (10.7), that is, the eigenfunction $\psi_{k'}(\boldsymbol{r})$ corresponding to the eigenvalue $\epsilon = k'^2/2m$ (assuming that ϵ lies in the quasi-continuum of eigenvalues of (10.7)). We then have

$$\psi(\boldsymbol{r}) = \psi_{k'}(\boldsymbol{r}) + \int G(\boldsymbol{r},\boldsymbol{r}';\epsilon)V(\boldsymbol{r}')\psi(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}'. \tag{10.33}$$

This integral equation may be solved by iteration. The zeroth approximation ignores the integral term and is just $\psi_{k'}(\boldsymbol{r})$. The first approximation is then

$$\psi^{(1)}(\boldsymbol{r}) = \psi_{k'}(\boldsymbol{r}) + \int G(\boldsymbol{r},\boldsymbol{r}';\epsilon)V(\boldsymbol{r}')\psi_{k'}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \tag{10.34}$$

obtained by substituting the zeroth approximation in the integrand, and the second approximation is obtained by substituting $\psi^{(1)}(\boldsymbol{r})$ in the integrand, and so on. The first approximation is known as the *Born approximation* in scattering theory.

10.2. The Density of States in Energy

For systems, such as metals, whose one-electron energy levels form a quasi-continuum, a function of great importance is the density of states in energy $\mathcal{N}(\epsilon)$. This may be defined as follows:

$\mathcal{N}(\epsilon)\,\mathrm{d}\epsilon$ *is the number of (orbital) states with energies lying between* ϵ *and* $\epsilon + \mathrm{d}\epsilon$, *where* $\mathrm{d}\epsilon$ *is small yet still large enough to contain many energy levels.* More precisely, if $\nu(\epsilon)$ is the number of states with energies less than ϵ, then

$$\mathcal{N}(\epsilon) = \frac{\mathrm{d}\nu}{\mathrm{d}\epsilon}. \tag{10.35}$$

The density of states may be expressed in terms of a Green function or some transform of it. Before showing this, we shall give a more general definition of the Green function, which applies to any system. What we have considered up till now is the *free-electron Green function.*

The time-independent Schrödinger equation for a one-electron system is

$$(H-\epsilon)\psi(\boldsymbol{r}) = 0, \tag{10.36}$$

subject to given boundary conditions. Let us denote by ψ_n and ϵ_n the eigenfunctions and corresponding eigenvalues of this equation, so that

$$(H-\epsilon_n)\psi_n(\boldsymbol{r}) = 0. \tag{10.37}$$

We assume that the functions $\psi_n(\boldsymbol{r})$ form a complete, orthonormal set: that is,

$$\int \psi_n^*(\boldsymbol{r})\psi_m(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \delta_{nm}, \tag{10.38}$$

the integral being taken throughout the volume Ω (which may be all space) over which the boundary conditions are applied.

We now *define* a Green function for the system as the solution of the equation

$$(H-\epsilon-\mathrm{i}\eta)G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta) = -\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{10.39}$$

subject to the same boundary conditions, η being, as before, a small positive constant which will ultimately tend to zero. Then

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta) = \sum_n \frac{\psi_n^*(\boldsymbol{r}')\psi_n(\boldsymbol{r})}{\epsilon-\epsilon_n+\mathrm{i}\eta}, \tag{10.40}$$

since

$$\begin{aligned}(H-\epsilon-\mathrm{i}\eta)\sum_n \frac{\psi_n^*(\boldsymbol{r}')\psi_n(\boldsymbol{r})}{\epsilon-\epsilon_n+\mathrm{i}\eta} &= \sum_n \frac{\psi_n^*(\boldsymbol{r}')}{\epsilon-\epsilon_n+\mathrm{i}\eta}(H-\epsilon-\mathrm{i}\eta)\psi_n(\boldsymbol{r}) \\ &= -\sum_n \psi_n^*(\boldsymbol{r}')\psi_n(\boldsymbol{r}) = -\delta(\boldsymbol{r}-\boldsymbol{r}').\end{aligned} \tag{10.41}$$

It follows from (10.38) and (10.40) that

$$\int G(\boldsymbol{r},\boldsymbol{r};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{r}=\sum_n\frac{1}{\epsilon-\epsilon_n+\mathrm{i}\eta}. \tag{10.42}$$

Now, if the energy levels form a quasi-continuum, with density of states $\mathcal{N}(\epsilon)$, the sum on the right-hand side may be replaced by an integral, and the equation becomes

$$\int G(\boldsymbol{r},\boldsymbol{r};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{r}=\int\frac{\mathcal{N}(\epsilon_n)\,\mathrm{d}\epsilon_n}{\epsilon-\epsilon_n+\mathrm{i}\eta}. \tag{10.43}$$

Hence, according to eq. (III.10), Appendix III,

$$\lim_{\eta\to0}\int G(\boldsymbol{r},\boldsymbol{r};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{r}=\mathscr{P}\int\frac{\mathcal{N}(\epsilon_n)\,\mathrm{d}\epsilon_n}{\epsilon-\epsilon_n}-\mathrm{i}\pi\mathcal{N}(\epsilon), \tag{10.44}$$

where $\mathscr{P}$ denotes the Cauchy principal value of the integral.

We thus have the result

$$\lim_{\eta\to0}\operatorname{Im}\int G(\boldsymbol{r},\boldsymbol{r};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{r}=-\pi\mathcal{N}(\epsilon), \tag{10.45}$$

and, provided the resulting integral exists, the limit may be taken immediately, to give

$$\mathcal{N}(\epsilon)=-\frac{1}{\pi}\operatorname{Im}\int G(\boldsymbol{r},\boldsymbol{r};\epsilon)\,\mathrm{d}\boldsymbol{r}. \tag{10.46}$$

That this formula does indeed work may be simply demonstrated in the case of free electrons: that is, when H_0 is substituted for H in eq. (10.36). From eq. (10.29), we have

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=-\frac{m\exp(\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}{2\pi|\boldsymbol{r}-\boldsymbol{r}'|}, \tag{10.47}$$

where $\epsilon=k^2/2m$, and

$$\operatorname{Im}G(\boldsymbol{r},\boldsymbol{r}';\epsilon)=-\frac{m\sin(k|\boldsymbol{r}-\boldsymbol{r}'|)}{2\pi|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{10.48}$$

Taking the limit of this as $\boldsymbol{r}\to\boldsymbol{r}'$ gives

$$\operatorname{Im}G(\boldsymbol{r},\boldsymbol{r};\epsilon)=-\frac{mk}{2\pi}. \tag{10.49}$$

Hence, from (10.46),

$$\mathcal{N}(\epsilon)=\frac{mk}{2\pi^2}\int\mathrm{d}\boldsymbol{r}=\frac{m\Omega k}{2\pi^2}=\frac{\Omega}{4\pi^2}(2m)^{3/2}\epsilon^{1/2}, \tag{10.50}$$

which is correct (see WM, eq. (7.5), putting $\hbar=1$).

10.3. The Fourier Transform of the Free-electron Green Function

The Schrödinger equation for a free electron is

$$(H_0-\epsilon)\psi(\boldsymbol{r})=0, \tag{10.51}$$

which, as before, we assume to be subject to periodic boundary conditions over a large cube of side L and volume Ω. As shown in § 10.1, we may define a simplified Green function for this problem by the equation

$$(H_0-\epsilon-\mathrm{i}\eta)G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)=-\delta(\boldsymbol{r}), \tag{10.52}$$

which gives

$$G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)=\frac{1}{8\pi^3}\int\frac{\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})}{\epsilon-\epsilon_k+\mathrm{i}\eta}\mathrm{d}\boldsymbol{k} \tag{10.53}$$

(i.e. eqs. (10.22) and (10.23)), where $\epsilon_k=k^2/2m$. The more general Green function $G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)$ is obtained by substituting $\boldsymbol{r}-\boldsymbol{r}'$ for $\boldsymbol{r}$ in (10.53).

The Fourier transform (see Appendix V) of $G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)$ is[†], by definition,

$$G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)=\int G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}, \tag{10.54}$$

so that

$$G(\boldsymbol{r};\epsilon+\mathrm{i}\eta)=\frac{1}{8\pi^3}\int G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k}. \tag{10.55}$$

Comparing eqs. (10.53) and (10.55), we deduce that

$$G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)=\frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta}. \tag{10.56}$$

[†] It is customary to use the same symbol (with different arguments, however) for the Green function and its Fourier transform, and either may be referred to as the Green function.

It should be noted that, in the present work, all integrals with respect to $\boldsymbol{r}$ are assumed to be throughout the region over which the boundary conditions are applied – in this case a cube of side L. In the definition of the Fourier transform, however, eq. (V.8), Appendix V, the integral is throughout all space. If L is large, as postulated, the difference may be neglected.

Thus

$$\sum_{\boldsymbol{k}} G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)=\sum_{\boldsymbol{k}}\frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta}, \tag{10.57}$$

and, replacing both sums by integrals, we have

$$\frac{\Omega}{8\pi^3}\int G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{k}=\int\frac{\mathcal{N}(\epsilon_k)\,\mathrm{d}\epsilon_k}{\epsilon-\epsilon_k+\mathrm{i}\eta}. \tag{10.58}$$

On the left-hand side we have used the fact that the density of states in $\boldsymbol{k}$-space is $\Omega/8\pi^3$, while on the right-hand side we have integrated with respect to ϵ_k by using the density of states in energy. Taking the limit as $\eta\to0$, we then find, as in eq. (10.45),

$$\mathcal{N}(\epsilon)=-\frac{\Omega}{8\pi^4}\lim_{\eta\to0}\operatorname{Im}\int G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)\,\mathrm{d}\boldsymbol{k}. \tag{10.59}$$

Clearly, in this case, with $G(\boldsymbol{k};\epsilon+\mathrm{i}\eta)$ given by (10.56), we cannot put $\eta=0$ in the integrand immediately, since the integral would then be real (and divergent). Although here we have been specifically concerned with free electrons (and to use this method of determining $N(\epsilon)$ for free electrons is like using a sledge-hammer to crack a walnut), we shall see in the following section that the same result applies in more complicated cases. The present section is intended merely as a gentle introduction to what follows.

10.4. The Density of States in Energy of a System Subject to a Small Perturbation

A generalization of the foregoing work may be used to obtain the density of states of a system subject to a small perturbation, on the assumption that the eigenvalues form a quasi-continuum. An instructive application[†] of this is to an electron in an almost periodic field, such as is found in a dilute metallic solid solution, where the periodic field of the solvent metal is perturbed by the presence of a small concentration of solute ions. In the following account we shall have in mind a perturbed periodic field, but we need not specify the cause of the perturbation.

Let us first consider the unperturbed system. Suppose that H is the Hamiltonian for an electron in the unperturbed periodic field, and that, subject to the usual periodic boundary conditions, the eigenfunctions (Bloch functions) and eigenvalues of this Hamiltonian are $u_k(\boldsymbol{r})$ and ϵ_k, respectively, where $\boldsymbol{k}$ is the wave vector. Thus

$$(H-\epsilon_k)u_k=0. \tag{10.60}$$

We now define a Green function $G_0(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)$ for this unperturbed system by the equation

$$(H-\epsilon-\mathrm{i}\eta)G_0(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)=-\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{10.61}$$

and, as for eq. (10.39), it follows that

$$G_0(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)=\sum_{\boldsymbol{k}}\frac{u_k^*(\boldsymbol{r}')u_k(\boldsymbol{r})}{\epsilon-\epsilon_k+\mathrm{i}\eta}. \tag{10.62}$$

It is convenient to transform this Green function with respect to the functions $u_k(\boldsymbol{r})$, according to the equation

$$G_0(\boldsymbol{l},\boldsymbol{m};\epsilon+\mathrm{i}\eta)=\iint G_0(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)u_l^*(\boldsymbol{r})u_m(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'. \tag{10.63}$$

This is a generalization of what was done in the previous section – if the $u_k(\boldsymbol{r})$ were free-electron functions (that is, if H were just H_0), then (10.63) would define a type of Fourier transform of the Green function.

Using eq. (10.62) and assuming that the u_k are orthonormal, so that

$$\int u_k^*u_l\,\mathrm{d}\boldsymbol{r}=\delta_{kl}, \tag{10.64}$$

we find

$$G_0(\boldsymbol{l},\boldsymbol{m};\epsilon+\mathrm{i}\eta)=\sum_{\boldsymbol{k}}\frac{\delta_{lk}\delta_{km}}{\epsilon-\epsilon_k+\mathrm{i}\eta}. \tag{10.65}$$

This is zero unless $\boldsymbol{l}=\boldsymbol{m}$, and then the sum reduces to the single term with $\boldsymbol{k}=\boldsymbol{l}=\boldsymbol{m}$. Thus

$$G_0(\boldsymbol{k},\boldsymbol{k};\epsilon+\mathrm{i}\eta)=\frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta}, \tag{10.66}$$

which we shall denote simply by $G_0(\boldsymbol{k})$. This is the same expression as in (10.56), but here, of course, the ϵ_k are not free-electron energy levels.

It follows, as before, that the density of states $\mathcal{N}_0(\epsilon)$ for the unperturbed system is

$$\mathcal{N}_0(\epsilon)=-\frac{\Omega}{8\pi^4}\lim_{\eta\to0}\operatorname{Im}\int G_0(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}. \tag{10.67}$$

[†] The method is essentially that developed by S. F. Edwards in a number of papers beginning with Edwards, S. F., 1958, Phil. Mag. **3**, 1020, but the application to dilute solid solutions, upon which the present account is based, was made by Jones, H., 1965, Proc. Roy. Soc. A**285**, 461; 1966, Proc Roy. Soc. A**294**, 405.

Now let us consider the perturbed system, and suppose that $V(\boldsymbol{r})$ is the potential energy due to the perturbation. We may assume the same periodic boundary conditions as before, but (unless $V(\boldsymbol{r})$ is also periodic) the eigenfunctions will no longer be Bloch functions; we denote them by ψ_n and the corresponding eigenvalues by ϵ_n, so that

$$(H+V-\epsilon_n)\psi_n = 0. \tag{10.68}$$

A Green function may be defined in the usual way by the equation

$$(H+V-\epsilon-\mathrm{i}\eta)G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta) = -\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{10.69}$$

from which it follows that

$$G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta) = \sum_n \frac{\psi_n^*(\boldsymbol{r}')\psi_n(\boldsymbol{r})}{\epsilon-\epsilon_n+\mathrm{i}\eta}, \tag{10.70}$$

on the assumption that the functions ψ_n are orthonormal.

Transforming this with respect to the unperturbed eigenfunctions $u_{\boldsymbol{k}}$, as for G_0 in eq. (10.63), we have

$$\begin{aligned} G(\boldsymbol{l},\boldsymbol{m};\epsilon+\mathrm{i}\eta) &= \iint G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)u_{\boldsymbol{l}}^*(\boldsymbol{r})u_{\boldsymbol{m}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}' \\ &= \sum_n \frac{1}{\epsilon-\epsilon_n+\mathrm{i}\eta}\int \psi_n(\boldsymbol{r})u_{\boldsymbol{l}}^*(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\int \psi_n^*(\boldsymbol{r}')u_{\boldsymbol{m}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \\ &= \sum_n \frac{a_n^*(\boldsymbol{l})a_n(\boldsymbol{m})}{\epsilon-\epsilon_n+\mathrm{i}\eta}, \end{aligned} \tag{10.71}$$

where

$$a_n(\boldsymbol{m}) = \int \psi_n^*(\boldsymbol{r})u_{\boldsymbol{m}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}. \tag{10.72}$$

In particular,

$$G(\boldsymbol{k},\boldsymbol{k};\epsilon+\mathrm{i}\eta) = \sum_n \frac{|a_n(\boldsymbol{k})|^2}{\epsilon-\epsilon_n+\mathrm{i}\eta}, \tag{10.73}$$

which we shall denote simply by $G(\boldsymbol{k})$. Thus

$$\sum_{\boldsymbol{k}} G(\boldsymbol{k}) = \sum_n\sum_{\boldsymbol{k}} \frac{|a_n(\boldsymbol{k})|^2}{\epsilon-\epsilon_n+\mathrm{i}\eta}. \tag{10.74}$$

Now,

$$\begin{aligned} \sum_{\boldsymbol{k}} |a_n(\boldsymbol{k})|^2 &= \sum_{\boldsymbol{k}} \int \psi_n(\boldsymbol{r})u_{\boldsymbol{k}}^*(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\int \psi_n^*(\boldsymbol{r}')u_{\boldsymbol{k}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \\ &= \int \psi_n(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\int \psi_n^*(\boldsymbol{r}')\sum_{\boldsymbol{k}} u_{\boldsymbol{k}}^*(\boldsymbol{r})u_{\boldsymbol{k}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \\ &= \int \psi_n(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\int \psi_n^*(\boldsymbol{r}')\delta(\boldsymbol{r}-\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \\ &= \int |\psi_n(\boldsymbol{r})|^2\mathrm{d}\boldsymbol{r} = 1, \end{aligned} \tag{10.75}$$

using the completeness condition for the functions $u_{\boldsymbol{k}}$. Hence

$$\sum_{\boldsymbol{k}} G(\boldsymbol{k}) = \sum_n \frac{1}{\epsilon-\epsilon_n+\mathrm{i}\eta}, \tag{10.76}$$

and, if we replace the sums by integrals, it follows that, as in eq. (10.67), the density of states $\mathcal{N}(\epsilon)$ for the perturbed system is

$$\mathcal{N}(\epsilon) = -\frac{\Omega}{8\pi^4}\lim_{\eta\to 0}\operatorname{Im}\int G(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}. \tag{10.77}$$

The assumption here is that the perturbed energy levels ϵ_n also form a quasi-continuum.

The function $G(\boldsymbol{k})$ in eq. (10.77) depends upon both the perturbed and the unperturbed eigenfunctions, as well as the perturbed eigenvalues. We shall now show how to express $G(\boldsymbol{k})$ in terms of $G_0(\boldsymbol{k})$, so that $\mathcal{N}(\epsilon)$ may be found in terms of the unperturbed eigenfunctions and eigenvalues only.

Writing simply $G(\boldsymbol{r},\boldsymbol{r}')$ for $G(\boldsymbol{r},\boldsymbol{r}';\epsilon+\mathrm{i}\eta)$, we note, first of all, that this function satisfies the integral equation

$$G(\boldsymbol{r},\boldsymbol{r}') = G_0(\boldsymbol{r},\boldsymbol{r}') + \int G_0(\boldsymbol{r},\boldsymbol{r}_1)V(\boldsymbol{r}_1)G(\boldsymbol{r}_1,\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}_1. \tag{10.78}$$

To prove this we operate on both sides with $H-\epsilon-\mathrm{i}\eta$, using eqs. (10.61) and (10.69), and obtain, from the left-hand side,

$$-\delta(\boldsymbol{r}-\boldsymbol{r}')-V(\boldsymbol{r})G(\boldsymbol{r},\boldsymbol{r}'),$$

and, from the right-hand side,

$$\begin{aligned} -\delta(\boldsymbol{r}-\boldsymbol{r}') &- \int \delta(\boldsymbol{r}-\boldsymbol{r}_1)V(\boldsymbol{r}_1)G(\boldsymbol{r}_1,\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}_1 \\ &= -\delta(\boldsymbol{r}-\boldsymbol{r}')-V(\boldsymbol{r})G(\boldsymbol{r},\boldsymbol{r}'). \end{aligned} \tag{10.79}$$

Iterating eq. (10.78), that is, repeatedly substituting for G in the integral, we find

$$G(\boldsymbol{r},\boldsymbol{r}') = G_0(\boldsymbol{r},\boldsymbol{r}') + \int G_0(\boldsymbol{r},\boldsymbol{r}_1)V(\boldsymbol{r}_1)G_0(\boldsymbol{r}_1,\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}_1$$
$$+ \iint G_0(\boldsymbol{r},\boldsymbol{r}_1)V(\boldsymbol{r}_1)G_0(\boldsymbol{r}_1,\boldsymbol{r}_2)V(\boldsymbol{r}_2)G_0(\boldsymbol{r}_2,\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2 + \dots \tag{10.80}$$

Now $G(\boldsymbol{k})$, or $G(\boldsymbol{k},\boldsymbol{k};\epsilon+\mathrm{i}\eta)$, is obtained by multiplying $G(\boldsymbol{r},\boldsymbol{r}')$ by $u_k^*(\boldsymbol{r})u_k(\boldsymbol{r}')$ and integrating with respect to $\boldsymbol{r}$ and $\boldsymbol{r}'$. Thus, using eq. (10.63),

$$G(\boldsymbol{k}) = G_0(\boldsymbol{k}) + \iiint G_0(\boldsymbol{r},\boldsymbol{r}_1)V(\boldsymbol{r}_1)G_0(\boldsymbol{r}_1,\boldsymbol{r}')u_k^*(\boldsymbol{r})u_k(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\mathrm{d}\boldsymbol{r}_1$$
$$+ \iiiint G_0(\boldsymbol{r},\boldsymbol{r}_1)V(\boldsymbol{r}_1)G_0(\boldsymbol{r}_1,\boldsymbol{r}_2)V(\boldsymbol{r}_2)G_0(\boldsymbol{r}_2,\boldsymbol{r}')$$
$$\times u_k^*(\boldsymbol{r})u_k(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2 + \dots \tag{10.81}$$

According to eqs. (10.62) and (10.64), the second term on the right-hand side becomes

$$\iiint \sum_l \frac{u_l^*(\boldsymbol{r}_1)u_l(\boldsymbol{r})}{\epsilon-\epsilon_l+\mathrm{i}\eta}V(\boldsymbol{r}_1) \sum_m \frac{u_m^*(\boldsymbol{r}')u_m(\boldsymbol{r}_1)}{\epsilon-\epsilon_m+\mathrm{i}\eta}\, u_k^*(\boldsymbol{r})u_k(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\mathrm{d}\boldsymbol{r}_1$$
$$= \int \frac{u_k^*(\boldsymbol{r}_1)V(\boldsymbol{r}_1)u_k(\boldsymbol{r}_1)}{(\epsilon-\epsilon_k+\mathrm{i}\eta)^2}\,\mathrm{d}\boldsymbol{r}_1$$
$$= [G_0(\boldsymbol{k})]^2 V_{kk}\,. \tag{10.82}$$

Here, and in what follows, we use the compact notation

$$V_{kl} = \int u_k^*(\boldsymbol{r})V(\boldsymbol{r})u_l(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}. \tag{10.83}$$

Similarly, the third term on the right-hand side of eq. (10.81) becomes

$$[G_0(\boldsymbol{k})]^2 \iint u_k^*(\boldsymbol{r}_1)V(\boldsymbol{r}_1)G_0(\boldsymbol{r}_1,\boldsymbol{r}_2)V(\boldsymbol{r}_2)u_k(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2$$
$$= [G_0(\boldsymbol{k})]^2 \iint u_k^*(\boldsymbol{r}_1)V(\boldsymbol{r}_1) \sum_l \frac{u_l^*(\boldsymbol{r}_2)u_l(\boldsymbol{r}_1)}{\epsilon-\epsilon_l+\mathrm{i}\eta}V(\boldsymbol{r}_2)u_k(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2$$
$$= [G_0(\boldsymbol{k})]^2 \sum_l G_0(\boldsymbol{l}) \int u_k^*(\boldsymbol{r}_1)V(\boldsymbol{r}_1)u_l(\boldsymbol{r}_1)\,\mathrm{d}\boldsymbol{r}_1 \int u_l^*(\boldsymbol{r}_2)V(\boldsymbol{r}_2)u_k(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_2$$
$$= [G_0(\boldsymbol{k})]^2 \sum_l G_0(\boldsymbol{l})V_{kl}\,V_{lk}\,. \tag{10.84}$$

It is fairly easy to see that the next term, not shown in eq. (10.81), must be

$$[G_0(\boldsymbol{k})]^2 \sum_l \sum_m G_0(\boldsymbol{l})G_0(\boldsymbol{m})V_{kl}V_{lm}\,V_{mk}\,, \tag{10.85}$$

and the evolution of the terms becomes obvious. We thus have

$$G(\boldsymbol{k}) = G_0(\boldsymbol{k})\Big[1+G_0(\boldsymbol{k})V_{kk} + G_0(\boldsymbol{k}) \sum_l G_0(\boldsymbol{l})V_{kl}\,V_{lk}$$
$$+ G_0(\boldsymbol{k}) \sum_l \sum_m G_0(\boldsymbol{l})G_0(\boldsymbol{m})V_{kl}\,V_{lm}\,V_{mk} + \dots\Big], \tag{10.86}$$

which expresses $G(\boldsymbol{k})$ in terms of the unperturbed Green function and the matrix elements of V with respect to the unperturbed eigenfunctions.

If (10.86) is substituted in (10.77) we obtain a series for the density of states $\mathcal{N}(\epsilon)$ in the perturbed system, which can be evaluated, in principle, using the unperturbed eigenfunctions and eigenvalues only. Clearly, this is a type of perturbation series, and, if the perturbation V is small (that is, if the matrix elements of V are small), one would expect to obtain sufficiently accurate results by considering only the first few terms†.

There would be little point in proceeding with the application of the foregoing to specific cases of practical interest, such as dilute solid solutions, since the calculations are complicated and the results are not relevant to the main theme of the book. Instead, we shall give two trivial but instructive examples which involve little calculation.

First, let us suppose that *V is constant*, so that

$$V_{kl} = V\delta_{kl}. \tag{10.87}$$

Then, from (10.86),

$$G(\boldsymbol{k}) = G_0(\boldsymbol{k})\{1+G_0(\boldsymbol{k})V+[G_0(\boldsymbol{k})V]^2+\dots\}$$
$$= \frac{G_0(\boldsymbol{k})}{1-G_0(\boldsymbol{k})V} = \frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta}\left(1-\frac{V}{\epsilon-\epsilon_k+\mathrm{i}\eta}\right)^{-1}$$
$$= \frac{1}{\epsilon-\epsilon_k-V+\mathrm{i}\eta}. \tag{10.88}$$

---

† It has been pointed out by A. D. Brailsford (Brailsford, A. D., 1966, Proc. Roy. Soc. A**292**, 433) that the use of the expansion in (10.86) to any given order in V must yield results identical with those given by ordinary perturbation theory of the same order, and the use of the latter may sometimes be more convenient than the Green function method. In the present chapter, however, our concern is merely to illustrate Green function techniques, regardless of the practical advantages to be gained or lost by their use in specific cases.

This is just $G_0(\boldsymbol{k})$ with $\epsilon - V$ substituted for ϵ. Hence, from eqs. (10.67) and (10.77), we deduce, without further calculation, that

$$\mathscr{N}(\epsilon) = \mathscr{N}_0(\epsilon - V). \tag{10.89}$$

This is obviously correct, since each energy level of the unperturbed system has been raised by an amount V.

Secondly, as a slightly less trivial example, let us suppose that $V(\boldsymbol{r})$ is not constant in space, but is so small that its off-diagonal matrix elements may be neglected. In other words, suppose that we may take

$$V_{kl} = V_{kk}\,\delta_{kl}. \tag{10.90}$$

A calculation similar to (10.88) then gives immediately

$$G(\boldsymbol{k}) = \frac{1}{\epsilon - \epsilon_k - V_{kk} + \mathrm{i}\eta}. \tag{10.91}$$

If V_{kk} varies with $\boldsymbol{k}$, therefore, the perturbed density of states curve will have a different shape from the unperturbed one. If, however, the unperturbed eigenfunctions are free-electron functions or, more generally, Wigner–Seitz functions (see WM, p. 255) of the type

$$u_k(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,u(\boldsymbol{r}), \tag{10.92}$$

where $u(\boldsymbol{r})$ is independent of $\boldsymbol{k}$, then

$$V_{kk} = \int u_k^* V u_k \,\mathrm{d}\boldsymbol{r} = \int V|u|^2 \,\mathrm{d}\boldsymbol{r} = v, \tag{10.93}$$

say, the same for all $\boldsymbol{k}$, and it follows, as before, that

$$\mathscr{N}(\epsilon) = \mathscr{N}_0(\epsilon - v). \tag{10.94}$$

This is an example of what is known in the theory of alloys as the *rigid-band model*, according to which the *shape* of the density of states curve is not changed by the perturbation. Again, the result is obvious from elementary considerations.

10.5. Green Function Operators

In manipulations such as those described in the foregoing section it is often found convenient to make use of the *Green function operators*†, defined by

$$G_0 = \frac{1}{\epsilon - H + \mathrm{i}\eta}, \tag{10.95}$$

$$G = \frac{1}{\epsilon - H - V + \mathrm{i}\eta}. \tag{10.96}$$

We shall now show that the functions $G_0(\boldsymbol{k})$ and $G(\boldsymbol{k})$ in eq. (10.86) are diagonal matrix elements of these operators with respect to the functions $u_k(\boldsymbol{r})$.

First, we have

$$\begin{aligned}\langle \boldsymbol{k}|G_0|\boldsymbol{k}\rangle &\equiv \int u_k^*(\boldsymbol{r})\frac{1}{\epsilon - H + \mathrm{i}\eta}u_k(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \\ &= \int \frac{|u_k(\boldsymbol{r})|^2}{\epsilon - \epsilon_k + \mathrm{i}\eta}\mathrm{d}\boldsymbol{r} = \frac{1}{\epsilon - \epsilon_k + \mathrm{i}\eta} = G_0(\boldsymbol{k}),\end{aligned} \tag{10.97}$$

which proves that $G_0(\boldsymbol{k})$ is a diagonal matrix element of G_0.

Now,

$$\langle \boldsymbol{k}|G|\boldsymbol{k}\rangle \equiv \int u_k^*(\boldsymbol{r})\frac{1}{\epsilon - H - V + \mathrm{i}\eta}u_k(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}. \tag{10.98}$$

Expanding u_k in terms of the perturbed eigenfunctions ψ_n, eq. (10.68), we have

$$u_k(\boldsymbol{r}) = \sum_n a_n(\boldsymbol{k})\psi_n(\boldsymbol{r}), \tag{10.99}$$

where $a_n(\boldsymbol{k})$ is defined by eq. (10.72). Thus

$$\langle \boldsymbol{k}|G|\boldsymbol{k}\rangle = \sum_{m,n}\int a_m^*(\boldsymbol{k})\psi_m^*(\boldsymbol{r})\frac{1}{\epsilon - H - V + \mathrm{i}\eta}a_n(\boldsymbol{k})\psi_n(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$$

---

† These operators are closely related to the so-called *resolvent* operators. The resolvent of a Hamiltonian H, for example, is defined as $(z - H)^{-1}$, or the negative of this, where z is a complex variable. Resolvents have been used by N. M. Hugenholtz (Hugenholtz, N. M., 1957, Physica **23**, 481) to develop a graphical analysis of the perturbation series for many-particle systems similar to that of Goldstone described in the present work.

$$= \sum_{m,n} a_m^*(\boldsymbol{k}) a_n(\boldsymbol{k}) \int \frac{\psi_m^*(\boldsymbol{r})\psi_n(\boldsymbol{r})}{\epsilon-\epsilon_n+\mathrm{i}\eta}\mathrm{d}\boldsymbol{r}$$

$$= \sum_{n} \frac{|a_n(\boldsymbol{k})|^2}{\epsilon-\epsilon_n+\mathrm{i}\eta} = G(\boldsymbol{k}). \tag{10.100}$$

from the definition of $G(\boldsymbol{k})$ in eq. (10.73). This proves that $G(\boldsymbol{k})$ is a diagonal matrix element of G.

We are now in a position to obtain the expansion (10.68) in a very simple manner. By elementary operator algebra, we have

$$\begin{aligned}
&\frac{1}{\epsilon-H+\mathrm{i}\eta}+\frac{1}{\epsilon-H+\mathrm{i}\eta}V\frac{1}{\epsilon-H-V+\mathrm{i}\eta}\\
&\quad=\frac{1}{\epsilon-H+\mathrm{i}\eta}\left(1+V\frac{1}{\epsilon-H-V+\mathrm{i}\eta}\right)\\
&\quad=\frac{1}{\epsilon-H+\mathrm{i}\eta}(\epsilon-H-V+\mathrm{i}\eta+V)\frac{1}{\epsilon-H-V+\mathrm{i}\eta}\\
&\quad=\frac{1}{\epsilon-H-V+\mathrm{i}\eta}.
\end{aligned} \tag{10.101}$$

In other words,

$$G = G_0 + G_0VG, \tag{10.102}$$

and iteration of this operator equation gives

$$G = G_0 + G_0VG + G_0VG_0VG_0 + \dots. \tag{10.103}$$

Taking diagonal matrix elements of both sides with respect to u_k, we have

$$G(\boldsymbol{k}) = G_0(\boldsymbol{k}) + \langle \boldsymbol{k}|G_0VG_0|\boldsymbol{k}\rangle + \dots. \tag{10.104}$$

The second term on the right-hand side is

$$\begin{aligned}
\left\langle \boldsymbol{k}\left|\frac{1}{\epsilon-H+\mathrm{i}\eta}V\frac{1}{\epsilon-H+\mathrm{i}\eta}\right|\boldsymbol{k}\right\rangle &= \frac{1}{(\epsilon-\epsilon_k+\mathrm{i}\eta)^2}\langle \boldsymbol{k}|V|\boldsymbol{k}\rangle\\
&= [G_0(\boldsymbol{k})]^2V_{kk},
\end{aligned} \tag{10.105}$$

using the result expressed in eq. (I.24), Appendix I. The other terms may be found similarly and it is seen that eq. (10.104) reduces to eq. (10.86).

10.6. Time-dependent Green Functions

Although we have been specifically concerned with the time-independent Schrödinger equation for a one-electron system, that is, an equation in three spatial dimensions, the Green function method can be formally extended to any number of dimensions. In particular, the time can be included, and we shall very briefly consider this case as it will be convenient to refer to it in the following chapter.

The time-dependent Schrödinger equation for a single electron moving in an electrostatic field with potential energy $V(\boldsymbol{r},t)$ is

$$\left[H_0 + V(\boldsymbol{r},t) - \mathrm{i}\frac{\partial}{\partial t}\right]\psi(\boldsymbol{r},t) = 0, \tag{10.106}$$

where, as before,

$$H_0 = -\frac{1}{2m}\nabla^2. \tag{10.107}$$

Eq. (10.106) may be written in the form (cf. eq. (10.2))

$$\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)\psi(\boldsymbol{r},t) = -V(\boldsymbol{r},t)\psi(\boldsymbol{r},t), \tag{10.108}$$

and, by an obvious extension of the work of §10.1, we define a Green function for this problem by means of the equation

$$\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{r},t,\boldsymbol{r}',t') = -\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t'), \tag{10.109}$$

subject to the same boundary conditions as (10.108). The operator on the left is not affected by a change of origin of either $\boldsymbol{r}$ or t, so that we must have, as in eq. (10.16),

$$G_0(\boldsymbol{r},t,\boldsymbol{r}',t') = G_0(\boldsymbol{r}-\boldsymbol{r}', t-t'), \tag{10.110}$$

and, as in eq. (10.17), we may confine ourselves to the simplified function $G_0(\boldsymbol{r},t)$ satisfying the equation

$$\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{r},t) = -\delta(\boldsymbol{r})\delta(t). \tag{10.111}$$

$G_0(\boldsymbol{r},t)$ is the time-dependent free-electron Green function†.

---

† We use the symbol G_0, rather than G as in §10.1, merely to emphasize that this is the *unperturbed* (in this case, free-electron) Green function, according to the notation of §10.4.

Eq. (10.108) may now be expressed as the integral equation

$$\psi(\boldsymbol{r},t) = u(\boldsymbol{r},t) + \iint_{-\infty}^{\infty} G_0(\boldsymbol{r}-\boldsymbol{r}', t-t')V(\boldsymbol{r}',t')\psi(\boldsymbol{r}',t')\,\mathrm{d}t'\,\mathrm{d}\boldsymbol{r}', \tag{10.112}$$

where $u(\boldsymbol{r},t)$ is any solution of the equation

$$\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)u(\boldsymbol{r},t) = 0. \tag{10.113}$$

We prove this by operating upon both sides of (10.112) with $(H_0 - \mathrm{i}\partial/\partial t)$, thus:

$$\begin{aligned}\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)\psi(\boldsymbol{r},t) &= \left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)u(\boldsymbol{r},t) \\ &\quad + \iint_{-\infty}^{\infty}\left(H_0 - \mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{r}-\boldsymbol{r}', t-t') \\ &\quad \times V(\boldsymbol{r}',t')\psi(\boldsymbol{r}',t')\,\mathrm{d}t'\,\mathrm{d}\boldsymbol{r}' \\ &= -\iint_{-\infty}^{\infty}\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t')V(\boldsymbol{r}',t')\psi(\boldsymbol{r}',t')\,\mathrm{d}t'\,\mathrm{d}\boldsymbol{r}' \\ &= -V(\boldsymbol{r},t)\psi(\boldsymbol{r},t),\end{aligned} \tag{10.114}$$

which agrees with eq. (10.108). The solution $u(\boldsymbol{r},t)$ of (10.113) to be used depends upon the physical conditions. For example, a plane wave solution is

$$u_k(\boldsymbol{r},t) = \exp\left[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \epsilon_k t)\right], \tag{10.115}$$

where, as before,

$$\epsilon_k = k^2/2m, \tag{10.116}$$

and, if this is used in (10.112), $\psi(\boldsymbol{r},t)$ will give the perturbation of this plane wave by $V(\boldsymbol{r},t)$ – in other words, the *scattering* of an initially freely-moving electron, with momentum $\boldsymbol{k}$, by $V(\boldsymbol{r},t)$. We shall not elaborate upon this as the further development is not relevant to our future work.

In conclusion, we consider the Fourier transform $G_0(\boldsymbol{k},t)$ of $G_0(\boldsymbol{r},t)$ with respect to $\boldsymbol{r}$, which is defined by

$$G_0(\boldsymbol{r},t) = \frac{1}{8\pi^3}\int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})G_0(\boldsymbol{k},t)\,\mathrm{d}\boldsymbol{k}. \tag{10.117}$$

Substituting this in eq. (10.111) gives

$$\frac{1}{8\pi^3}\int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\left(\frac{k^2}{2m} - \mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{k},t)\,\mathrm{d}\boldsymbol{k} = -\frac{\delta(t)}{8\pi^3}\int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k}, \tag{10.118}$$

since, from eq. (V.11), Appendix V,

$$\delta(\boldsymbol{r}) = \frac{1}{8\pi^3}\int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k}, \tag{10.119}$$

and we deduce that the equation satisfied by $G_0(\boldsymbol{k},t)$ is

$$\left(\frac{k^2}{2m} - \mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{k},t) = -\delta(t). \tag{10.120}$$

The solution of this equation will be found in the following chapter.

CHAPTER 11

GREEN FUNCTIONS FOR MANY-ELECTRON SYSTEMS

11.1. Introduction

In the previous chapter we have seen how some of the properties of a one-electron system (or, indeed, of a many-electron system in which the electronic interactions are ignored) may be expressed in terms of Green functions. Similar functions may be defined for interacting many-electron systems, and these are also called Green functions, although, as we shall see, they do not in general satisfy the same sort of differential equations as the Green functions we have already met.

Of particular interest are the *one-particle* and *two-particle* Green functions, so-called because they describe the motion of one electron and of two electrons, respectively, added to the many-electron system. Most of the important properties of the system can, in fact, be expressed in terms of the one-particle Green function, and it is this which will mainly concern us in the present chapter, although there will be a brief mention of the two-particle Green function in the final section.

11.2. The One-particle Green Function

We shall use the field operators $\psi(x)$ and $\psi^\dagger(x)$ which were introduced in §2.8. The definitions given there (eqs. (2.130) and (2.131)) were

$$\psi(x) = \sum_i \phi_i(x) c_i, \tag{11.1}$$

$$\psi^\dagger(x) = \sum_i \phi_i^*(x) c_i^\dagger, \tag{11.2}$$

where $\phi_i(x)$ are the one-electron functions or *basis functions* (including spin) to be used in constructing the wave functions of our many-electron

system, and the destruction and creation operators c_i, $c_i^\dagger$ refer to these functions. Although it had not been made explicit in ch. 2, it should be clear from the work of ch. 5 that here we are using the *Schrödinger picture*, in which the operators are independent of time. In fact, in most of the present chapter, we wish to use the *Heisenberg picture*, in which the operators are time-dependent and the wave functions time-independent. If H is the Hamiltonian operator (which we shall assume does not depend explicitly upon the time) of our interacting many-electron system, then, according to eq. (5.13), we have†

$$\psi(x,t) = \exp(\mathrm{i}Ht)\psi(x)\exp(-\mathrm{i}Ht), \tag{11.3}$$

$$\psi^\dagger(x,t) = \exp(\mathrm{i}Ht)\psi^\dagger(x)\exp(-\mathrm{i}Ht), \tag{11.4}$$

or

$$\psi(x,t) = \sum_i \phi_i(x) c_i(t), \tag{11.5}$$

$$\psi^\dagger(x,t) = \sum_i \phi_i^*(x) c_i^\dagger(t), \tag{11.6}$$

where

$$c_i(t) = \exp(\mathrm{i}Ht) c_i \exp(-\mathrm{i}Ht), \tag{11.7}$$

$$c_i^\dagger(t) = \exp(\mathrm{i}Ht) c_i^\dagger \exp(-\mathrm{i}Ht). \tag{11.8}$$

We note that

$$\begin{aligned} \psi(x) &\equiv \psi(x,0), \quad & \psi^\dagger(x) &\equiv \psi^\dagger(x,0), \\ c_i &\equiv c_i(0), \quad & c_i^\dagger &\equiv c_i^\dagger(0). \end{aligned} \tag{11.9}$$

It is easily verified that all the commutation relations of §§2.3 and 2.8 remain valid when the operators are time-dependent, provided the operators are taken at the same time. For example,

$$\begin{aligned} \{\psi(x,t), \psi^\dagger(x',t)\} &= \exp(\mathrm{i}Ht)\{\psi(x), \psi^\dagger(x')\}\exp(-\mathrm{i}Ht) \\ &= \delta(x-x'). \end{aligned} \tag{11.10}$$

This is not true if the operators do not have the same time argument.

We shall also make use of the time-ordering operator T, due to Wick.

---

† In this chapter again we set $\hbar = 1$. All time-dependent operators will be in the Heisenberg picture, unless otherwise stated.

For any two *creation or destruction* operators $A(t)$, $B(t')$ we have[†]

$$T\{A(t)B(t')\} = \begin{cases} A(t)B(t'), & \text{if } t > t', \\ -B(t')A(t), & \text{if } t' > t. \end{cases} \tag{11.11}$$

The operator T thus orders the product of operators so that their times decrease from left to right, but differs from the operator P defined in §5.5 in that a minus sign is introduced when the two operators have to be permuted in order to obtain this time-ordering. In general, for the product of any number of creation or destruction operators, $A(t_1)B(t_2)C(t_3)\ldots$, the operator T orders them with times decreasing from left to right and multiplies them by $(-1)^p$, where p is the number of permutations of pairs of adjacent operators required to bring about this ordering.

We may now define the one-particle Green function $G(xt, x't')$ for the system as

$$G(xt,x't') = -\mathrm{i}\langle\Psi_0|T\{\boldsymbol{\psi}(x,t)\boldsymbol{\psi}^\dagger(x',t')\}|\Psi_0\rangle, \tag{11.12}$$

where Ψ_0 is the true ground-state wave function of the interacting system in the Heisenberg picture – that is, the Schrödinger wave function at time $t = 0$. Thus

$$G(xt,x't') = \begin{cases} -\mathrm{i}\langle\Psi_0|\boldsymbol{\psi}(x,t)\boldsymbol{\psi}^\dagger(x',t')|\Psi_0\rangle & \text{if } t > t', \\ +\mathrm{i}\langle\Psi_0|\boldsymbol{\psi}^\dagger(x',t')\boldsymbol{\psi}(x,t)|\Psi_0\rangle, & \text{if } t' > t. \end{cases} \tag{11.13}$$

Notice that $G(xt,x't')$ is not defined at $t = t'$, since the operator T itself is not defined there.

Another way of writing $G(xt,x't')$ in a single expression, which does not involve the operator T explicitly, utilizes the *Heaviside unit step function*, defined by

$$\theta(t) = \begin{cases} 0, & t < 0, \\ 1, & t > 0. \end{cases} \tag{11.14}$$

This gives

$$\begin{aligned} G(xt,x't') = &-\mathrm{i}\theta(t-t')\langle\Psi_0|\boldsymbol{\psi}(x,t)\boldsymbol{\psi}^\dagger(x',t')|\Psi_0\rangle \\ &+\mathrm{i}\theta(t'-t)\langle\Psi_0|\boldsymbol{\psi}^\dagger(x',t')\boldsymbol{\psi}(x,t)|\Psi_0\rangle. \end{aligned} \tag{11.15}$$

No logical reason for this definition is yet apparent, but it will be seen below that the function so defined is convenient for describing many properties of the system. For example, the operator representing the density of electrons at position $\boldsymbol{r}$ and time t, with given spin, is, from eq. (2.142),

$$\rho(\boldsymbol{x},t) = \boldsymbol{\psi}^\dagger(\boldsymbol{x},t)\boldsymbol{\psi}(\boldsymbol{x},t). \tag{11.16}$$

The expectation value of this in the ground state is

$$\begin{aligned} \langle\rho(\boldsymbol{x},t)\rangle &= \langle\Psi_0|\boldsymbol{\psi}^\dagger(\boldsymbol{x},t)\boldsymbol{\psi}(\boldsymbol{x},t)|\Psi_0\rangle \\ &= -\mathrm{i}\lim_{t'\to t+0} G(\boldsymbol{x}t,\boldsymbol{x}t'). \end{aligned} \tag{11.17}$$

Taking the limit as $t' \to t+0$ ensures that $t' > t$, which gives the correct order for the operators.

The expectation value of any sum of one-electron operators may similarly be expressed in terms of the Green function. Consider such a sum of one-electron operators $\Sigma_i f(\boldsymbol{x}_i,t)$, one for each electron of the system. In the formalism of second quantization this sum becomes

$$F = \int \boldsymbol{\psi}^\dagger(\boldsymbol{x},t)f(\boldsymbol{x},t)\boldsymbol{\psi}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x} \tag{11.18}$$

(e.g. eq. (2.137) in the Heisenberg picture). The expectation value of this operator in the ground state is

$$\langle F\rangle = \left\langle\Psi_0\middle|\int \boldsymbol{\psi}^\dagger(\boldsymbol{x},t)f(\boldsymbol{x},t)\boldsymbol{\psi}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}\middle|\Psi_0\right\rangle, \tag{11.19}$$

which may be written

$$\langle F\rangle = -\mathrm{i}\int \lim_{\substack{x'\to x \\ t'\to t+0}} f(\boldsymbol{x},t)G(\boldsymbol{x}t,\boldsymbol{x}'t')\,\mathrm{d}\boldsymbol{x}. \tag{11.20}$$

Here, since $\boldsymbol{x}$ consists of the position vector $\boldsymbol{r}$ and spin coordinate ζ, the limit as $\boldsymbol{x}' \to \boldsymbol{x}$ means the limit as $\boldsymbol{r}' \to \boldsymbol{r}$ with $\zeta' = \zeta$. It is necessary to use some such device, rather than simply setting $\boldsymbol{x}' = \boldsymbol{x}$ immediately, because we wish $f(\boldsymbol{x},t)$ to act upon $\boldsymbol{\psi}(\boldsymbol{x},t)$ only, not upon $\boldsymbol{\psi}^\dagger(\boldsymbol{x},t)$. Another of the many ways of expressing this is

$$\langle F\rangle = -\mathrm{i}\int \lim_{t'\to t+0} [f(\boldsymbol{x},t)G(\boldsymbol{x}t,\boldsymbol{x}'t')]_{x'=x}\,\mathrm{d}\boldsymbol{x}. \tag{11.21}$$

We shall now show that for a time-independent system (that is, one for which H does not depend explicitly upon the time) $G(\boldsymbol{x}t,\boldsymbol{x}'t')$ depends

---

[†] The curly brackets here do *not* imply the anticommutator of $A(t)$ and $B(t')$ – they merely enclose the product of operators upon which T acts.

only upon the time difference $t-t'$ and not upon t and t' separately. We have, for $t > t'$,

$$\begin{aligned} G(xt,x't') &= -\mathrm{i}\langle\Psi_0|\boldsymbol{\psi}(x,t)\boldsymbol{\psi}^\dagger(x',t')|\Psi_0\rangle \\ &= -\mathrm{i}\langle\Psi_0|\exp(\mathrm{i}Ht)\boldsymbol{\psi}(x)\exp(-\mathrm{i}Ht)\exp(\mathrm{i}Ht')\boldsymbol{\psi}^\dagger(x') \\ &\qquad \times\exp(-\mathrm{i}Ht')|\Psi_0\rangle \\ &= -\mathrm{i}\langle\Psi_0|\exp(\mathrm{i}Ht')\exp[\mathrm{i}H(t-t')]\boldsymbol{\psi}(x)\exp[-\mathrm{i}H(t-t')] \\ &\qquad \times\boldsymbol{\psi}^\dagger(x')|\Psi_0\rangle\exp(-\mathrm{i}Et'), \end{aligned} \tag{11.22}$$

where E is the ground-state energy, so that

$$H\Psi_0 = E\Psi_0. \tag{11.23}$$

Also, from Appendix I, eq. (I.24), we have, for any function Ψ,

$$\langle\Psi_0|\exp(\mathrm{i}Ht')|\Psi\rangle = \exp(\mathrm{i}Et')\langle\Psi_0|\Psi\rangle. \tag{11.24}$$

Hence, for $t > t'$,

$$\begin{aligned} G(xt,x't') &= -\mathrm{i}\langle\Psi_0|\exp[\mathrm{i}H(t-t')]\boldsymbol{\psi}(x)\exp[-\mathrm{i}H(t-t')]\boldsymbol{\psi}^\dagger(x')|\Psi_0\rangle \\ &= -\mathrm{i}\langle\Psi_0|\boldsymbol{\psi}(x,t-t')\boldsymbol{\psi}^\dagger(x',0)|\Psi_0\rangle. \end{aligned} \tag{11.25}$$

Similarly, for $t' > t$,

$$G(xt,x't') = +\mathrm{i}\langle\Psi_0|\boldsymbol{\psi}^\dagger(x',0)\boldsymbol{\psi}(x,t-t')|\Psi_0\rangle. \tag{11.26}$$

We thus have

$$\begin{aligned} G(xt,x't') &= -\mathrm{i}\langle\Psi_0|T\{\boldsymbol{\psi}(x,t-t')\boldsymbol{\psi}^\dagger(x',0)\}|\Psi_0\rangle \\ &= G(x,t-t',x'0). \end{aligned} \tag{11.27}$$

11.3. Elimination of the Spin

In some texts on Green functions electron spin is simply ignored, while in others it is treated in a rather cavalier fashion which is not very helpful to the non-expert. Our definition of the one-particle Green function, eq. (11.12), includes the spin, since $\boldsymbol{x}$ comprises both space and spin coordinates and the one-electron functions $\phi_i(\boldsymbol{x})$ have spin factors. However, it is mathematically convenient to eliminate the spin. At the risk of appearing tedious it may be advisable to show in detail how this can be done.

First, in anticipation of our application to a free-electron gas and, in principle at least, to metals, where the one-electron functions are characterized by a wave vector $\boldsymbol{k}$ and spin quantum number σ, we shall write, instead of $\phi_i(\boldsymbol{x})$,

$$\phi_{k\sigma}(\boldsymbol{x}) \equiv \psi_k(\boldsymbol{r})\chi_\sigma(\zeta), \tag{11.28}$$

as in ch. 3, eq. (3.39), where σ can have the values $\pm\frac{1}{2}$ and

$$\chi_{1/2}(\zeta) = \alpha(\zeta), \qquad \chi_{-1/2}(\zeta) = \beta(\zeta), \tag{11.29}$$

with

$$\begin{aligned} \alpha(1) &= 1, & \alpha(-1) &= 0, \\ \beta(1) &= 0, & \beta(-1) &= 1. \end{aligned} \tag{11.30}$$

The definitions of $\boldsymbol{\psi}(\boldsymbol{x})$ and $\boldsymbol{\psi}^\dagger(\boldsymbol{x})$ then become

$$\boldsymbol{\psi}(\boldsymbol{x}) \equiv \boldsymbol{\psi}(\boldsymbol{r},\zeta) = \sum_k\sum_\sigma \psi_k(\boldsymbol{r})\chi_\sigma(\zeta)c_{k\sigma}, \tag{11.31}$$

$$\boldsymbol{\psi}^\dagger(\boldsymbol{x}) \equiv \boldsymbol{\psi}^\dagger(\boldsymbol{r},\zeta) = \sum_k\sum_\sigma \psi_k^*(\boldsymbol{r})\chi_\sigma(\zeta)c_{k\sigma}^\dagger. \tag{11.32}$$

Thus, since each $\psi_k(\boldsymbol{r})$ appears in two spin-orbitals, with spin factors α and β respectively, we have

$$\boldsymbol{\psi}(\boldsymbol{r},1) = \sum_k \psi_k(\boldsymbol{r})c_{k,1/2} \tag{11.33}$$

and

$$\boldsymbol{\psi}(\boldsymbol{r},-1) = \sum_k \psi_k(\boldsymbol{r})c_{k,-1/2}. \tag{11.34}$$

Let us denote these by $\boldsymbol{\psi}_{1/2}(\boldsymbol{r})$ and $\boldsymbol{\psi}_{-1/2}(\boldsymbol{r})$, respectively – in other words, let

$$\boldsymbol{\psi}_\sigma(\boldsymbol{r}) = \sum_k \psi_k(\boldsymbol{r})c_{k\sigma}, \tag{11.35}$$

so that

$$\boldsymbol{\psi}(\boldsymbol{x}) = \sum_\sigma \chi_\sigma(\zeta)\boldsymbol{\psi}_\sigma(\boldsymbol{r}). \tag{11.36}$$

In the Heisenberg picture, similarly, we have

$$\boldsymbol{\psi}_\sigma(\boldsymbol{r},t) = \sum_k \psi_k(\boldsymbol{r})c_{k\sigma}(t) \tag{11.37}$$

and

$$\boldsymbol{\psi}(\boldsymbol{x},t) = \sum_\sigma \chi_\sigma(\zeta)\boldsymbol{\psi}_\sigma(\boldsymbol{r},t). \tag{11.38}$$

Also,

$$\boldsymbol{\psi}_\sigma^\dagger(\boldsymbol{r},t) = \sum_k \psi_k^*(\boldsymbol{r}) c_{k\sigma}^\dagger(t) \tag{11.39}$$

and

$$\boldsymbol{\psi}^\dagger(\boldsymbol{x},t) = \sum_\sigma \chi_\sigma(\zeta)\boldsymbol{\psi}_\sigma^\dagger(\boldsymbol{r},t). \tag{11.40}$$

The one-particle Green function thus becomes

$$G(\boldsymbol{x}t,\boldsymbol{x}'t') = \sum_\sigma \sum_{\sigma'} \chi_\sigma(\zeta)\chi_{\sigma'}(\zeta')G_{\sigma\sigma'}(\boldsymbol{r}t,\boldsymbol{r}'t'), \tag{11.41}$$

where

$$\begin{aligned} G_{\sigma\sigma'}(\boldsymbol{r}t,\boldsymbol{r}'t') &= -\mathrm{i}\langle\Psi_0|T\{\boldsymbol{\psi}_\sigma(\boldsymbol{r},t)\boldsymbol{\psi}_{\sigma'}^\dagger(\boldsymbol{r}',t')\}|\Psi_0\rangle \\ &= -\mathrm{i}\sum_k\sum_{k'} \langle\Psi_0|T\{c_{k\sigma}(t)c_{k'\sigma'}^\dagger(t')\}|\Psi_0\rangle\psi_k(\boldsymbol{r})\psi_{k'}^*(\boldsymbol{r}'). \end{aligned} \tag{11.42}$$

We may therefore confine our attention to the latter function.

For a non-ferromagnetic system, in the absence of an external magnetic field, the matrix elements in (11.42) vanish if $\sigma \neq \sigma'$, since the spins in Ψ_0 must balance in pairs. We shall only consider such systems, and so, from (11.41), we have

$$G(\boldsymbol{x}t,\boldsymbol{x}'t') = \begin{cases} G_{1/2,1/2}(\boldsymbol{r}t,\boldsymbol{r}'t'), & \text{if } \zeta=\zeta'=1, \\ G_{-1/2,-1/2}(\boldsymbol{r}t,\boldsymbol{r}'t'), & \text{if } \zeta=\zeta'=-1, \\ 0, & \text{if } \zeta\neq\zeta'. \end{cases} \tag{11.43}$$

Also, it is easy to see that

$$G_{1/2,1/2}(\boldsymbol{r}t,\boldsymbol{r}'t') = G_{-1/2,-1/2}(\boldsymbol{r}t,\boldsymbol{r}'t'), \tag{11.44}$$

since each orbital function $\psi_k(\boldsymbol{r})$ is associated with both α and β spin functions in the set of basis functions $\phi_{k\sigma}(\boldsymbol{x})$.

The Green function $G(\boldsymbol{x}t,\boldsymbol{x}'t')$ is therefore completely determined by

$$G_{\sigma\sigma}(\boldsymbol{r}t,\boldsymbol{r}'t') = -\mathrm{i}\sum_k\sum_{k'} \langle\Psi_0|T\{c_{k\sigma}(t)c_{k'\sigma}^\dagger(t')\}|\Psi_0\rangle\psi_k(\boldsymbol{r})\psi_{k'}^*(\boldsymbol{r}'), \tag{11.45}$$

with *either* $\sigma=\frac{1}{2}$ *or* $\sigma=-\frac{1}{2}$. It is customary to omit the σs and simply write this as

$$G(\boldsymbol{r}t,\boldsymbol{r}'t') = -\mathrm{i}\sum_k\sum_{k'} \langle\Psi_0|T\{c_k(t)c_{k'}^\dagger(t')\}|\Psi_0\rangle\psi_k(\boldsymbol{r})\psi_{k'}^*(\boldsymbol{r}'), \tag{11.46}$$

or

$$G(\boldsymbol{r}t,\boldsymbol{r}'t') = -\mathrm{i}\langle\Psi_0|T\{\boldsymbol{\psi}(\boldsymbol{r},t)\boldsymbol{\psi}^\dagger(\boldsymbol{r}',t')\}|\Psi_0\rangle. \tag{11.47}$$

This is the one-particle Green function with spin eliminated.

The matrix element (multiplied by $-\mathrm{i}$) in (11.46) is the transform of $G(\boldsymbol{r}t,\boldsymbol{r}'t')$ with respect to the functions $\psi_k(\boldsymbol{r})$ (cf. eq. (10.63)). As in the previous chapter, we refer to this also as the Green function and use the same symbol G for it, thus:

$$G(\boldsymbol{k}t,\boldsymbol{k}'t') = -\mathrm{i}\langle\Psi_0|T\{c_k(t)c_{k'}^\dagger(t')\}|\Psi_0\rangle. \tag{11.48}$$

For time-independent systems, as shown previously for $G(\boldsymbol{x}t,\boldsymbol{x}'t')$, it follows that

$$G(\boldsymbol{k}t,\boldsymbol{k}'t') = -\mathrm{i}\langle\Psi_0|T\{c_k(t-t')c_{k'}^\dagger(0)\}|\Psi_0\rangle. \tag{11.49}$$

It is therefore sufficient to consider the simplified Green function

$$G(\boldsymbol{k},\boldsymbol{k}',t) = -\mathrm{i}\langle\Psi_0|T\{c_k(t)c_{k'}^\dagger(0)\}|\Psi_0\rangle. \tag{11.50}$$

In the same way we may write

$$G(\boldsymbol{r},\boldsymbol{r}',t) = -\mathrm{i}\langle\Psi_0|T\{\boldsymbol{\psi}(\boldsymbol{r},t)\boldsymbol{\psi}^\dagger(\boldsymbol{r}',0)\}|\Psi_0\rangle. \tag{11.51}$$

11.4. Physical Interpretation of the Green Function

It is not strictly necessary to give a physical interpretation of the Green function, since its properties and applications follow directly from its definition. In fact, however, it does have a simple physical interpretation which is often a useful intuitive guide in calculations, particularly those involved in perturbation theory. We shall consider both $G(\boldsymbol{k},\boldsymbol{k}',t)$ and $G(\boldsymbol{r},\boldsymbol{r}',t)$.

Let us first recall from elementary wave mechanics that, if $\Psi_n(0)$ are the stationary states of a system at time $t=0$, then the (Schrödinger) wave function $\Psi(t)$ at time t can be expanded in the form (see WM, p. 313)

$$\Psi(t) = \sum_n A_n(t)\exp(-\mathrm{i}E_nt)\Psi_n(0), \tag{11.52}$$

where

$$A_n(t) = \exp(\mathrm{i}E_nt)\langle\Psi_n(0)|\Psi(t)\rangle. \tag{11.53}$$

$|A_n(t)|^2$ is interpreted as the probability that the system is in the state $\Psi_n(0)$ at time t. $A_n(t)$ is called the *probability amplitude* of the system being in the state $\Psi_n(0)$ at time t.

If the system is in the state $\Psi(0)$ at $t=0$, the (Schrödinger) wave function at time t is (eq. (5.4))

$$\Psi(t) = \exp(-\mathrm{i}Ht)\Psi(0), \tag{11.54}$$

so that, in this case,

$$A_n(t) = \exp(\mathrm{i}E_n t)\langle\Psi_n(0)|\exp(-\mathrm{i}Ht)|\Psi(0)\rangle. \tag{11.55}$$

Now, for $t > 0$, we have†

$$\begin{aligned} G(\boldsymbol{k},\boldsymbol{k}',t) &= -\mathrm{i}\langle\Psi_0|c_{\boldsymbol{k}}(t)c^\dagger_{\boldsymbol{k}'}(0)|\Psi_0\rangle \\ &= -\mathrm{i}\langle\Psi_0|\exp(\mathrm{i}Ht)c_{\boldsymbol{k}}\exp(-\mathrm{i}Ht)c^\dagger_{\boldsymbol{k}'}|\Psi_0\rangle \\ &= -\mathrm{i}\exp(\mathrm{i}Et)\langle\Psi_0|c_{\boldsymbol{k}}\exp(-\mathrm{i}Ht)c^\dagger_{\boldsymbol{k}'}|\Psi_0\rangle \\ &= -\mathrm{i}\exp(\mathrm{i}Et)\langle c^\dagger_{\boldsymbol{k}}\Psi_0|\exp(-\mathrm{i}Ht)|c^\dagger_{\boldsymbol{k}'}\Psi_0\rangle. \end{aligned} \tag{11.56}$$

$c^\dagger_{\boldsymbol{k}'}\Psi_0$ represents the state of an $(N+1)$-electron system obtained from the ground state of the N-electron system at $t = 0$ by adding one electron in the state $\phi_{\boldsymbol{k}'\sigma}$ (the σ is implied in $c^\dagger_{\boldsymbol{k}'}$). Similarly, $c^\dagger_{\boldsymbol{k}}\Psi_0$ is the state obtained from the ground state at $t = 0$ by adding one electron in the state $\phi_{\boldsymbol{k}\sigma}$. Comparing (11.55) and (11.56), we see that, apart from a numerical factor of modulus unity, we may loosely‡ interpret $G(\boldsymbol{k},\boldsymbol{k}',t)$ as the probability amplitude of the $(N+1)$-electron system, which is in the state $c_{\boldsymbol{k}'}\Psi_0$ at $t = 0$, being in the state $c^\dagger_{\boldsymbol{k}}\Psi_0$ at a later time t.

$G(\boldsymbol{r},\boldsymbol{r}',t)$ can be interpreted similarly. For $t > 0$, we have

$$\begin{aligned} G(\boldsymbol{r},\boldsymbol{r}',t) &= -\mathrm{i}\langle\Psi_0|\exp(\mathrm{i}Ht)\boldsymbol{\psi}(\boldsymbol{r})\exp(-\mathrm{i}Ht)\boldsymbol{\psi}^\dagger(\boldsymbol{r}')|\Psi_0\rangle \\ &= -\mathrm{i}\exp(\mathrm{i}Et)\langle\Psi_0|\boldsymbol{\psi}(\boldsymbol{r})\exp(-\mathrm{i}Ht)\boldsymbol{\psi}^\dagger(\boldsymbol{r}')|\Psi_0\rangle \\ &= -\mathrm{i}\exp(\mathrm{i}Et)\langle\boldsymbol{\psi}^\dagger(\boldsymbol{r})\Psi_0|\exp(-\mathrm{i}Ht)|\boldsymbol{\psi}^\dagger(\boldsymbol{r}')\Psi_0\rangle. \end{aligned} \tag{11.57}$$

According to the interpretation of the field operators given in §2.8, $\boldsymbol{\psi}^\dagger(\boldsymbol{r}')\Psi_0$ is the state of an $(N+1)$-electron system obtained from the ground state of the N-electron system at $t = 0$ by adding an electron localized at $\boldsymbol{r}'$. Comparing (11.55) and (11.57), we see that, again apart from a numerical factor of modulus unity, we may loosely interpret $G(\boldsymbol{r},\boldsymbol{r}',t)$ as the probability amplitude of the $(N+1)$-electron system, which is in the state $\boldsymbol{\psi}^\dagger(\boldsymbol{r}')\Psi_0$ at $t = 0$, being in the state $\boldsymbol{\psi}^\dagger(\boldsymbol{r})\Psi_0$ at a later time t. In other words, it is the probability amplitude that an electron added at position $\boldsymbol{r}'$ to the N-electron system in its ground state at $t = 0$ will have moved or *propagated* to position $\boldsymbol{r}$ at a later time t. For this reason the one-particle Green function is often called the *one-particle propagator*.

We leave the interpretation in the case $t < 0$ as an exercise for the reader.

11.5. Free-electron Gas

We shall now consider the case of a free-electron gas, as described in §3.4, that is, a system of N electrons and a neutralizing uniform distribution of positive charge contained in a cubic box of side L and volume $\Omega = L^3$. Assuming periodic boundary conditions over the surface of the box, we take the normalized one-electron orbital functions to have the form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega}}\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}). \tag{11.58}$$

It will be convenient in general to take L to be very large, so that for many purposes the box may be regarded as infinite.

The Green function $G(\boldsymbol{r},\boldsymbol{r}',t)$ then depends only upon t and the difference $\boldsymbol{r}-\boldsymbol{r}'$. This is sufficiently obvious from the interpretation of $G(\boldsymbol{r},\boldsymbol{r}',t)$ given in the previous section, since, for the system described above, the probability amplitude for the propagation of an added electron from $\boldsymbol{r}'$ to $\boldsymbol{r}$ must clearly depend only upon the position of $\boldsymbol{r}$ relative to $\boldsymbol{r}'$, namely $\boldsymbol{r}-\boldsymbol{r}'$, and not upon the absolute positions of the two points. It should be noted, however, that this would not be the case if an external field, such as the periodic field of the ion-cores in a metal, were present.

Another instructive way of looking at the problem derives from eq. (11.46), which becomes, with $t' = 0$ and using (11.58),

$$G(\boldsymbol{r},\boldsymbol{r}',t) = -\frac{\mathrm{i}}{\Omega}\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'}\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\exp(-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}')\langle\Psi_0|T\{c_{\boldsymbol{k}}(t)c^\dagger_{\boldsymbol{k}'}(0)\}|\Psi_0\rangle. \tag{11.59}$$

The required result then follows from the fact that the matrix element, which is just $\mathrm{i}G(\boldsymbol{k},\boldsymbol{k}',t)$, vanishes if $\boldsymbol{k} \neq \boldsymbol{k}'$. A convincing proof of this is not quite so simple as the airy dismissal of the subject in most texts would suggest, and so a slight digression might be worth while (any reader who wishes to omit this may proceed immediately to eq. (11.63)).

---

† As in eq. (11.22) we retain E as the energy of the ground state Ψ_0, so that in the expansion (11.52) we must set $E_0 = E$.

‡ The comparison is not quite accurate since, except for a non-interacting system, $c^\dagger_{\boldsymbol{k}}\Psi_0$ is not a *stationary* state of the $(N+1)$-electron system. Also E is the ground-state energy of the N-electron system.

For $t > 0$, the matrix element is

$$\begin{aligned}\langle\Psi_0|c_{\boldsymbol{k}}(t)c^{\dagger}_{\boldsymbol{k}'}(0)|\Psi_0\rangle &= \langle\Psi_0|\exp(\mathrm{i}Ht)c_{\boldsymbol{k}}\exp(-\mathrm{i}Ht)c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle\\ &= \exp(\mathrm{i}Et)\langle\Psi_0|c_{\boldsymbol{k}}\exp(-\mathrm{i}Ht)c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle. \end{aligned}\tag{11.60}$$

We now expand $c^{\dagger}_{\boldsymbol{k}'}\Psi_0$ in terms of the orthonormal eigenfunctions Ψ^{N+1}_n of an $(N+1)$-electron system, thus:

$$c^{\dagger}_{\boldsymbol{k}'}\Psi_0 = \sum_n \langle\Psi^{N+1}_n|c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle\Psi^{N+1}_n, \tag{11.61}$$

and (11.60) becomes

$$\begin{aligned}&\langle\Psi_0|c_{\boldsymbol{k}}(t)c^{\dagger}_{\boldsymbol{k}'}(0)|\Psi_0\rangle\\ &\quad = \exp(\mathrm{i}Et)\sum_n \langle\Psi^{N+1}_n|c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle\langle\Psi_0|c_{\boldsymbol{k}}\exp(-\mathrm{i}Ht)|\Psi^{N+1}_n\rangle\\ &\quad = \sum_n \langle\Psi^{N+1}_n|c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle\langle\Psi_0|c_{\boldsymbol{k}}|\Psi^{N+1}_n\rangle\exp[\mathrm{i}(E-E^{N+1}_n)t]\\ &\quad = \sum_n \langle\Psi^{N+1}_n|c^{\dagger}_{\boldsymbol{k}'}|\Psi_0\rangle\langle\Psi^{N+1}_n|c^{\dagger}_{\boldsymbol{k}}|\Psi_0\rangle^*\exp[\mathrm{i}(E-E^{N+1}_n)t], \end{aligned}\tag{11.62}$$

where E^{N+1}_n is the energy of the state Ψ^{N+1}_n.

Now, if the system were *non-interacting*, so that Ψ_0 and Ψ^{N+1}_n were single determinants of orders N and $N+1$, respectively, the product of matrix elements in (11.62) would clearly vanish if $\boldsymbol{k} \neq \boldsymbol{k}'$, since $c^{\dagger}_{\boldsymbol{k}}$ and $c^{\dagger}_{\boldsymbol{k}'}$ would then operate upon Ψ_0 to give two *different* $(N+1)$-electron states, one at least of which would be orthogonal to Ψ^{N+1}_n. The general case of interacting electrons is not so straightforward, however, since both Ψ_0 and Ψ^{N+1}_n are then infinite series of determinants. To proceed in this case we must invoke the fact that Ψ_0 and Ψ^{N+1}_n are eigenfunctions of the total momentum operator†. This means, on the assumption that the total momentum of the ground state is zero, that the $\boldsymbol{k}$-vectors of every determinant in Ψ_0 must sum to zero. The effect of the operator $c^{\dagger}_{\boldsymbol{k}'}$ on Ψ_0 will thus be to produce a series of determinants of order $N+1$, the sum of the $\boldsymbol{k}$-vectors of each being $\boldsymbol{k}'$, and similarly for $c^{\dagger}_{\boldsymbol{k}}\Psi_0$. Since Ψ^{N+1}_n is also an eigenfunction of the total momentum operator, every determinant in its expansion must have the same $\boldsymbol{k}$-sum, and so must be orthogonal to $c^{\dagger}_{\boldsymbol{k}}\Psi_0$ or to $c^{\dagger}_{\boldsymbol{k}'}\Psi_0$, or to both, if $\boldsymbol{k} \neq \boldsymbol{k}'$. The product of matrix elements in (11.62) must therefore vanish if $\boldsymbol{k} \neq \boldsymbol{k}'$ – in other words, $G(\boldsymbol{k},\boldsymbol{k}',t)$, with $t > 0$, vanishes for a free-electron system if $\boldsymbol{k} \neq \boldsymbol{k}'$. Clearly, the same result applies if $t < 0$.

---

† The momentum of a free electron in the state $\phi_{\boldsymbol{k}\sigma}$ is $\boldsymbol{k}$. Hence (when we restore the spin indices which are omitted in the text for convenience) the operator representing the total momentum is

$$\sum_{\boldsymbol{k}\sigma}\boldsymbol{k}c^{\dagger}_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma}.$$

This operator commutes with the Hamiltonian (eq. (3.49)) and hence has simultaneous eigenfunctions with it. We may thus assume that Ψ_0 is an eigenfunction of the total momentum operator, with eigenvalue zero.

Eq. (11.59) thus reduces to

$$G(\boldsymbol{r},\boldsymbol{r}',t) = -\frac{\mathrm{i}}{\Omega}\sum_{\boldsymbol{k}}\exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')]\langle\Psi_0|T\{c_{\boldsymbol{k}}(t)c^{\dagger}_{\boldsymbol{k}}(0)\}|\Psi_0\rangle. \tag{11.63}$$

It is therefore sufficient to consider the simpler function

$$G(\boldsymbol{r},t) = -\frac{\mathrm{i}}{\Omega}\sum_{\boldsymbol{k}}\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\langle\Psi_0|T\{c_{\boldsymbol{k}}(t)c^{\dagger}_{\boldsymbol{k}}(0)\}|\Psi_0\rangle. \tag{11.64}$$

Expressing this as an integral, by means of (10.14), we have

$$G(\boldsymbol{r},t) = \frac{1}{8\pi^3}\int\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})G(\boldsymbol{k},t)\,\mathrm{d}\boldsymbol{k}, \tag{11.65}$$

where

$$G(\boldsymbol{k},t) = -\mathrm{i}\langle\Psi_0|T\{c_{\boldsymbol{k}}(t)c^{\dagger}_{\boldsymbol{k}}(0)\}|\Psi_0\rangle \tag{11.66}$$

(i.e. $G(\boldsymbol{k},\boldsymbol{k}',t)$ with $\boldsymbol{k}=\boldsymbol{k}'$). Thus, $G(\boldsymbol{k},t)$ is the Fourier transform of $G(\boldsymbol{r},t)$, and is given by

$$G(\boldsymbol{k},t) = \int\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})G(\boldsymbol{r},t)\,\mathrm{d}\boldsymbol{r}. \tag{11.67}$$

It is also useful to define the time transform of $G(\boldsymbol{k},t)$, or four-dimensional Fourier transform of $G(\boldsymbol{r},t)$, by

$$G(\boldsymbol{k},\epsilon) = \int_{-\infty}^{\infty}\exp(\mathrm{i}\epsilon t)G(\boldsymbol{k},t)\,\mathrm{d}t, \tag{11.68}$$

so that

$$G(\boldsymbol{k},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-\mathrm{i}\epsilon t)G(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon. \tag{11.69}$$

To add to the confusion, $G(\boldsymbol{k},\epsilon)$ is also called a one-particle Green function.

In this section we have gone to considerable trouble to relate the simplified functions $G(\boldsymbol{k},t), G(\boldsymbol{r},t)$ to the more general functions

$G(\boldsymbol{k},\boldsymbol{k}',t)$, $G(\boldsymbol{r},\boldsymbol{r}',t)$, for a free-electron system. We wish to remark, however, that this is by no means obligatory. Many texts *define* the one-particle Green function as $G(\boldsymbol{k},t)$, given in eq. (11.66), even for systems other than the free-electron system considered here. This is because $G(\boldsymbol{k},t)$, and its time transform $G(\boldsymbol{k},\epsilon)$, are the *most useful* Green functions – we shall see in §§11.9 and 11.10 that they give the expectation value of the number operator and the ground-state energy.

11.6. Non-interacting Particles

Let us now consider the case of a free-electron gas when the interaction of the electrons is neglected. The ground state is then Φ_0, a single determinant in which all the one-electron states within the Fermi sphere are occupied and all those outside the Fermi sphere are unoccupied. Thus, for a particular spin, the occupation number of the state $\boldsymbol{k}$ is given by

$$n_k = \begin{cases} 1, & \text{if } \ k < k_{\mathrm{F}}, \\ 0, & \text{if } \ k > k_{\mathrm{F}}. \end{cases} \tag{11.70}$$

We shall denote the Green function for this system by $G_0(\boldsymbol{k},t)$. Then

$$G_0(\boldsymbol{k},t) = -\mathrm{i}\langle\Phi_0|T\{c_k(t)c_k^\dagger(0)\}|\Phi_0\rangle. \tag{11.71}$$

The total Hamiltonian in this case is just H_0, so that

$$c_k(t) = \exp(\mathrm{i}H_0t)c_k\exp(-\mathrm{i}H_0t). \tag{11.72}$$

Hence, for $t > 0$,

$$\begin{aligned} G_0(\boldsymbol{k},t) &= -\mathrm{i}\langle\Phi_0|c_k(t)c_k^\dagger(0)|\Phi_0\rangle \\ &= -\mathrm{i}\langle\Phi_0|\exp(\mathrm{i}H_0t)c_k\exp(-\mathrm{i}H_0t)c_k^\dagger|\Phi_0\rangle \\ &= -\mathrm{i}\exp(\mathrm{i}E_0t)\langle\Phi_0|c_k\exp(-\mathrm{i}H_0t)c_k^\dagger|\Phi_0\rangle \\ &= -\mathrm{i}\langle\Phi_0|c_kc_k^\dagger|\Phi_0\rangle\exp(-\mathrm{i}\epsilon_kt) \\ &= -\mathrm{i}(1-n_k)\exp(-\mathrm{i}\epsilon_kt), \end{aligned} \tag{11.73}$$

since

$$H_0c_k^\dagger\Phi_0 = (E_0+\epsilon_k)c_k^\dagger\Phi_0, \tag{11.74}$$

where, as before, $\epsilon_k = k^2/2m$.

For $t < 0$, similarly,

$$\begin{aligned} G_0(\boldsymbol{k},t) &= \mathrm{i}\langle\Phi_0|c_k^\dagger(0)c_k(t)|\Phi_0\rangle \\ &= \mathrm{i}\langle\Phi_0|c_k^\dagger c_k|\Phi_0\rangle\exp(-\mathrm{i}\epsilon_kt) \\ &= \mathrm{i}n_k\exp(-\mathrm{i}\epsilon_kt). \end{aligned} \tag{11.75}$$

Thus

$$G_0(\boldsymbol{k},t) = \begin{cases} \mathrm{i}(n_k-1)\exp(-\mathrm{i}\epsilon_kt), & \text{if } \ t>0, \\ \mathrm{i}n_k\exp(-\mathrm{i}\epsilon_kt), & \text{if } \ t<0, \end{cases} \tag{11.76}$$

or, using the Heaviside unit step function defined in eq. (11.14),

$$G_0(\boldsymbol{k},t) = \mathrm{i}[\theta(t)(n_k-1)+\theta(-t)n_k]\exp(-\mathrm{i}\epsilon_kt). \tag{11.77}$$

We note that

$$\lim_{t\to+0}G_0(\boldsymbol{k},t) + \lim_{t\to-0}G_0(\boldsymbol{k},t) = -\mathrm{i}, \tag{11.78}$$

so that $G_0(\boldsymbol{k},t)$ has a finite discontinuity at $t = 0$.

Now, let us consider $G_0(\boldsymbol{k},\epsilon)$, the time transform of $G_0(\boldsymbol{k},t)$. From (11.70) and (11.77), we have, for $k > k_{\mathrm{F}}$,

$$G_0(\boldsymbol{k},t) = -\mathrm{i}\theta(t)\exp(-\mathrm{i}\epsilon_kt), \tag{11.79}$$

so that, from (11.68),

$$\begin{aligned} G_0(\boldsymbol{k},\epsilon) &= \int_{-\infty}^{\infty}\exp(\mathrm{i}\epsilon t)G_0(\boldsymbol{k},t)\,\mathrm{d}t \\ &= -\mathrm{i}\int_0^{\infty}\exp[\mathrm{i}(\epsilon-\epsilon_k)t]\,\mathrm{d}t. \end{aligned} \tag{11.80}$$

The integral does not converge, so that as in the previous chapter we introduce a factor $\mathrm{e}^{-\eta t}$, with η positive, into the integrand, ultimately taking the limit as $\eta \to +0$. Thus we write

$$\begin{aligned} G_0(\boldsymbol{k},\epsilon) &= \lim_{\eta\to+0} -\mathrm{i}\int_0^{\infty}\exp[\mathrm{i}(\epsilon-\epsilon_k+\mathrm{i}\eta)t]\,\mathrm{d}t \\ &= \lim_{\eta\to+0}\frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta}. \end{aligned} \tag{11.81}$$

Similarly, for $k < k_{\mathrm{F}}$,

$$G_0(\boldsymbol{k},t) = \mathrm{i}\theta(-t)\exp(-\mathrm{i}\epsilon_kt), \tag{11.82}$$

and

$$G_0(\boldsymbol{k},\epsilon) = \mathrm{i}\int_{-\infty}^{0} \exp\,[\mathrm{i}(\epsilon-\epsilon_k)t]\,\mathrm{d}t$$

$$= \lim_{\eta\to+0} \mathrm{i}\int_{-\infty}^{0} \exp\,[\mathrm{i}(\epsilon-\epsilon_k-\mathrm{i}\eta)t]\,\mathrm{d}t$$

$$= \lim_{\eta\to+0} \frac{1}{\epsilon-\epsilon_k-\mathrm{i}\eta}. \tag{11.83}$$

We may combine (11.81) and (11.83) by writing

$$G_0(\boldsymbol{k},\epsilon) = \lim_{\eta_k\to0} \frac{1}{\epsilon-\epsilon_k+\mathrm{i}\eta_k}, \tag{11.84}$$

where

$$\left.\begin{aligned} \eta_k > 0, &\quad \text{if}\quad k > k_\mathrm{F},\\ \eta_k < 0, &\quad \text{if}\quad k < k_\mathrm{F}. \end{aligned}\right\} \tag{11.85}$$

The justification for these η tricks is simply that the inverse time transform of $G_0(\boldsymbol{k},\epsilon)$, so found, is indeed $G_0(\boldsymbol{k},t)$, as we shall now show. Using eqs. (11.69), (11.84) and (III.11), Appendix III, we find, for $k > k_\mathrm{F}$,

$$G_0(\boldsymbol{k},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\,(-\mathrm{i}\epsilon t)\,G_0(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon$$

$$= \frac{1}{2\pi}\mathscr{P}\int_{-\infty}^{\infty} \frac{\exp\,(-\mathrm{i}\epsilon t)\,\mathrm{d}\epsilon}{\epsilon-\epsilon_k} - \tfrac{1}{2}\mathrm{i}\int_{-\infty}^{\infty} \exp\,(-\mathrm{i}\epsilon t)\,\delta(\epsilon-\epsilon_k)\,\mathrm{d}\epsilon$$

$$= \frac{1}{2\pi}\mathscr{P}\int_{-\infty}^{\infty} \frac{\exp\,(-\mathrm{i}\epsilon t)\,\mathrm{d}\epsilon}{\epsilon-\epsilon_k} - \tfrac{1}{2}\mathrm{i}\exp\,(-\mathrm{i}\epsilon_k t). \tag{11.86}$$

It is shown in Appendix VI, eqs. (VI.24) and (VI.25), that the Cauchy principal value $\mathscr{P}$ of the integral is given by

$$\mathscr{P}\int_{-\infty}^{\infty} \frac{\exp\,(-\mathrm{i}\epsilon t)\,\mathrm{d}\epsilon}{\epsilon-\epsilon_k} = \begin{cases} -\mathrm{i}\pi\exp\,(-\mathrm{i}\epsilon_k t), & \text{if}\quad t>0,\\ \mathrm{i}\pi\exp\,(-\mathrm{i}\epsilon_k t), & \text{if}\quad t<0. \end{cases} \tag{11.87}$$

Substitution in (11.86) gives, for $k > k_\mathrm{F}$,

$$G_0(\boldsymbol{k},t) = \begin{cases} -\mathrm{i}\exp\,(-\mathrm{i}\epsilon_k t), & \text{if}\quad t>0,\\ 0, & \text{if}\quad t<0, \end{cases} \tag{11.88}$$

which agrees with (11.79). Similarly, for $k < k_\mathrm{F}$, η_k is negative and, again according to eq. (III.11), this only changes the sign of the second term on the right-hand side of (11.86). Thus, for $k < k_\mathrm{F}$, we have

$$G_0(\boldsymbol{k},t) = \begin{cases} 0, & \text{if}\quad t>0,\\ \mathrm{i}\exp\,(-\mathrm{i}\epsilon_k t), & \text{if}\quad t<0, \end{cases} \tag{11.89}$$

in agreement with (11.82).

11.7. Differential Equation for Green Function of Non-interacting System

The equation of motion of the one-particle Green function for an interacting system will be discussed in §11.12. Meanwhile, however, we may show very easily that the Green function $G_0(\boldsymbol{k},t)$ for a system of non-interacting free particles satisfies eq. (10.120), which is the differential equation for the Green function of the one-electron time-dependent Schrödinger equation.

We have, for $k > k_\mathrm{F}$,

$$G_0(\boldsymbol{k},t) = -\mathrm{i}\theta(t)\exp\,(-\mathrm{i}\epsilon_k t). \tag{11.90}$$

Now the Heaviside unit step function may be expressed in the form

$$\theta(t) = \int_{-\infty}^{t} \delta(x)\,\mathrm{d}x, \tag{11.91}$$

which gives immediately

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \delta(t). \tag{11.92}$$

Hence, for $k > k_\mathrm{F}$,

$$\frac{\partial}{\partial t}G_0(\boldsymbol{k},t) = -\mathrm{i}\delta(t)\exp\,(-\mathrm{i}\epsilon_k t) - \theta(t)\epsilon_k\exp\,(-\mathrm{i}\epsilon_k t), \tag{11.93}$$

and it follows that

$$\left(\frac{k^2}{2m}-\mathrm{i}\frac{\partial}{\partial t}\right)G_0(\boldsymbol{k},t) = \left[-\frac{\mathrm{i}k^2}{2m}\theta(t)-\delta(t)+\mathrm{i}\theta(t)\epsilon_k\right]\exp\,(-\mathrm{i}\epsilon_k t)$$

$$= -\delta(t)\exp\,(-\mathrm{i}\epsilon_k t) = -\delta(t), \tag{11.94}$$

since $\epsilon_k = k^2/2m$, and, for any function $f(t)$ which is well-behaved at $t = 0$,

$$\delta(t)f(t) = \delta(t)f(0). \tag{11.95}$$

It is easily shown in the same way that $G_0(\boldsymbol{k},t)$ also satisfies (11.94) if $k < k_{\mathrm{F}}$.

$G_0(\boldsymbol{k},t)$ for a many-particle system therefore satisfies eq. (10.120), which is a justification of the use of the name *Green function* for the functions relating to many-particle systems defined in this chapter. It will be seen later, however, that the one-particle Green function for a system of interacting particles does not satisfy the same sort of differential equation as the Green function for a one-particle system.

11.8. Lehmann Representation

In this section we shall derive certain mathematical properties of one-particle Green functions which are frequently encountered in the literature. We shall, however, make no applications of these properties in the present work, so that any reader who merely wants a brief survey of the field of Green functions may omit this section.

We obtain immediately from eqs. (11.62) and (11.66) that, in the case of a free-electron gas, whether interacting or not,

for $t > 0$,

$$G(\boldsymbol{k},t) = -\mathrm{i} \sum_n |\langle \Psi_n^{N+1} | c_{\boldsymbol{k}}^\dagger | \Psi_0 \rangle|^2 \exp\left[\mathrm{i}(E - E_n^{N+1})t\right], \tag{11.96}$$

where E_n^{N+1} is the energy of the state† Ψ_n^{N+1} of an $(N+1)$-electron system and E, as before, is the ground-state energy of an N-electron system. We may, in fact, write

$$\begin{aligned} E_n^{N+1} - E &= (E_n^{N+1} - E_0^{N+1}) + (E_0^{N+1} - E) \\ &= \epsilon_n^{N+1} + \mu^{N+1}, \end{aligned} \tag{11.97}$$

where ϵ_n^{N+1} is the *excitation energy* of the state Ψ_n^{N+1} (i.e. the energy of that state minus the ground-state energy E_0^{N+1}) and μ^{N+1} is the *chemical potential*‡ for an $(N+1)$-electron system (i.e. the ground-state energy

---

† As explained in §11.5, only states with total momentum $\boldsymbol{k}$ need be included.

‡ For a *non-interacting* free-electron gas at the absolute zero of temperature the chemical potential is just the energy at the surface of the Fermi sphere, that is, $\epsilon_{\mathrm{F}} = k_{\mathrm{F}}^2/2m$.

of an $(N+1)$-electron system minus that of an N-electron system). Thus, for $t > 0$,

$$G(\boldsymbol{k},t) = -\mathrm{i} \sum_n |\langle \Psi_n^{N+1} | c_{\boldsymbol{k}}^\dagger | \Psi_0 \rangle|^2 \exp\left[-\mathrm{i}(\epsilon_n^{N+1} + \mu^{N+1})t\right]. \tag{11.98}$$

Similarly, we find that,

for $t < 0$,

$$G(\boldsymbol{k},t) = \mathrm{i} \sum_n |\langle \Psi_n^{N-1} | c_{\boldsymbol{k}} | \Psi_0 \rangle|^2 \exp\left[\mathrm{i}(\epsilon_n^{N-1} - \mu^N)t\right], \tag{11.99}$$

where

$$\epsilon_n^{N-1} = E_n^{N-1} - E_0^{N-1} \tag{11.100}$$

and

$$\mu^N = E - E_0^{N-1}, \tag{11.101}$$

E_n^{N-1}, Ψ_n^{N-1} being the energy and wave function of a state† of an $(N-1)$-electron system.

Now, clearly, μ^N must be virtually the same as μ^{N+1}, provided N is very large – as it is for metallic systems, for example – since the energy loss due to removing a single electron cannot depend appreciably upon whether N or $N+1$ electrons are originally present (it is, in fact, independent of N to order $1/N$). Therefore, we may write

$$\mu^{N+1} = \mu^N = \mu. \tag{11.102}$$

Finally, then, we have

$$G(\boldsymbol{k},t) = \begin{cases} -\mathrm{i} \sum_n |\langle \Psi_n^{N+1} | c_{\boldsymbol{k}}^\dagger | \Psi_0 \rangle|^2 \exp\left[-\mathrm{i}(\epsilon_n^{N+1} + \mu)t\right], & t > 0, \\ \mathrm{i} \sum_n |\langle \Psi_n^{N-1} | c_{\boldsymbol{k}} | \Psi_0 \rangle|^2 \exp\left[\mathrm{i}(\epsilon_n^{N-1} - \mu)t\right], & t < 0. \end{cases} \tag{11.103}$$

It is found convenient to convert these expressions to integrals, by means of the *spectral density functions*, defined by‡

$$A(\boldsymbol{k},\epsilon) = \sum_n |\langle \Psi_n^{N+1} | c_{\boldsymbol{k}}^\dagger | \Psi_0 \rangle|^2 \delta(\epsilon - \epsilon_n^{N+1}), \tag{11.104}$$

$$B(\boldsymbol{k},\epsilon) = \sum_n |\langle \Psi_n^{N-1} | c_{\boldsymbol{k}} | \Psi_0 \rangle|^2 \delta(\epsilon - \epsilon_n^{N-1}). \tag{11.105}$$

---

† Here only states of total momentum $-\boldsymbol{k}$ need be included.

‡ Other notations, such as $\rho^+(\boldsymbol{k},\epsilon), \rho^-(\boldsymbol{k},\epsilon)$ or $A_+(\boldsymbol{k},\epsilon), A_-(\boldsymbol{k},\epsilon)$ are also used.

In terms of these functions (11.103) becomes

$$G(k,t) = \begin{cases} -\mathrm{i}\displaystyle\int_0^\infty A(k,\epsilon)\exp[-\mathrm{i}(\epsilon+\mu)t]\,\mathrm{d}\epsilon, & t>0, \\ \mathrm{i}\displaystyle\int_0^\infty B(k,\epsilon)\exp[\mathrm{i}(\epsilon-\mu)t]\,\mathrm{d}\epsilon, & t<0, \end{cases} \tag{11.106}$$

since

$$\int_0^\infty \delta(\epsilon-\epsilon_n^{N+1})\exp[-\mathrm{i}(\epsilon+\mu)t]\,\mathrm{d}\epsilon = \exp[-\mathrm{i}(\epsilon_n^{N+1}+\mu)t], \tag{11.107}$$

etc. It should be noted that ϵ_n^{N+1} and ϵ_n^{N-1} are necessarily positive, since each is an *excitation energy* (the energy of an excited state minus the ground-state energy), so that it is only necessary to integrate from 0 to ∞.

Before going on it might be useful to consider what the functions $A(k,\epsilon)$, $B(k,\epsilon)$ become in the case of *non-interacting particles*. Here the matrix element $\langle\Psi_n^{N+1}|c_k^\dagger|\Psi_0\rangle$ in (11.104) vanishes unless $k > k_\mathrm{F}$ and Ψ_n^{N+1} is obtained from Ψ_0 by adding a single electron in the state k, so that

$$\epsilon_n^{N+1} = \epsilon_k - \epsilon_\mathrm{F} = \epsilon_k - \mu, \tag{11.108}$$

where $\epsilon_k = k^2/2m$ as usual. Also, the matrix element $\langle\Psi_n^{N-1}|c_k|\Psi_0\rangle$ in (11.105) vanishes unless $k < k_\mathrm{F}$ and Ψ_n^{N-1} is obtained from Ψ_0 by subtracting a single electron in the state k, so that

$$\epsilon_n^{N-1} = \epsilon_\mathrm{F} - \epsilon_k = \mu - \epsilon_k. \tag{11.109}$$

It follows that, for non-interacting particles,

$$A(k,\epsilon) = (1-n_k)\delta(\epsilon-\epsilon_k+\mu), \tag{11.110}$$

$$B(k,\epsilon) = n_k\delta(\epsilon+\epsilon_k-\mu), \tag{11.111}$$

where n_k is given by (11.70).

For interacting particles $c_k^\dagger\Psi_0$ is not an eigenfunction of the $(N+1)$-electron system. It can, however, be expressed as an infinite series of those Ψ_n^{N+1} with total momentum k. These Ψ_n^{N+1} will clearly be spread over a range of energies, so that A and (similarly) B will not be simple δ-functions as in (11.110) and (11.111). None the less, provided the interaction is not too strong, it is not unreasonable to expect A and B to be peaked near $\epsilon = \epsilon_k - \mu$ and $\epsilon = \mu - \epsilon_k$, respectively. This has an important bearing upon the general theory of *quasi-particles*, but we do not wish to elaborate this theory here as it is both mathematically more complicated and conceptually less rigorous than the intended scope of the present work.

Now, the Fourier transform of $G(k,t)$ with respect to t is, using (11.106),

$$\begin{aligned} G(k,\epsilon) &= \int_{-\infty}^\infty \exp(\mathrm{i}\epsilon t)G(k,t)\,\mathrm{d}t \\ &= -\mathrm{i}\int_0^\infty \exp(\mathrm{i}\epsilon t)\,\mathrm{d}t\int_0^\infty A(k,\epsilon')\exp[-\mathrm{i}(\epsilon'+\mu)t]\,\mathrm{d}\epsilon' \\ &\quad +\mathrm{i}\int_{-\infty}^0 \exp(\mathrm{i}\epsilon t)\,\mathrm{d}t\int_0^\infty B(k,\epsilon')\exp[\mathrm{i}(\epsilon'-\mu)t]\,\mathrm{d}\epsilon'. \end{aligned} \tag{11.112}$$

Inverting the order of integration, and changing the sign of t in the second term, gives

$$\begin{aligned} G(k,\epsilon) &= -\mathrm{i}\int_0^\infty A(k,\epsilon')\,\mathrm{d}\epsilon'\int_0^\infty \exp[\mathrm{i}(\epsilon-\epsilon'-\mu)t]\,\mathrm{d}t \\ &\quad +\mathrm{i}\int_0^\infty B(k,\epsilon')\,\mathrm{d}\epsilon'\int_0^\infty \exp[-\mathrm{i}(\epsilon+\epsilon'-\mu)t]\,\mathrm{d}t. \end{aligned} \tag{11.113}$$

The integrals with respect to t do not converge as they stand. To overcome this difficulty, as in previous sections, we include a factor $\mathrm{e}^{-\eta t}$, with η positive, in the integrands and ultimately take the limit as $\eta \to 0$. Thus we write

$$\begin{aligned} G(k,\epsilon) &= \lim_{\eta\to+0}\Big[-\mathrm{i}\int_0^\infty A(k,\epsilon')\,\mathrm{d}\epsilon'\int_0^\infty \exp[\mathrm{i}(\epsilon-\epsilon'-\mu+\mathrm{i}\eta)t]\,\mathrm{d}t \\ &\quad +\mathrm{i}\int_0^\infty B(k,\epsilon')\,\mathrm{d}\epsilon'\int_0^\infty \exp[-\mathrm{i}(\epsilon+\epsilon'-\mu-\mathrm{i}\eta)t]\,\mathrm{d}t\Big] \\ &= \lim_{\eta\to+0}\left[\int_0^\infty \frac{A(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon-\epsilon'-\mu+\mathrm{i}\eta} + \int_0^\infty \frac{B(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon+\epsilon'-\mu-\mathrm{i}\eta}\right]. \end{aligned} \tag{11.114}$$

This is the so-called *Lehmann representation* of $G(k,\epsilon)$, from which several properties of $G(k,\epsilon)$ can be deduced.

We note in passing that insertion of (11.110) and (11.111) in (11.114) gives immediately, for non-interacting particles,

$$G_0(k,\epsilon) = \lim_{\eta\to+0}\left[\frac{1-n_k}{\epsilon-\epsilon_k+\mathrm{i}\eta}+\frac{n_k}{\epsilon-\epsilon_k-\mathrm{i}\eta}\right], \tag{11.115}$$

which agrees with (11.84).

We shall now obtain some properties of $G(k,\epsilon)$ from eq. (11.114). First, using the formula given in Appendix III, eq. (III.10), we find

$$\begin{aligned}G(k,\epsilon) &= \mathscr{P}\int_0^\infty \frac{A(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon-\epsilon'-\mu}+\mathscr{P}\int_0^\infty \frac{B(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon+\epsilon'-\mu}\\ &\quad -\mathrm{i}\pi A(k,\epsilon-\mu)+\mathrm{i}\pi B(k,\mu-\epsilon),\end{aligned} \tag{11.116}$$

where $\mathscr{P}$ denotes the Cauchy principal value of the integral.

Now, according to their definitions, (11.104) and (11.105), $A(k,\epsilon)$ and $B(k,\epsilon)$ are both real. Therefore the *imaginary part* of $G(k,\epsilon)$ is

$$\operatorname{Im} G(k,\epsilon) = -\pi A(k,\epsilon-\mu)+\pi B(k,\mu-\epsilon). \tag{11.117}$$

Also, as we have noted before, ϵ_n^{N+1} and ϵ_n^{N-1} are necessarily positive, being excitation energies, so that $A(k,\epsilon)$ and $B(k,\epsilon)$ both vanish for $\epsilon < 0$. Hence

$$\operatorname{Im} G(k,\epsilon) = \begin{cases}-\pi A(k,\epsilon-\mu), & \text{if } \epsilon > \mu,\\ \pi B(k,\mu-\epsilon), & \text{if } \epsilon < \mu.\end{cases} \tag{11.118}$$

Since, again from their definitions, both A and B are either positive or zero, it follows that *the imaginary part of* $G(k,\epsilon)$ *changes sign at* $\epsilon=\mu$. Clearly, if A and B do not both vanish at $\epsilon=\mu$, $\operatorname{Im} G(k,\epsilon)$ has a discontinuity there.

The *real part* of $G(k,\epsilon)$ is, from eqs. (11.116) and (11.118),

$$\begin{aligned}\operatorname{Re} G(k,\epsilon) &= \mathscr{P}\int_0^\infty \frac{A(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon-\epsilon'-\mu}+\mathscr{P}\int_0^\infty \frac{B(k,\epsilon')\,\mathrm{d}\epsilon'}{\epsilon+\epsilon'-\mu}\\ &= -\frac{1}{\pi}\mathscr{P}\int_0^\infty \frac{\operatorname{Im} G(k,\epsilon'+\mu)\,\mathrm{d}\epsilon'}{\epsilon-\epsilon'-\mu}+\frac{1}{\pi}\mathscr{P}\int_0^\infty \frac{\operatorname{Im} G(k,\mu-\epsilon')\,\mathrm{d}\epsilon'}{\epsilon+\epsilon'-\mu}.\end{aligned} \tag{11.119}$$

Substituting $\omega=\epsilon'+\mu$ in the first integral, and $\omega=\mu-\epsilon'$ in the second, we obtain

$$\operatorname{Re} G(k,\epsilon) = \frac{1}{\pi}\mathscr{P}\int_\mu^\infty \frac{\operatorname{Im} G(k,\omega)\,\mathrm{d}\omega}{\omega-\epsilon}-\frac{1}{\pi}\mathscr{P}\int_{-\infty}^\mu \frac{\operatorname{Im} G(k,\omega)\,\mathrm{d}\omega}{\omega-\epsilon}. \tag{11.120}$$

This relation between the real and imaginary parts of $G(k,\epsilon)$ is called a *dispersion relation*, or *Kramers–Kronig relation*, because of its similarity to expressions obtained by Kramers and Kronig in their work on optical dispersion.

11.9. The Expectation Value of the Number Operator and the Existence of a Fermi Surface

As pointed out in §3.4, in the ground state of a free-electron gas, when the Coulomb interaction is neglected, all electrons occupy states lying within a sphere of radius k_F in k-space. This sphere is the *Fermi surface* for the non-interacting system.

Another way of stating this is given in eq. (11.70), which may be written more explicitly as

$$\langle\Phi_0|n_k|\Phi_0\rangle = \begin{cases}1, & \text{if } k<k_\mathrm{F},\\ 0, & \text{if } k>k_\mathrm{F},\end{cases} \tag{11.121}$$

where n_k is the *number operator* of the state k, given by

$$n_k = c_k^\dagger c_k \tag{11.122}$$

(*N.B.* k implies $k\sigma$) as in eq. (2.57). Eq. (11.121) thus gives the *expectation value of the number operator* of the state k in Φ_0, which is a single determinant.

When the interaction of the electrons is taken into account, the ground-state wave function Ψ_0 becomes an infinite sum of determinants, some of which will contain the state k and some of which will not. The expectation value of n_k will therefore not be simply 0 or 1, as in the non-interacting case, but some number lying in between.

Whether the system is interacting or not, the expectation value of n_k in the ground state Ψ_0 is given by

$$\begin{aligned}\langle\Psi_0|n_k|\Psi_0\rangle &= \langle\Psi_0|c_k^\dagger c_k|\Psi_0\rangle\\ &= -\mathrm{i}\lim_{t\to-0} G(k,t).\end{aligned} \tag{11.123}$$

This follows from the definition of $G(\boldsymbol{k},t)$ given in eq. (11.66), the restriction to negative values of t being to preserve the order of $c_k^\dagger c_k$. That this leads immediately to (11.121) in the non-interacting case ($G = G_0$) is obtained from eq. (11.76).

The sharp drop in $\langle\Phi_0|n_k|\Phi_0\rangle$ from 1 to 0 at $k = k_F$, as shown in fig. 11.1(a), unequivocally defines the Fermi surface for the non-interacting case. For interacting electrons, however, one might have expected this discontinuity to be smoothed out, so that a well-defined Fermi surface would no longer exist. An analysis† of the situation using Green functions reveals, however, that a discontinuity at $k = k_F$ still exists for a normal system (that is, for a non-superconducting system) when the electrons are interacting. This discontinuity of $\langle\Psi_0|n_k|\Psi_0\rangle$, shown schematically in fig. 11.1(b), is not as deep as in the non-interacting case, but is, none the less, sufficient to define the Fermi surface. It is sometimes said that the interacting system retains a 'memory' of the non-interacting system.

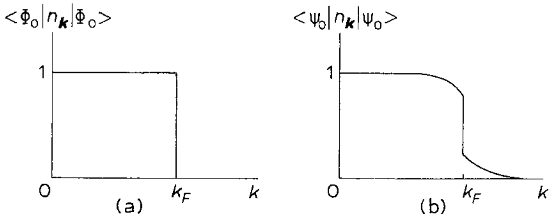


Fig. 11.1. Expectation value of the number operator n_k for (a) the non-interacting ground state Φ_0, (b) the interacting ground state Ψ_0 (schematic). The distribution is isotropic in both cases.

The expectation value of n_k may also be usefully expressed in terms of $G(\boldsymbol{k},\epsilon)$, as we shall now demonstrate. From eq. (11.69) we have

$$G(\boldsymbol{k},t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathrm{i}\epsilon t)\, G(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon, \qquad (11.124)$$

and this may be written as a contour integral in the complex ϵ-plane, thus:

$$G(\boldsymbol{k},t) = \frac{1}{2\pi}\int_{\mathrm{C}} \exp(-\mathrm{i}\epsilon t)\, G(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon, \qquad (11.125)$$

where, for $t < 0$, the contour C consists of the real axis and an infinite semicircle in the upper half-plane. This is because, on a semicircle of radius R in the upper half-plane,

$$\exp(-\mathrm{i}\epsilon t) = \exp[-\mathrm{i}tR(\cos\theta + \mathrm{i}\sin\theta)]$$
$$= \exp(-\mathrm{i}tR\cos\theta)\exp(tR\sin\theta),$$

and, since $\sin\theta$ is positive and t negative, this tends to zero as R tends to infinity. The contour integral is therefore equal to the integral along the real axis. Inserting (11.125) in (11.123), and taking the limit as $t \to -0$, we find

$$\langle\Psi_0|n_k|\Psi_0\rangle = -\frac{\mathrm{i}}{2\pi}\int_{\mathrm{C}} G(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon. \qquad (11.126)$$

This will be used in the following section. Meanwhile, it is easy to verify that it leads to the correct result for the non-interacting case. From (11.84), we have

$$G_0(\boldsymbol{k},\epsilon) = \lim_{\eta\to+0}\frac{1}{\epsilon - \epsilon_k + \mathrm{i}\eta}, \quad \text{if} \quad k > k_F, \qquad (11.127)$$

and the function $(\epsilon - \epsilon_k + \mathrm{i}\eta)^{-1}$ has no pole inside the contour C. Hence

$$\langle\Phi_0|n_k|\Phi_0\rangle = 0, \quad \text{if} \quad k > k_F. \qquad (11.128)$$

Also,

$$G_0(\boldsymbol{k},\epsilon) = \lim_{\eta\to+0}\frac{1}{\epsilon - \epsilon_k - \mathrm{i}\eta}, \quad \text{if} \quad k < k_F, \qquad (11.129)$$

and $(\epsilon - \epsilon_k - \mathrm{i}\eta)^{-1}$ has a simple pole at $\epsilon = \epsilon_k + \mathrm{i}\eta$ inside C, the residue being unity. Hence

$$\langle\Phi_0|n_k|\Phi_0\rangle = 2\pi\mathrm{i}\left(-\frac{\mathrm{i}}{2\pi}\right) = 1, \quad \text{if} \quad k < k_F. \qquad (11.130)$$

11.10. The Energy of the Ground State

In this section we shall obtain an expression for the energy of the ground state (and hence the correlation energy) of an interacting free-electron gas in terms of $G(\boldsymbol{k},\epsilon)$.

The Hamiltonian is

$$H = H_0 + H', \qquad (11.131)$$

---

† The mathematics is rather complicated and we will not include it. See, for example, Roman, P., 1965, *Advanced Quantum Theory* (Addison Wesley, Reading, Mass.) p. 464.

where†

$$H_0 = \sum_k \epsilon_k c_k^\dagger c_k \tag{11.132}$$

and

$$H' = \tfrac{1}{2} \sum_{k_1 k_2 k_3 k_4} \langle \boldsymbol{k}_1 \boldsymbol{k}_2 | v | \boldsymbol{k}_3 \boldsymbol{k}_4 \rangle c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3}. \tag{11.133}$$

Now, using the anticommutation relations, eqs. (2.53)–(2.55), we have

$$\begin{aligned}\sum_k c_k^\dagger [H_0, c_k] &= \sum_k (c_k^\dagger H_0 c_k - c_k^\dagger c_k H_0) \\ &= \sum_k \sum_{k'} \epsilon_{k'} (c_k^\dagger c_{k'}^\dagger c_{k'} c_k - c_k^\dagger c_k c_{k'}^\dagger c_{k'}) \\ &= \sum_k \sum_{k'} \epsilon_{k'} [-c_k^\dagger c_{k'}^\dagger c_k c_{k'} + c_k^\dagger (c_{k'}^\dagger c_k - \delta_{kk'}) c_{k'}] \\ &= -\sum_k \epsilon_k c_k^\dagger c_k = -H_0.\end{aligned} \tag{11.134}$$

Also,

$$\begin{aligned}\sum_k c_k^\dagger [H', c_k] = \tfrac{1}{2} \sum_k \sum_{k_1 k_2 k_3 k_4} &\langle \boldsymbol{k}_1 \boldsymbol{k}_2 | v | \boldsymbol{k}_3 \boldsymbol{k}_4 \rangle \\ &\times (c_k^\dagger c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3} c_k - c_k^\dagger c_k c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3}),\end{aligned} \tag{11.135}$$

and

$$\begin{aligned}c_k^\dagger c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3} c_k &= c_k^\dagger c_{k_1}^\dagger c_{k_2}^\dagger c_k c_{k_4} c_{k_3} \\ &= c_k^\dagger c_{k_1}^\dagger (\delta_{kk_2} - c_k c_{k_2}^\dagger) c_{k_4} c_{k_3} \\ &= c_k^\dagger c_{k_1}^\dagger c_{k_4} c_{k_3} \delta_{kk_2} - c_k^\dagger (\delta_{kk_1} - c_k c_{k_1}^\dagger) c_{k_2}^\dagger c_{k_4} c_{k_3} \\ &= c_k^\dagger c_{k_1}^\dagger c_{k_4} c_{k_3} \delta_{kk_2} - c_k^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3} \delta_{kk_1} + c_k^\dagger c_k c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3}.\end{aligned} \tag{11.136}$$

Substitution of (11.136) in (11.135) gives

$$\begin{aligned}\sum_k c_k^\dagger [H', c_k] &= \tfrac{1}{2} \sum_{k_1 k_2 k_3 k_4} \langle \boldsymbol{k}_1 \boldsymbol{k}_2 | v | \boldsymbol{k}_3 \boldsymbol{k}_4 \rangle \\ &\qquad \times (c_{k_2}^\dagger c_{k_1}^\dagger c_{k_4} c_{k_3} - c_{k_1}^\dagger c_{k_2}^\dagger c_{k_4} c_{k_3}) \\ &= -2H'.\end{aligned} \tag{11.137}$$

It follows from (11.134) and (11.137) that

$$\sum_k \langle \Psi_0 | c_k^\dagger [H, c_k] | \Psi_0 \rangle = -\langle \Psi_0 | H_0 + 2H' | \Psi_0 \rangle, \tag{11.138}$$

---

† These expressions are simplified forms of those given in eqs. (3.51) and (3.59). It must be remembered that, in the present chapter, the subscript k stands for $k\sigma$ and a sum over k also includes a sum over the two values of the spin index. Here, since we are taking $\hbar = 1$, $\epsilon_k = k^2/2m$.

and the energy of the ground state is thus

$$\begin{aligned}E = \langle \Psi_0 | H | \Psi_0 \rangle &= \tfrac{1}{2} \langle \Psi_0 | 2H_0 + 2H' | \Psi_0 \rangle \\ &= \tfrac{1}{2} \langle \Psi_0 | H_0 | \Psi_0 \rangle - \tfrac{1}{2} \sum_k \langle \Psi_0 | c_k^\dagger [H, c_k] | \Psi_0 \rangle.\end{aligned} \tag{11.139}$$

We are now in a position to express this in terms of $G(\boldsymbol{k}, \epsilon)$. First, from eqs. (11.122), (11.126) and (11.132), we have

$$\begin{aligned}\langle \Psi_0 | H_0 | \Psi_0 \rangle &= \sum_k \epsilon_k \langle \Psi_0 | n_k | \Psi_0 \rangle \\ &= -\frac{\mathrm{i}}{2\pi} \sum_k \epsilon_k \int_C G(\boldsymbol{k}, \epsilon)\, \mathrm{d}\epsilon.\end{aligned} \tag{11.140}$$

Now, the equation of motion of the Heisenberg operator $c_k(t)$ is, from eq. (5.15),

$$[H, c_k(t)] = -\mathrm{i} \frac{\mathrm{d}c_k(t)}{\mathrm{d}t}. \tag{11.141}$$

Therefore

$$\begin{aligned}\langle \Psi_0 | c_k^\dagger [H, c_k(t)] | \Psi_0 \rangle &= -\mathrm{i} \left\langle \Psi_0 \middle| c_k^\dagger \frac{\mathrm{d}c_k(t)}{\mathrm{d}t} \middle| \Psi_0 \right\rangle \\ &= -\frac{\mathrm{d}}{\mathrm{d}t} G(\boldsymbol{k}, t), \quad \text{for} \quad t < 0,\end{aligned} \tag{11.142}$$

and it follows from this and eq. (11.125) that

$$\begin{aligned}\langle \Psi_0 | c_k^\dagger [H, c_k] | \Psi_0 \rangle &= -\lim_{t \to -0} \frac{\mathrm{d}}{\mathrm{d}t} G(\boldsymbol{k}, t) \\ &= \frac{\mathrm{i}}{2\pi} \int_C G(\boldsymbol{k}, \epsilon) \epsilon\, \mathrm{d}\epsilon.\end{aligned} \tag{11.143}$$

Substituting (11.140) and (11.143) in (11.139) gives

$$E = -\frac{\mathrm{i}}{4\pi} \sum_k \int_C (\epsilon_k + \epsilon) G(\boldsymbol{k}, \epsilon)\, \mathrm{d}\epsilon. \tag{11.144}$$

Writing the sum as an integral in the usual way, *and multiplying by 2 to account for the sum over spins* (see §11.3), we have, finally,

$$E = -\frac{\mathrm{i}\Omega}{(2\pi)^4} \int \mathrm{d}\boldsymbol{k} \int_C (\epsilon_k + \epsilon) G(\boldsymbol{k}, \epsilon)\, \mathrm{d}\epsilon, \tag{11.145}$$

with $\epsilon_k = k^2/2m$ for free electrons.

This is an exact expression for the ground-state energy. It will be noted that the explicit form of the Coulomb interaction has not been used in its derivation, so that it is true for *any* two-particle interaction – the particular form of the interaction only affects $G(\boldsymbol{k},\epsilon)$. It is easy to verify, for example, that it is true for non-interacting particles, when $G(\boldsymbol{k},\epsilon)$ reduces to $G_0(\boldsymbol{k},\epsilon)$, as given in eq. (11.84). We have

$$(\epsilon_k+\epsilon)G_0(\boldsymbol{k},\epsilon) = \lim_{\eta\to+0}\frac{\epsilon_k+\epsilon}{\epsilon-\epsilon_k+\mathrm{i}\eta}, \quad \text{if} \quad k > k_{\mathrm{F}}, \tag{11.146}$$

and the function $(\epsilon_k+\epsilon)(\epsilon-\epsilon_k+\mathrm{i}\eta)^{-1}$ has no pole inside the contour C. However,

$$(\epsilon_k+\epsilon)G_0(\boldsymbol{k},\epsilon) = \lim_{\eta\to+0}\frac{\epsilon_k+\epsilon}{\epsilon-\epsilon_k-\mathrm{i}\eta}, \quad \text{if} \quad k < k_{\mathrm{F}}, \tag{11.146a}$$

and $(\epsilon_k+\epsilon)(\epsilon-\epsilon_k-\mathrm{i}\eta)^{-1}$ has a simple pole at $\epsilon=\epsilon_k+\mathrm{i}\eta$ inside C, the residue being $2\epsilon_k+\mathrm{i}\eta$ or $2\epsilon_k$ when the limit is taken. The ground-state energy of the non-interacting system is thus

$$\begin{aligned} E_0 &= -\frac{\mathrm{i}\Omega}{(2\pi)^4}\int \mathrm{d}\boldsymbol{k}\int_{\mathrm{C}}(\epsilon_k+\epsilon)G_0(\boldsymbol{k},\epsilon)\,\mathrm{d}\epsilon \\ &= -\frac{\mathrm{i}\Omega}{(2\pi)^4}\int_0^{k_{\mathrm{F}}} 4\pi k^2(2\pi\mathrm{i})(2\epsilon_k)\,\mathrm{d}k \\ &= \frac{\Omega}{10\pi^2 m}k_{\mathrm{F}}^5, \end{aligned} \tag{11.147}$$

which is the value found by elementary means.

Formula (11.145) is, in fact, of even wider generality than we have indicated. It would, in principle, be applicable to a metallic system for which the one-electron functions are the usual Bloch functions (see §7.3), provided that $G(\boldsymbol{k},t)$ is *defined by* (11.66). The one-electron energy ϵ_k would, however, no longer be $k^2/2m$, and its accurate calculation, for all $\boldsymbol{k}$, as well as that of $G(\boldsymbol{k},\epsilon)$, would be prohibitively laborious. Even for interacting free electrons the calculation of $G(\boldsymbol{k},\epsilon)$ is extremely difficult (precisely as difficult, in fact, as the calculation of the correlation energy described in ch. 9). In the remainder of the book we shall give a brief outline of two methods of calculating Green functions which are often used, at least for idealized systems or with the aid of approximations.

11.11. The Calculation of Green Functions by Perturbation Theory

The most popular method† of calculating Green functions is by means of a perturbation expansion involving the use of Feynman graphs. The treatment is similar to that of the ground-state energy described in detail in earlier chapters. We shall therefore confine ourselves to the initial stages of the calculation, pointing out some small differences from our previous method.

We again make use of the adiabatic hypothesis discussed in ch. 6. Here, however, since we shall be integrating over both negative and positive values of t, we introduce an exponential factor $\mathrm{e}^{-\alpha|t|}$ into the interaction term, where α is again a small positive constant ultimately tending to zero. We thus consider the time-dependent Hamiltonian

$$H(t) = H_0 + H'(t) = H_0 + \mathrm{e}^{-\alpha|t|}H', \tag{11.148}$$

which reduces to the ordinary time-independent Hamiltonian (11.131) when $t=0$ and to the non-interacting Hamiltonian H_0 when $t\to\pm\infty$. As before, we treat $H'(t)$ as a perturbation.

Although until now, in the present chapter, we have expressed the operators in the Heisenberg picture, here, as in chs. 6 and 7, it is more convenient to use the interaction picture. First, we wish to show that the representations $O_{\mathrm{H}}(t)$ and $O_{\mathrm{I}}(t)$ of any operator O in the Heisenberg and interaction pictures, respectively, are related by the equation

$$O_{\mathrm{H}}(t) = U^{\dagger}(t,0)O_{\mathrm{I}}(t)U(t,0), \tag{11.149}$$

where $U(t,t')$ is the unitary time-development operator defined by eq. (5.24), namely,

$$\Psi_{\mathrm{I}}(t) = U(t,t')\Psi_{\mathrm{I}}(t'). \tag{11.150}$$

For any wave function $\Psi_n(t)$ we have

$$\Psi_{n\mathrm{I}}(t) = U(t,0)\Psi_{n\mathrm{I}}(0) = U(t,0)\Psi_{n\mathrm{H}}, \tag{11.151}$$

since, as shown by eqs. (5.12) and (6.12), at $t=0$ any wave function is the same in all pictures. Also, we know (see Appendix II) that the matrix

---

† The whole approach to Green function techniques is often made through what is essentially perturbation theory. See, for example, Mattuck R. D., 1967, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill, New York), for an interesting and readable account of this alternative approach.

element of an operator $O(t)$ between any pair of wave functions $\Psi_n(t)$, $\Psi_m(t)$ must have the same value in all pictures, so that, using (11.151),

$$\begin{aligned}\langle\Psi_{mI}(t)|O_I(t)|\Psi_{nI}(t)\rangle &= \langle\Psi_{mH}|U^\dagger(t,0)O_I(t)U(t,0)|\Psi_{nH}\rangle\\ &= \langle\Psi_{mH}|O_H(t)|\Psi_{nH}\rangle.\end{aligned}\tag{11.152}$$

This is true for all Ψ_{nH},Ψ_{mH}, so that (11.149) follows[†].

Now the Green function $G(\boldsymbol{k},t)$ is defined by eq. (11.66), for $t > 0$, as

$$G(\boldsymbol{k},t) = -i\langle\Psi_0|c_{kH}(t)c^\dagger_{kH}(0)|\Psi_0\rangle,\tag{11.153}$$

where Ψ_0 is the ground state at $t = 0$, the same in all pictures. Using (11.149) we have

$$c^\dagger_{kH}(0) = c^\dagger_{kI}(0),\tag{11.154}$$

and

$$c_{kH}(t) = U^\dagger(t,0)c_{kI}(t)U(t,0).\tag{11.155}$$

Also, from eq. (6.14),

$$\Psi_0 = U(0,-\infty)\Phi_0,\tag{11.156}$$

where Φ_0 is the non-interacting ground state (i.e. the ground state when $t \to -\infty$). Hence, for $t > 0$,

$$\begin{aligned}G(\boldsymbol{k},t) = -i\langle\Phi_0|U^\dagger(0,-\infty)U^\dagger(t,0)c_{kI}(t)&\\ \times U(t,0)c^\dagger_{kI}(0)U(0,-\infty)|\Phi_0\rangle&.\end{aligned}\tag{11.157}$$

We now define the so-called *S-matrix* by

$$S = U(\infty,-\infty),\tag{11.158}$$

and, according to the properties of $U(t,t')$ given in eqs. (5.27), (5.28), (5.29) and (5.31), it follows that

$$S = U(\infty,0)U(0,-\infty)\tag{11.159}$$

[†] It is easy to verify that this is correct when H does not depend explicitly upon the time (i.e. for a conservative system). In this case $U(t,t')$ has the form (see eq. (5.36))

$$U(t,t') = \exp(iH_0t)\exp[-iH(t-t')]\exp(-iH_0t'),$$

so that (11.149) gives

$$O_H(t) = \exp[i(H-H_0)t]O_I(t)\exp[-i(H-H_0)t] = \exp(iH't)O_I(t)\exp(-iH't),$$

which is correct by (5.13) and (5.20). In the present case the Hamiltonian (11.148) does depend explicitly upon the time, but (11.149) still holds.

and

$$S^\dagger = U(-\infty,0)U(0,\infty).\tag{11.160}$$

Thus

$$\begin{aligned}U(-\infty,0) &= U^\dagger(0,-\infty)\\ &= S^\dagger U^\dagger(0,\infty) = S^\dagger U(\infty,0),\end{aligned}\tag{11.161}$$

and

$$\begin{aligned}U^\dagger(0,-\infty)U^\dagger(t,0) &= S^\dagger U(\infty,0)U(0,t)\\ &= S^\dagger U(\infty,t).\end{aligned}\tag{11.162}$$

Substituting in (11.157) gives,

for $t > 0$,

$$\begin{aligned}G(\boldsymbol{k},t) = -i\langle\Phi_0|S^\dagger U(\infty,t)c_{kI}(t)&\\ \times U(t,0)c^\dagger_{kI}(0)U(0,-\infty)|\Phi_0\rangle&.\end{aligned}\tag{11.163}$$

In ch. 6 we did not consider positive values of t. However, we may assume that switching on the interaction slowly when $t \to -\infty$ and ultimately switching it off when $t \to +\infty$ (which is accomplished by the factor $e^{-\alpha|t|}$) will cause the ground state Φ_0 finally to return to Φ_0, at least within a multiplying factor of modulus unity. In other words, we may assume that

$$S\Phi_0 = e^{i\phi}\Phi_0,\tag{11.164}$$

which gives

$$e^{i\phi} = \langle\Phi_0|S|\Phi_0\rangle.\tag{11.165}$$

Thus, for $t > 0$,

$$G(\boldsymbol{k},t) = -\frac{i\langle\Phi_0|U(\infty,t)c_{kI}(t)U(t,0)c^\dagger_{kI}(0)U(0,-\infty)|\Phi_0\rangle}{\langle\Phi_0|S|\Phi_0\rangle}.\tag{11.166}$$

We recall from eq. (5.52) that

$$U(t,t') = P\exp\left[-i\int_{t'}^{t} H_I(t_1)\,dt_1\right],\tag{11.167}$$

where in the present case (cf. eq. (7.17)) $H_I(t)$, which is what $H'e^{-\alpha|t|}$ becomes in the interaction picture, is given by

$$H_I(t) = \tfrac{1}{2}\sum_{k_1k_2k_3k_4}\langle\boldsymbol{k}_1\boldsymbol{k}_2|v|\boldsymbol{k}_3\boldsymbol{k}_4\rangle c^\dagger_{k_1I}(t)c^\dagger_{k_2I}(t)c_{k_4I}(t)c_{k_3I}(t)\,e^{-\alpha|t|}.\tag{11.168}$$

It should now be noticed that, since $H_{\mathrm{I}}(t)$ is the sum of terms, each containing the product of *four* (i.e. an even number of) creation and destruction operators, we must have

$$T\{H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)\} = \begin{cases} H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2), & \text{if} \quad t_1 > t_2, \\ H_{\mathrm{I}}(t_2)H_{\mathrm{I}}(t_1), & \text{if} \quad t_2 > t_1, \end{cases} \tag{11.169}$$

with the same sign in both cases. This is because the definition of T in eq. (11.11) applies to creation or destruction operators only, and inverting the order of $H_{\mathrm{I}}(t_1)H_{\mathrm{I}}(t_2)$ involves an *even* number of permutations of pairs of adjacent creation or destruction operators. Thus the operator T in this case is the same as the operator P, so that (11.167) can be written

$$U(t,t') = T\exp\left[-\mathrm{i}\int_{t'}^{t} H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1\right], \tag{11.170}$$

and, in particular,

$$S = T\exp\left[-\mathrm{i}\int_{-\infty}^{\infty} H_{\mathrm{I}}(t_1)\,\mathrm{d}t_1\right]. \tag{11.171}$$

Now, in (11.166) we have $t > 0$, and all times in the integrands of the expansion of $U(\infty,t)$ must therefore be $> t$, all those in $U(t,0)$ must be > 0, and all those in $U(0,-\infty)$ must be < 0. Also, from the properties of $U(t,t')$, we have

$$S = U(\infty,t)\,U(t,0)\,U(0,-\infty). \tag{11.172}$$

We may thus write

$$\begin{aligned} U(\infty,t)c_{kl}(t)U(t,0)c_{kl}^{\dagger}U(0,-\infty) &= T\{c_{kl}(t)c_{kl}^{\dagger}(0)\,U(\infty,t)\,U(t,0)\,U(0,-\infty)\} \\ &= T\{c_{kl}(t)c_{kl}^{\dagger}(0)S\}. \end{aligned} \tag{11.173}$$

Finally, then,

$$G(\boldsymbol{k},t) = -\frac{\mathrm{i}\langle\Phi_0|T\{c_{kl}(t)c_{kl}^{\dagger}(0)S\}|\Phi_0\rangle}{\langle\Phi_0|S|\Phi_0\rangle}. \tag{11.174}$$

Although we have assumed $t > 0$, it is easy to show that the result is exactly the same for $t < 0$. Clearly, since S depends upon α, (11.174) should strictly be written

$$G(\boldsymbol{k},t) = -\mathrm{i}\lim_{\alpha\to 0}\frac{\langle\Phi_0|T\{c_{kl}(t)c_{kl}^{\dagger}(0)S_\alpha\}|\Phi_0\rangle}{\langle\Phi_0|S_\alpha|\Phi_0\rangle}. \tag{11.175}$$

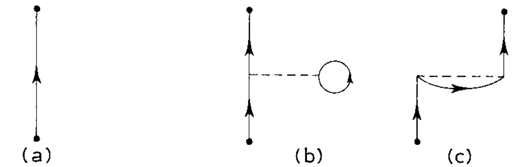
This equation gives the perturbation expansion of $G(\boldsymbol{k},t)$, just as eq. (6.16) gave the perturbation expansion of the ground-state energy. Written in full, it becomes

$$G(\boldsymbol{k},t) = -\mathrm{i}\lim_{\alpha\to 0}$$

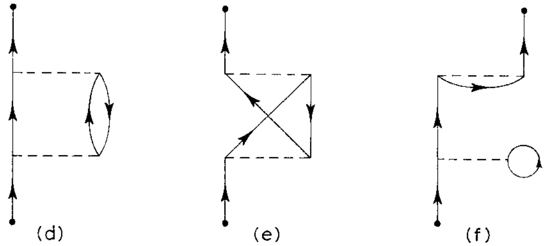
$$\times\frac{\displaystyle\sum_{n=0}^{\infty}\frac{(-\mathrm{i})^n}{n!}\int_{-\infty}^{\infty}\mathrm{d}t_1\ldots\int_{-\infty}^{\infty}\mathrm{d}t_n\langle\Phi_0|T\{c_{kl}(t)c_{kl}^{\dagger}(0)H_{\mathrm{I}}(t_1)\ldots H_{\mathrm{I}}(t_n)\}|\Phi_0\rangle}{\displaystyle\sum_{n=0}^{\infty}\frac{(-\mathrm{i})^n}{n!}\int_{-\infty}^{\infty}\mathrm{d}t_1\ldots\int_{-\infty}^{\infty}\mathrm{d}t_n\langle\Phi_0|T\{H_{\mathrm{I}}(t_1)\ldots H_{\mathrm{I}}(t_n)\}|\Phi_0\rangle}, \tag{11.176}$$

the terms with $n = 0$ being simply $\langle\Phi_0|T\{c_{kl}(t)c_{kl}^{\dagger}(0)\}|\Phi_0\rangle$ in the numerator and unity in the denominator.

As in the case of the ground-state energy, the terms of the numerator and denominator of (11.176) may be represented by Feynman graphs. In the case of the denominator the graphs are exactly the same as those described in ch. 7, since the different limits of integration in $U_\alpha(t,-\infty)$ and S_α do not affect the graphical analysis. In the case of the numerator the graphs are similar, but each has two free vertices (i.e. points not attached to an interaction line) at times 0 and t, respectively, with a single hole or particle line entering or leaving, these lines representing $c_{kl}^{\dagger}(0)$ and $c_{kl}(t)$. Some types of graphs of low orders are shown in fig. 11.2 (the free vertices are indicated by large dots). The zeroth-order graph clearly represents $G_0(\boldsymbol{k},t)$.



Zeroth order First order



Second order

Fig. 11.2. Some types of graphs representing terms in the numerator of $G(\mathbf{k}, t)$, eq. (11.176).

The linked graph theorem again applies, the net effect of it being that the denominator in eq. (11.176) may be ignored provided that only those terms represented by linked graphs are included in the numerator (all the graphs shown in fig. 11.2 are linked).

It should be noted that graphs may be drawn in many different ways – in fact, there are almost as many ways as there are writers on the subject. As a trivial example, the time axis is often made horizontal instead of vertical – figs. 11.3(a), (b), (c) show the second-order graphs of fig. 11.2 drawn in this way. In addition, however, many distortions, which none the less retain the essential topology of a graph, are in current use. For example, continuous hole and particle lines may be joined in one straight line, preserving the correct sense of the arrows, so that the interaction lines may have to be curved in order to connect up the proper vertices. Figs. 11.3(d), (e), (f) show the second-order graphs of fig. 11.2 drawn in this way. (*N.B.* Even when the interaction lines need not be curved they are often so drawn, as in fig. 11.3(d).) The significance of other conventions used in drawing graphs are best discovered from the texts in which they appear.

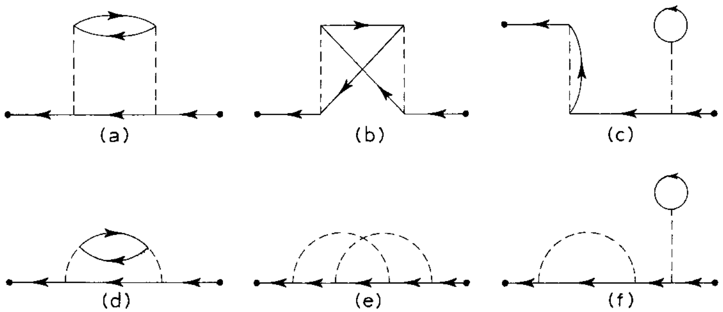


Fig. 11.3. Two alternative ways of drawing the second-order graphs of fig. 11.2.

Although we have been concerned here with $G(\boldsymbol{k},t)$, the function $G(\boldsymbol{k},\epsilon)$ can be expanded similarly, as indeed can the more general functions $G(\boldsymbol{k}t,\boldsymbol{k}'t')$ and $G(\boldsymbol{r}t,\boldsymbol{r}'t')$.

11.12. Equations of Motion and the Two-particle Green Function

Another way of attempting to calculate Green functions is by solving their equations of motion. This is by no means straightforward, however, for we shall see that the equation of motion of a one-particle Green function in general involves a two-particle Green function, and the equation of motion of the latter in general involves a three-particle Green function, and so on. The problem is therefore to solve an infinite set of coupled equations of motion. Clearly, this can only be done with the aid of approximations or for idealized systems.

In §11.7 we found the equation of motion of $G_0(\boldsymbol{k},t)$ for a system of non-interacting free particles. We shall, in this final section, derive the equation of motion of the Green function $G(\boldsymbol{x}t,\boldsymbol{x}'t')$, including spin, for a system of interacting free electrons (the general treatment remains valid for other systems, however – in particular, it is valid for interactions other than the Coulomb interaction).

We now return to the Heisenberg picture, as in the rest of this chapter apart from the previous section. According to §§2.8 and 11.2, the Hamiltonian (11.131) may be written

$$H = H_0 + H' = -\int \psi^\dagger(\boldsymbol{x},t)\frac{1}{2m}\nabla^2\psi(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}$$

$$+\tfrac{1}{2}\iint \psi^\dagger(\boldsymbol{x}_1,t)\psi^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\psi(\boldsymbol{x}_2,t)\psi(\boldsymbol{x}_1,t)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2, \qquad (11.177)$$

and, from (5.15), the equation of motion of the Heisenberg operator $\psi(\boldsymbol{x},t)$ is

$$[H,\psi(\boldsymbol{x},t)] = -\mathrm{i}\frac{\partial}{\partial t}\psi(\boldsymbol{x},t). \qquad (11.178)$$

Now,

$$[H_0,\psi(\boldsymbol{x},t)] = -\frac{1}{2m}\int [\psi^\dagger(\boldsymbol{x}',t)\nabla'^2\psi(\boldsymbol{x}',t),\psi(\boldsymbol{x},t)]\,\mathrm{d}\boldsymbol{x}'$$

$$= \frac{1}{2m}\int \{\psi^\dagger(\boldsymbol{x}',t),\psi(\boldsymbol{x},t)\}\nabla'^2\psi(\boldsymbol{x}',t)\,\mathrm{d}\boldsymbol{x}', \qquad (11.179)$$

since, with ∇'^2 operating upon $\psi(\boldsymbol{x}',t)$ only

$$[\psi^\dagger(\boldsymbol{x}',t)\nabla'^2\psi(\boldsymbol{x}',t),\psi(\boldsymbol{x},t)]$$

$$= \psi^\dagger(\boldsymbol{x}',t)\nabla'^2\psi(\boldsymbol{x}',t)\psi(\boldsymbol{x},t) - \psi(\boldsymbol{x},t)\psi^\dagger(\boldsymbol{x}',t)\nabla'^2\psi(\boldsymbol{x}',t)$$

$$= -\{\psi^\dagger(\boldsymbol{x}',t),\psi(\boldsymbol{x},t)\}\nabla'^2\psi(\boldsymbol{x}',t), \qquad (11.180)$$

where we have used the anticommutation relation

$$\{\psi(\boldsymbol{x}',t),\psi(\boldsymbol{x},t)\} = 0. \qquad (11.181)$$

Also, from eq. (11.10), we have

$$\{\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t),\boldsymbol{\psi}(\boldsymbol{x},t)\} = \delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{11.182}$$

Hence

$$\begin{aligned}[H_0,\boldsymbol{\psi}(\boldsymbol{x},t)] &= \frac{1}{2m}\int \delta(\boldsymbol{x}-\boldsymbol{x}')\nabla'^2\boldsymbol{\psi}(\boldsymbol{x}',t)\,\mathrm{d}\boldsymbol{x}' \\ &= \frac{1}{2m}\nabla^2\boldsymbol{\psi}(\boldsymbol{x},t). \end{aligned} \tag{11.183}$$

For the interaction term H', we have

$$\begin{aligned}[H',\boldsymbol{\psi}(\boldsymbol{x},t)] = \tfrac{1}{2}\iint\, &[\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t) \\ &\times v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t),\boldsymbol{\psi}(\boldsymbol{x},t)]\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2. \end{aligned} \tag{11.184}$$

Again using the anticommutation relations, we obtain

$$\begin{aligned}&[\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t),\boldsymbol{\psi}(\boldsymbol{x},t)] \\ &\quad= \boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x},t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\qquad - \boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\quad= -\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\qquad + \delta(\boldsymbol{x}_2-\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\qquad - \boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\quad= \delta(\boldsymbol{x}_2-\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t) \\ &\qquad - \delta(\boldsymbol{x}_1-\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t). \end{aligned} \tag{11.185}$$

Hence

$$\begin{aligned}[H',\boldsymbol{\psi}(\boldsymbol{x},t)] &= \tfrac{1}{2}\iint \delta(\boldsymbol{x}_2-\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &\quad - \tfrac{1}{2}\iint \delta(\boldsymbol{x}_1-\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2 \\ &= \tfrac{1}{2}\int \boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)v(\boldsymbol{x}_1,\boldsymbol{x})\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\,\mathrm{d}\boldsymbol{x}_1 \\ &\quad - \tfrac{1}{2}\int \boldsymbol{\psi}^\dagger(\boldsymbol{x}_2,t)v(\boldsymbol{x},\boldsymbol{x}_2)\boldsymbol{\psi}(\boldsymbol{x}_2,t)\boldsymbol{\psi}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}_2 \\ &= -\int v(\boldsymbol{x}_1,\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}_1, \end{aligned} \tag{11.186}$$

where we have used the fact that

$$v(\boldsymbol{x}_1,\boldsymbol{x}_2) = v(\boldsymbol{x}_2,\boldsymbol{x}_1). \tag{11.187}$$

Substituting (11.183) and (11.186) in (11.178) gives

$$\begin{aligned}\mathrm{i}\frac{\partial}{\partial t}\boldsymbol{\psi}(\boldsymbol{x},t) &= -\frac{1}{2m}\nabla^2\boldsymbol{\psi}(\boldsymbol{x},t) \\ &\quad + \int v(\boldsymbol{x}_1,\boldsymbol{x})\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}_1. \end{aligned} \tag{11.188}$$

Now, as in eq. (11.15), we may write

$$\begin{aligned}T\{\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\} &= \theta(t-t')\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t') \\ &\quad -\theta(t'-t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\boldsymbol{\psi}(\boldsymbol{x},t), \end{aligned} \tag{11.189}$$

so that, using (11.92),

$$\begin{aligned}&\frac{\partial}{\partial t}T\{\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\} \\ &\quad= \theta(t-t')\frac{\partial\boldsymbol{\psi}(\boldsymbol{x},t)}{\partial t}\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t') + \frac{\partial\theta(t-t')}{\partial t}\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t') \\ &\qquad -\theta(t'-t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\frac{\partial\boldsymbol{\psi}(\boldsymbol{x},t)}{\partial t} - \frac{\partial\theta(t'-t)}{\partial t}\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\boldsymbol{\psi}(\boldsymbol{x},t) \\ &\quad= T\left\{\frac{\partial\boldsymbol{\psi}(\boldsymbol{x},t)}{\partial t}\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\right\} + \{\boldsymbol{\psi}(\boldsymbol{x},t),\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\}\delta(t-t') \\ &\quad= T\left\{\frac{\partial\boldsymbol{\psi}(\boldsymbol{x},t)}{\partial t}\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\right\} + \delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'). \end{aligned} \tag{11.190}$$

Thus, using (11.188),

$$\begin{aligned}&\mathrm{i}\frac{\partial}{\partial t}\langle\Psi_0|T\{\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\}|\Psi_0\rangle \\ &\quad= \mathrm{i}\left\langle\Psi_0\left|T\left\{\frac{\partial\boldsymbol{\psi}(\boldsymbol{x},t)}{\partial t}\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\right\}\right|\Psi_0\right\rangle + \mathrm{i}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t') \\ &\quad= -\frac{1}{2m}\nabla^2\langle\Psi_0|T\{\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\}|\Psi_0\rangle \\ &\qquad + \left\langle\Psi_0\left|\int v(\boldsymbol{x}_1,\boldsymbol{x})T\{\boldsymbol{\psi}^\dagger(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x}_1,t)\boldsymbol{\psi}(\boldsymbol{x},t)\boldsymbol{\psi}^\dagger(\boldsymbol{x}',t')\}\,\mathrm{d}\boldsymbol{x}_1\right|\Psi_0\right\rangle \\ &\qquad + \mathrm{i}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t'), \end{aligned} \tag{11.191}$$

where ∇^2 acts upon $\psi(x,t)$ only.

In other words, from the definition (11.12),

$$i\frac{\partial}{\partial t}G(xt,x't') = -\frac{1}{2m}\nabla^2 G(xt,x't') - i\int v(x_1,x)\langle\Psi_0|T\{\psi^\dagger(x_1,t)\psi(x_1,t)\psi(x,t) \times \psi^\dagger(x',t')\}|\Psi_0\rangle\,dx_1 + \delta(x-x')\delta(t-t'). \tag{11.192}$$

We now define† the *two-particle Green function* by

$$K(x_1t_1,x_2t_2;x_3t_3,x_4t_4) = \langle\Psi_0|T\{\psi(x_1,t_1)\psi(x_2,t_2)\psi^\dagger(x_3,t_3)\psi^\dagger(x_4,t_4)\}|\Psi_0\rangle. \tag{11.193}$$

This may be used in (11.192) provided that we write $\psi^\dagger(x_1,t)$ as $\psi^\dagger(x_1,t_+)$, where t_+ is infinitesimally greater than t, in order to preserve the correct order of the operators; that is, we write

$$\begin{aligned}\langle\Psi_0|T\{\psi^\dagger(x_1,t)\psi(x_1,t)\psi(x,t)\psi^\dagger(x',t')\}|\Psi_0\rangle &= \langle\Psi_0|T\{\psi(x_1,t)\psi(x,t)\psi^\dagger(x_1,t_+)\psi^\dagger(x',t')\}|\Psi_0\rangle \\ &= K(x_1t,xt;x_1t_+,x't').\end{aligned} \tag{11.194}$$

The equation of motion of the one-particle Green function thus becomes

$$\left(i\frac{\partial}{\partial t}+\frac{1}{2m}\nabla^2\right)G(xt,x't') = \delta(x-x')\delta(t-t') - i\int v(x_1,x)K(x_1t,xt;x_1t_+,x't')\,dx_1. \tag{11.195}$$

If there is no interaction, so that $v = 0$, this reduces to the equation for $G_0(xt,x't')$, namely,

$$\left(i\frac{\partial}{\partial t}+\frac{1}{2m}\nabla^2\right)G_0(xt,x't') = \delta(x-x')\delta(t-t'), \tag{11.196}$$

which, apart from the inclusion of spin, is the same as eq. (10.109) for the time-dependent Green function of a single particle. This again justifies the name Green function. However, for an interacting system, we see that the equation of motion of the one-particle Green function depends upon the two-particle Green function. Similarly, the equation of motion of the latter may be shown to depend upon both the one-particle Green function and a three-particle Green function (defined by an obvious extension of (11.193)), and so on. We thus obtain an infinite set of coupled equations, as mentioned previously.

The total energy of the system may be expressed in terms of the one-particle and two-particle Green functions. However, as we have already seen, it can more usefully be expressed in terms of the one-particle Green function only. We shall therefore say no more about the two-particle Green function, except to remark that it determines the *collective motion* of the system, which we have already met in another context in ch. 4.

---

† It should be noted that sometimes a factor -1 or $-i$ is included in the definition.

Appendix I

HERMITIAN OPERATORS

The *adjoint* or *Hermitian conjugate* $A^\dagger$ of a linear‡ operator A is defined by the relation

$$\int f^* A g \,\mathrm{d}\tau = \int (A^\dagger f)^* g \,\mathrm{d}\tau, \tag{I.1}$$

where f and g are arbitrary functions (such that the integral exists, however). More compactly, we may write

$$\langle f|A|g\rangle = \langle A^\dagger f|g\rangle, \tag{I.2}$$

or, alternatively,

$$\langle f|A|g\rangle = \langle g|A^\dagger|f\rangle^*, \tag{I.3}$$

since

$$\int (A^\dagger f)^* g \,\mathrm{d}\tau = \left(\int g^* A^\dagger f \mathrm{d}\tau\right)^*. \tag{I.4}$$

In order to be able to find the Hermitian conjugate of any combination of operators we shall now derive certain properties of Hermitian conjugates.

(1) Hermitian conjugation is a reciprocal operation, that is,

$$(A^\dagger)^\dagger = A. \tag{I.5}$$

We have, from (I.3),

$$\langle f|A|g\rangle = \langle g|A^\dagger|f\rangle^* = \langle f|(A^\dagger)^\dagger|g\rangle, \tag{I.6}$$

---

‡ A linear operator A is one such that

$$A(c_1 f + c_2 g) = c_1 Af + c_2 Ag,$$

where c_1 and c_2 are constants.

and (I.5) follows, since f and g are arbitrary functions.

(2) If c is a number, then

$$c^\dagger = c^*, \tag{I.7}$$

since, again from (I.3),

$$\langle f|c^\dagger|g\rangle = \langle g|c|f\rangle^* = \langle f|c^*|g\rangle. \tag{I.8}$$

(3) For any two operators A and B,

$$(A+B)^\dagger = A^\dagger + B^\dagger, \tag{I.9}$$

since

$$\begin{aligned}\langle f|(A+B)^\dagger|g\rangle &= \langle g|A+B|f\rangle^* \\ &= \langle g|A|f\rangle^* + \langle g|B|f\rangle^* \\ &= \langle f|A^\dagger|g\rangle + \langle f|B^\dagger|g\rangle \\ &= \langle f|A^\dagger + B^\dagger|g\rangle.\end{aligned} \tag{I.10}$$

(4) For any two operators A and B,

$$(AB)^\dagger = B^\dagger A^\dagger \tag{I.11}$$

(note the reversal of order, which is important). We have, from (I.2) and (I.5),

$$\begin{aligned}\langle f|(AB)^\dagger|g\rangle &= \langle ABf|g\rangle \\ &= \langle Bf|A^\dagger|g\rangle = \langle f|B^\dagger A^\dagger|g\rangle.\end{aligned} \tag{I.12}$$

An operator A is said to be *self-adjoint* or *Hermitian* if

$$A = A^\dagger. \tag{I.13}$$

It follows from (I.3) that, for such an operator,

$$\langle f|A|g\rangle = \langle g|A|f\rangle^*. \tag{I.14}$$

Conversely, if (I.14) holds, then A is Hermitian, since f and g are arbitrary functions.

It is easily seen that the eigenvalues of a Hermitian operator are real. Suppose that

$$A\psi = \lambda\psi, \tag{I.15}$$

and A is Hermitian. Then, if ψ is normalized,

$$\lambda = \langle\psi|A|\psi\rangle = \langle\psi|A|\psi\rangle^* = \lambda^*, \tag{I.16}$$

from (I.14), so that the eigenvalue λ is real.

It is a postulate of quantum mechanics that an operator representing a physical quantity must be Hermitian, and its real eigenvalues represent the possible results of physical measurements of that quantity. In particular, it is easily shown that the Hamiltonian operator H for a given system is Hermitian, and its real eigenvalues are the energy levels of the system

Another important property is that the eigenfunctions of a Hermitian operator, belonging to different eigenvalues, are orthogonal. Suppose that

$$A\psi_1 = \lambda_1\psi_1, \tag{I.17}$$

$$A\psi_2 = \lambda_2\psi_2, \tag{I.18}$$

and $\lambda_1 \neq \lambda_2$. Then

$$\langle\psi_2|A|\psi_1\rangle = \lambda_1\langle\psi_2|\psi_1\rangle, \tag{I.19}$$

and

$$\langle\psi_1|A|\psi_2\rangle = \lambda_2\langle\psi_1|\psi_2\rangle. \tag{I.20}$$

Using (I.14) and (I.16), we thus find

$$\begin{aligned}\langle\psi_2|A|\psi_1\rangle &= \langle\psi_1|A|\psi_2\rangle^* \\ &= \lambda_2^*\langle\psi_1|\psi_2\rangle^* = \lambda_2\langle\psi_2|\psi_1\rangle = \lambda_1\langle\psi_2|\psi_1\rangle,\end{aligned} \tag{I.21}$$

which gives

$$(\lambda_1-\lambda_2)\langle\psi_2|\psi_1\rangle = 0. \tag{I.22}$$

Since $\lambda_1 \neq \lambda_2$, this implies that

$$\langle\psi_2|\psi_1\rangle = 0, \tag{I.23}$$

or ψ_1 and ψ_2 are orthogonal.

The final step is not valid if $\lambda_1 = \lambda_2$, that is, if ψ_1 and ψ_2 are degenerate. However, a set of degenerate eigenfunctions can always be orthogonalized (see, for example, WM p. 47), so that no generality is lost by assuming that the eigenfunctions of a Hermitian operator form an orthogonal set.

In conclusion, it may be useful to mention a simple property which is used many times throughout the book, particularly in ch. 11. If ψ is an eigenfunction of a Hermitian operator A, belonging to the eigenvalue λ, so that eq. (I.15) holds, and f is any other function, then, from (I.14) and (I.16),

$$\begin{aligned}\langle\psi|A|f\rangle &= \langle f|A|\psi\rangle^* = \lambda^*\langle f|\psi\rangle^* \\ &= \lambda\langle\psi|f\rangle.\end{aligned} \tag{I.24}$$

Appendix II

UNITARY OPERATORS AND TRANSFORMATIONS

An operator U is said to be *unitary* if

$$UU^\dagger = U^\dagger U = 1. \tag{II.1}$$

That is to say, for a unitary operator,

$$U^\dagger = U^{-1}. \tag{II.2}$$

The transformation of an operator A into an operator A', where

$$A' = UAU^\dagger, \tag{II.3}$$

or

$$A = U^\dagger A'U, \tag{II.4}$$

and U is a unitary operator, is called a *unitary transformation.*

Let us define the functions F and G in terms of the functions f and g, so that

$$F = Uf, \quad G = Ug, \tag{II.5}$$

or

$$f = U^\dagger F, \quad g = U^\dagger G. \tag{II.6}$$

The matrix element of the operator A with respect to the functions f and g is then, using eqs. (I.2) and (II.3),

$$\begin{aligned}\langle f|A|g\rangle &= \langle U^\dagger F|A|U^\dagger G\rangle \\ &= \langle F|UAU^\dagger|G\rangle = \langle F|A'|G\rangle.\end{aligned} \tag{II.7}$$

Thus the matrix elements of the operator A with respect to a set of functions $f, g, \ldots$, are the same as those of the transformed operator A'

with respect to the transformed functions $Uf, Ug, \ldots$. In particular, if ψ is an eigenfunction of A corresponding to the eigenvalue λ, then $U\psi$ is an eigenfunction of A' corresponding to the same eigenvalue. This can be seen directly, for

$$A\psi = \lambda\psi, \tag{II.8}$$

so that, from (II.4),

$$U^{\dagger}A'U\psi = \lambda\psi. \tag{II.9}$$

Hence

$$UU^{\dagger}A'U\psi = U\lambda\psi = \lambda U\psi, \tag{II.10}$$

or

$$A'(U\psi) = \lambda(U\psi). \tag{II.11}$$

Since we are essentially concerned with finding the eigenvalues of operators representing physical quantities, and of the Hamiltonian operator in particular, it is clear that we are permitted to make any unitary transformation which may prove to be convenient.

Appendix III

A USEFUL INTEGRAL FORMULA

It is often required, particularly in work involving Green functions, as seen in chs. 10 and 11, to evaluate an expression of the type

$$\lim_{\eta\to 0}\int_{-\infty}^{\infty}\frac{f(x)\,\mathrm{d}x}{x+\mathrm{i}\eta}.$$

We shall now prove that, provided $f(x)$ is a well-behaved function (in particular, continuous at $x = 0$),

$$\lim_{\eta\to 0}\int_{-\infty}^{\infty}\frac{f(x)\,\mathrm{d}x}{x\pm\mathrm{i}\eta} = \mathscr{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x \mp \mathrm{i}\pi f(0), \tag{III.1}$$

where $\mathscr{P}$ denotes the *Cauchy principal value* of the integral following it, defined by

$$\mathscr{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x = \lim_{\eta\to 0}\left[\int_{-\infty}^{-\eta}\frac{f(x)}{x}\mathrm{d}x + \int_{\eta}^{\infty}\frac{f(x)}{x}\mathrm{d}x\right], \tag{III.2}$$

η being a positive constant.

The formula (III.1) is often expressed in the compact form

$$\lim_{\eta\to 0}\frac{1}{x\pm\mathrm{i}\eta} = \mathscr{P}\frac{1}{x}\mp\mathrm{i}\pi\delta(x), \tag{III.3}$$

it being understood that both sides are to appear in an integrand multiplied by a well-behaved function of x. Here $\delta(x)$ is the Dirac *delta function*, which vanishes when $x \neq 0$, and for which

$$\int_{-b}^{b} f(x)\delta(x)\,\mathrm{d}x = f(0), \tag{III.4}$$

where b is any positive number.

To prove (III.1), we make use of the fact that

$$\frac{1}{x\pm \mathrm{i}\eta}=\frac{x\mp \mathrm{i}\eta}{x^2+\eta^2}=\frac{x}{x^2+\eta^2}\mp\frac{\mathrm{i}\eta}{x^2+\eta^2}. \tag{III.5}$$

Now,

$$\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2}=0, \quad \text{when} \quad x\neq 0, \tag{III.6}$$

and

$$\int_{-b}^{b}\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2}\mathrm{d}x=\lim_{\eta\to 0}\left(\tan^{-1}\frac{b}{\eta}-\tan^{-1}\frac{-b}{\eta}\right)=\pi. \tag{III.7}$$

Therefore, from the definition of the delta function, we have

$$\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2}=\pi\delta(x), \tag{III.8}$$

which gives the second term in (III.3).

Also,

$$\lim_{\eta\to 0}\int_{-\infty}^{\infty}\frac{f(x)x\,\mathrm{d}x}{x^2+\eta^2}=\lim_{\eta\to 0}\left[\int_{-\infty}^{-\eta}\frac{f(x)x\,\mathrm{d}x}{x^2+\eta^2}+\int_{\eta}^{\infty}\frac{f(x)x\,\mathrm{d}x}{x^2+\eta^2}+\int_{-\eta}^{\eta}\frac{f(x)x\,\mathrm{d}x}{x^2+\eta^2}\right]$$

$$=\lim_{\eta\to 0}\left[\int_{-\infty}^{-\eta}\frac{f(x)}{x}\mathrm{d}x+\int_{\eta}^{\infty}\frac{f(x)}{x}\mathrm{d}x\right]+\lim_{\eta\to 0}\int_{-\eta}^{\eta}\frac{f(x)x\,\mathrm{d}x}{x^2+\eta^2}$$

$$=\mathscr{P}\int_{-\infty}^{\infty}\frac{f(x)}{x}\mathrm{d}x+f(0)\lim_{\eta\to 0}\int_{-\eta}^{\eta}\frac{x\,\mathrm{d}x}{x^2+\eta^2}. \tag{III.9}$$

The integral in the second term vanishes, since the integrand is odd, and so we are left with the first term of eq. (III.1). This completes the proof.

It follows immediately that, if x_0 is any constant,

$$\lim_{\eta\to 0}\int_{-\infty}^{\infty}\frac{f(x)\,\mathrm{d}x}{x-x_0+\mathrm{i}\eta}=\mathscr{P}\int_{-\infty}^{\infty}\frac{f(x)\,\mathrm{d}x}{x-x_0}\mp \mathrm{i}\pi f(x_0), \tag{III.10}$$

or, compactly,

$$\lim_{\eta\to 0}\frac{1}{x-x_0+\mathrm{i}\eta}=\mathscr{P}\frac{1}{x-x_0}\mp \mathrm{i}\pi\delta(x-x_0). \tag{III.11}$$

Appendix IV

THE SOLUTION OF INHOMOGENEOUS DIFFERENTIAL EQUATIONS BY MEANS OF GREEN FUNCTIONS

Let us consider the differential equation

$$(L-\lambda)u(\boldsymbol{r})=f(\boldsymbol{r}), \tag{IV.1}$$

where L is a linear differential operator, λ a given constant and $f(\boldsymbol{r})$ a given function. Such an equation is said to be *inhomogeneous* because $u(\boldsymbol{r})$ does not appear as a factor in the term on the right-hand side. It is clear that, provided $f(\boldsymbol{r})$ is not identically zero, any multiple of a solution is not itself a solution. The equation

$$(L-\lambda)\psi(\boldsymbol{r})=0, \tag{IV.2}$$

on the other hand, is homogeneous, and any multiple of a solution *is* itself a solution.

We wish to solve eq. (IV.1) within some region of space Ω, subject to given boundary conditions. However, although the equation is inhomogeneous, we shall assume that the boundary conditions are homogeneous; that is to say, any multiple of $u(\boldsymbol{r})$ satisfies the same boundary conditions as $u(\boldsymbol{r})$. For example, $u(\boldsymbol{r})=0$ over the bounding surface of Ω is a homogeneous boundary condition, and so are the periodic boundary conditions familiar in the theory of metals.

Now, let us suppose that $\psi_n(\boldsymbol{r})$ is an eigenfunction of eq. (IV.2) corresponding to the eigenvalue λ_n, that is,

$$(L-\lambda_n)\psi_n(\boldsymbol{r})=0, \tag{IV.3}$$

assuming the same boundary conditions as on eq. (IV.1), and let us expand $u(\boldsymbol{r})$ and $f(\boldsymbol{r})$ in terms of the complete set of $\psi_n(\boldsymbol{r})$. Thus

$$u(\boldsymbol{r})=\sum_n a_n\psi_n(\boldsymbol{r}), \quad f(\boldsymbol{r})=\sum_n b_n\psi_n(\boldsymbol{r}), \tag{IV.4}$$

where we have assumed for simplicity that the eigenvalues λ_n form a discrete set. Eq. (IV.1) then becomes

$$\sum_n a_n(\lambda_n-\lambda)\psi_n(\boldsymbol{r}) = \sum_n b_n\psi_n(\boldsymbol{r}), \tag{IV.5}$$

giving

$$a_n = \frac{b_n}{\lambda_n-\lambda}. \tag{IV.6}$$

Now, if ψ_n is normalized within Ω, so that

$$\int |\psi_n|^2 \mathrm{d}\boldsymbol{r} = 1, \tag{IV.7}$$

we have

$$b_n = \int \psi_n^*(\boldsymbol{r}) f(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}, \tag{IV.8}$$

the integrals being taken throughout Ω. Hence

$$\begin{aligned} u(\boldsymbol{r}) &= \sum_n \frac{\psi_n(\boldsymbol{r}) \int \psi_n^*(\boldsymbol{r}) f(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}}{\lambda_n-\lambda} \\ &= -\int G(\boldsymbol{r},\boldsymbol{r}';\lambda) f(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \end{aligned} \tag{IV.9}$$

where

$$G(\boldsymbol{r},\boldsymbol{r}';\lambda) = \sum_n \frac{\psi_n(\boldsymbol{r})\psi_n^*(\boldsymbol{r}')}{\lambda-\lambda_n}, \tag{IV.10}$$

which is called the *Green function*† for the problem.

In order to find the differential equation satisfied by the Green function, let us set

$$f(\boldsymbol{r}) = -\delta(\boldsymbol{r}-\boldsymbol{r}_0).$$

Then eq. (IV.9) gives

$$\begin{aligned} u(\boldsymbol{r}) &= \int G(\boldsymbol{r},\boldsymbol{r}';\lambda)\delta(\boldsymbol{r}-\boldsymbol{r}_0)\,\mathrm{d}\boldsymbol{r}' \\ &= G(\boldsymbol{r},\boldsymbol{r}_0;\lambda), \end{aligned} \tag{IV.11}$$

---

† The name *Green's function* is perhaps more common, but the omission of the possessive is consistent with the use of the names Feynman graph, Fermi surface, Schrödinger equation, etc. It should be noted also that the Green function is sometimes defined as the negative of (IV.10). Clearly, this is unimportant, and merely requires a change of sign in (IV.9). The sign chosen here is that which is most frequently encountered in the literature relevant to ch. 10.

and we deduce that $G(\boldsymbol{r},\boldsymbol{r}';\lambda)$ satisfies the equation

$$(L-\lambda)G(\boldsymbol{r},\boldsymbol{r}';\lambda) = -\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{IV.12}$$

subject to the same boundary conditions as $u(\boldsymbol{r})$.

On the assumption that the λ_n are real (as will be the case if L is a Hermitian operator) and that λ is real, it follows from eq. (IV.10) that

$$G(\boldsymbol{r},\boldsymbol{r}';\lambda) = [G(\boldsymbol{r}',\boldsymbol{r};\lambda)]^*. \tag{IV.13}$$

A difficulty occurs if λ is equal to one of the eigenvalues, λ_m say, of L. In this case, eq. (IV.5) becomes

$$\sum_n a_n(\lambda_n-\lambda_m)\psi_n(\boldsymbol{r}) = \sum_n b_n\psi_n(\boldsymbol{r}), \tag{IV.14}$$

which can be so only if $b_m = 0$. That is to say, eq. (IV.1) has a solution in this case only if

$$\int \psi_m^*(\boldsymbol{r}) f(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = 0. \tag{IV.15}$$

A useful application of the Green function method is in converting a *homogeneous* differential equation into an integral equation. For example, the equation

$$[L+V(\boldsymbol{r})-\lambda]u(\boldsymbol{r}) = 0,$$

or

$$(L-\lambda)u(\boldsymbol{r}) = -V(\boldsymbol{r})u(\boldsymbol{r}), \tag{IV.16}$$

is homogeneous. However, simply applying the procedure described above, with $-V(\boldsymbol{r})u(\boldsymbol{r})$ substituted for $f(\boldsymbol{r})$, eq. (IV.9) becomes

$$u(\boldsymbol{r}) = \int G(\boldsymbol{r},\boldsymbol{r}';\lambda)V(\boldsymbol{r}')u(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \tag{IV.17}$$

with $G(\boldsymbol{r},\boldsymbol{r}';\lambda)$ again given by eq. (IV.10) and satisfying eq. (IV.12). Eq. (IV.17) is an integral equation for $u(\boldsymbol{r})$. An application of this to the Schrödinger equation is given in ch. 10.

Although here we have specified three spatial dimensions, clearly the procedure is the same for any number of dimensions.

Appendix V

FOURIER TRANSFORMS

The *Fourier integral theorem* states that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \int_{-\infty}^{\infty} f(x') \exp\left[\mathrm{i}(x-x')k\right] \mathrm{d}x', \tag{V.1}$$

provided that the integral exists. This may be proved by expanding $f(x)$ in complex Fourier series in the interval $(-L,L)$ and then letting $L \to \infty$.

If we define the *Fourier transform* $\phi(k)$ of $f(x)$ by

$$\phi(k) = \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \tag{V.2}$$

then it follows that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(k)\, \mathrm{e}^{\mathrm{i}kx} \mathrm{d}k. \tag{V.3}$$

Sometimes the constant $1/2\pi$, or its square root, is included in the definition of $\phi(k)$.

As an application of this we consider the delta function†. We have

$$\phi(k) = \int_{-\infty}^{\infty} \delta(x)\, \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x = 1, \tag{V.4}$$

so that

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx} \mathrm{d}k. \tag{V.5}$$

---

† Strictly speaking, $\delta(x)$ is not a function, but a *distribution*. However, it may be treated in exactly the same way as a function, provided that it is remembered that it is only significant when it appears in an integrand.

This is a useful form for the delta function. Although the integral does not exist in a rigorous sense, none the less, when it appears in the integrand of an integral with respect to x, it has the same effect as $\delta(x)$.

The foregoing may easily be generalized to functions of any number of variables. We define the Fourier transform $\phi(k_1, k_2, \ldots, k_n)$ of a function $f(x_1, x_2, \ldots, x_n)$ by

$$\phi(k_1,, k_2, \ldots, k_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n) \times \exp\left[-\mathrm{i}(k_1x_1 + \ldots + k_nx_n)\right] \mathrm{d}x_1 \ldots \mathrm{d}x_n, \tag{V.6}$$

whence

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(k_1, k_2, \ldots, k_n) \times \exp\left[\mathrm{i}(k_1x_1 + \ldots + k_nx_n)\right] \mathrm{d}k_1 \ldots \mathrm{d}k_n. \tag{V.7}$$

In particular, for a function $f(\boldsymbol{r})$ of position in space, we may use the condensed notation

$$\phi(\boldsymbol{k}) = \int f(\boldsymbol{r}) \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}, \tag{V.8}$$

$$f(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \phi(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\, \mathrm{d}\boldsymbol{k}, \tag{V.9}$$

the integrals being taken throughout the respective spaces.

As an application of this we consider the three-dimensional delta function $\delta(\boldsymbol{r})$ (defined by an obvious generalization of (III.4)). We have

$$\phi(\boldsymbol{k}) = \int \delta(\boldsymbol{r}) \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} = 1, \tag{V.10}$$

so that

$$\delta(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\, \mathrm{d}\boldsymbol{k}. \tag{V.11}$$

As a second important example, we shall find the Fourier transform of the function $1/r$. We make use of the fact that

$$\nabla^2\left(\frac{1}{r}\right) = -4\pi\delta(\boldsymbol{r}). \tag{V.12}$$

This may be proved by noting that $\nabla^2(1/r)$ is zero everywhere except at $r = 0$, and applying the divergence theorem of Gauss to a sphere S, of volume v, centred at $r = 0$. This gives

$$\int_v \nabla^2\left(\frac{1}{r}\right) \mathrm{d}v = \int_S \nabla\left(\frac{1}{r}\right) \cdot \mathrm{d}\boldsymbol{S}$$

$$= \int_S \frac{\mathrm{d}S}{r^2} = -4\pi, \tag{V.13}$$

which is what is obtained also from (V.12).

Thus, from eqs. (V.11) and (V.12), we have

$$\nabla^2 \int \phi(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k} = -4\pi \int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k} \tag{V.14}$$

or

$$\int \phi(\boldsymbol{k})(-k^2) \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k} = -4\pi \int \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k}, \tag{V.15}$$

where $\phi(\boldsymbol{k})$ is here the Fourier transform of $1/r$. We deduce that

$$\phi(\boldsymbol{k}) = \frac{4\pi}{k^2}, \tag{V.16}$$

and

$$\frac{1}{r} = \frac{1}{2\pi^2} \int \frac{1}{k^2} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\,\mathrm{d}\boldsymbol{k}, \tag{V.17}$$

which direct integration proves to be correct. The method used here avoids the awkward question of the convergence of the integral in eq. (V.8) when $f(\boldsymbol{r}) = 1/r$.

Appendix VI

CONTOUR INTEGRALS

Definite integrals of functions of a real variable may often be evaluated most conveniently by regarding them as part of a contour integral in the complex z-plane. This is particularly so for the integrals occurring in Green function theory, as will be seen in chs. 10 and 11. We shall state here the relevant definitions and theorems.

A one-valued function $f(z)$ of the complex variable $z = x + \mathrm{i}y$ is said to be *differentiable* at a point z if

$$f'(z) = \lim_{\delta z \to 0} \frac{f(z+\delta z) - f(z)}{\delta z} \tag{VI.1}$$

exists and is independent of the path by which $\delta z \to 0$. (*N.B.* $z + \delta z$ can approach z from any direction in the z-plane.) A function which is one-valued and differentiable in a given region of the z-plane is said to be *analytic* in that region (other names, such as *regular* or *holomorphic*, sometimes with slightly varying definitions, are often used).

Any one-valued function $f(z)$ may be expressed in the form

$$f(z) = u + \mathrm{i}v, \tag{VI.2}$$

where u and v are real, one-valued functions of x and y. A necessary and sufficient condition for $f(z)$ to be analytic in a given region is that u and v satisfy the equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{VI.3}$$

in the region. These are known as the *Cauchy–Riemann equations*.

A function $f(z)$ which is analytic within a circle centred at the point $z=a$ can be expanded in a series of the form

$$f(z) = A_0+A_1(z-a)+A_2(z-a)^2+\dots, \tag{VI.4}$$

within the circle, where

$$A_n = \frac{1}{n!}f^{(n)}(a). \tag{VI.5}$$

This is just Taylor's series for the function near $z=a$.

A point at which a function fails to be analytic is called a *singular point* or *singularity* of the function. Singularities can be of several types, the most important of which, as far as the present work is concerned, is called a *pole*, which may be defined as follows:

If, in the neighbourhood of a point $z=a$, a function $f(z)$ may be expanded in a series of the form

$$\begin{aligned} f(z) &= A_0+A_1(z-a)+A_2(z-a)^2+\dots \\ &+\frac{B_1}{z-a}+\frac{B_2}{(z-a)^2}+\dots+\frac{B_n}{(z-a)^n}, \end{aligned} \tag{VI.6}$$

where the A_i and B_i are constants and n is *finite* but not zero, $f(z)$ is said to have a *pole of order n* at $z=a$. If $n=1$, it is called a *simple pole*.

If n is not finite, but the series of reciprocal powers of $z-a$ is infinite, the singularity at $z=a$ is not a pole, but is called an *essential singularity*. For example,

$$e^{1/z} = 1+\frac{1}{z}+\frac{1}{2!z^2}+\frac{1}{3!z^3}+\dots, \tag{VI.7}$$

has an essential singularity at $z=0$.

The series (VI.6) is called *Laurent's series* for $f(z)$ near $z=a$. Clearly, if $f(z)$ were analytic at $z=a$, all the coefficients B_i would be zero and Laurent's series would reduce to Taylor's series.

Whatever the value of $n(\geqslant 1)$, the coefficient B_1 of $(z-a)^{-1}$ is called the *residue* at $z=a$. This applies to both poles and essential singularities.

The *residue theorem*, which is so important in contour integration, may be stated as follows:

If $f(z)$ is analytic inside and on a simple contour C, except for a finite number of poles within C, then

$$\oint f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\sum(\text{residues}). \tag{VI.8}$$

(*N.B.* A simple contour is a closed curve which does not intersect itself.) The integral is a line integral, or contour integral, around C in an anti-clockwise sense, and Σ (residues) is the sum of the residues at all the poles within C.

If $f(z)$ has no poles inside or on C, the right-hand side of (VI.8) reduces to zero. In other words, if $f(z)$ is analytic *everywhere* inside and on C, the contour integral is zero. This is known as *Cauchy's theorem.*

If $z=a$ is a *simple pole* of $f(z)$, the residue is given by the formula

$$B_1 = \lim_{z\to a}(z-a)f(z), \tag{VI.9}$$

which follows immediately from eq. (VI.6) with $n=1$. For example, consider the function $(z^2-4)^{-1}$. This has two simple poles, at $z=\pm 2$. The residue at $z=2$ is given by

$$\lim_{z\to 2}\frac{z-2}{z^2-4} = \lim_{z\to 2}\frac{1}{z+2} = \tfrac{1}{4}, \tag{VI.10}$$

and that at $z=-2$ is given by

$$\lim_{z\to -2}\frac{z+2}{z^2-4} = -\tfrac{1}{4}. \tag{VI.11}$$

Both residues are obtained immediately by writing the function in the form

$$\frac{1}{z^2-4} = \frac{1}{4}\left(\frac{1}{z-2}-\frac{1}{z+2}\right), \tag{VI.12}$$

since $1/(z+2)$ is analytic near $z=2$ (and hence expressible in Taylor's series), while $1/(z-2)$ is analytic near $z=-2$.

As a second example, let us take that occurring in §10.1 (following eq. (10.24)). The problem (in our present notation) is to evaluate the contour integral

$$\oint\frac{\exp(\mathrm{i}rz)\,\mathrm{d}z}{(x_0-z+\mathrm{i}\eta')(x_0+z+\mathrm{i}\eta')}, \tag{VI.13}$$

where r, x_0 and η' are positive real constants, around a semicircle of radius R in the upper half-plane and along the real axis from $-R$ to R. The integrand (on the assumption that $R > |x_0+\mathrm{i}\eta'|$) has one simple pole inside the contour, at $z = x_0+\mathrm{i}\eta'$, the residue being

$$\lim_{z\to x_0+\mathrm{i}\eta'}\frac{-\exp(\mathrm{i}rz)}{x_0+z+\mathrm{i}\eta'} = -\frac{\exp[\mathrm{i}r(x_0+\mathrm{i}\eta')]}{2(x_0+\mathrm{i}\eta')}, \tag{VI.14}$$

and the value of the contour integral is $2\pi\mathrm{i}$ times this.

The integral around the semicircle, where $z = R\,\mathrm{e}^{\mathrm{i}\theta}$, becomes

$$\int_0^{\pi} \frac{\exp\,[\mathrm{i}rR(\cos\theta+\mathrm{i}\sin\theta)]R\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta}{(x_0-R\,\mathrm{e}^{\mathrm{i}\theta}+\mathrm{i}\eta')(x_0+R\,\mathrm{e}^{\mathrm{i}\theta}+\mathrm{i}\eta')}. \tag{VI.15}$$

This tends to zero as $R \to \infty$, since $\sin\theta > 0$ for $0 < \theta < \pi$. We therefore obtain, for the integral along the real axis,

$$\int_{-\infty}^{\infty} \frac{\exp\,(\mathrm{i}rx)\,\mathrm{d}x}{(x_0-x+\mathrm{i}\eta')(x_0+x+\mathrm{i}\eta')} = -\frac{\pi\mathrm{i}\exp\,[\mathrm{i}r(x_0+\mathrm{i}\eta')]}{x_0+\mathrm{i}\eta'}. \tag{VI.16}$$

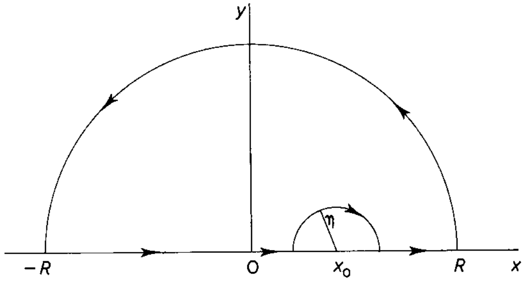
Contour integration in the z-plane may often be used to evaluate the Cauchy principal value (see Appendix III) of the integral of a function $F(x)$ which has a simple pole at some point on the x-axis. To be definite, suppose that we wish to evaluate

$$\mathscr{P}\int_{-\infty}^{\infty} F(x)\,\mathrm{d}x,$$

given that $F(x)$ has one simple pole at $x = x_0$. We consider the function $F(z)$ of the complex variable z which reduces to $F(x)$ when z is real. Since $F(x)$ has a simple pole at x_0, it follows that, near $z = x_0$, we can write $F(z)$ in the form

$$F(z) = G(z) + \frac{B_1}{z-x_0}, \tag{VI.17}$$

where $G(z)$ is analytic and B_1 is the residue at $z = x_0$. We now integrate $F(z)$ around the closed contour shown below, which consists of a large semicircle, of radius R, and a small semicircle, of radius η, in the upper half-plane, together with those parts of the x-axis from $-R$ to $x_0-\eta$ and $x_0+\eta$ to R, ultimately letting $R \to \infty$ and $\eta \to 0$. Let us assume, as is generally the case, that the integral around the large semicircle vanishes in the limit as $R \to \infty$. The integral around the small semicircle, on which $z - x_0 = \eta\,\mathrm{e}^{\mathrm{i}\theta}$, is



$$\begin{aligned}\int_{\pi}^{0} F(z)\eta\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta &= \int_{\pi}^{0}\left[G(z)+\frac{B_1}{z-x_0}\right]\eta\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta \\ &= \int_{\pi}^{0}[\eta G(z)+B_1\mathrm{e}^{-\mathrm{i}\theta}]\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta \\ &= -B_1\pi\mathrm{i},\end{aligned} \tag{VI.18}$$

in the limit as $\eta \to 0$. Again in the limits as $R \to \infty$ and $\eta \to 0$, therefore, we have

$$\begin{aligned}\oint F(z)\,\mathrm{d}z &= \lim_{\eta\to 0}\left[\int_{-\infty}^{x_0-\mathrm{i}\eta} F(x)\,\mathrm{d}x + \int_{x_0+\mathrm{i}\eta}^{\infty} F(x)\,\mathrm{d}x\right] - B_1\pi\mathrm{i} \\ &= \mathscr{P}\int_{-\infty}^{\infty} F(x)\,\mathrm{d}x - B_1\pi\mathrm{i} \\ &= 2\pi\mathrm{i}\sum\text{(residues)},\end{aligned} \tag{VI.19}$$

where Σ (residues) is the sum of the residues at the poles of $f(z)$ in the upper half-plane (but not on the x-axis). In other words,

$$\mathscr{P}\int_{-\infty}^{\infty} F(x)\,\mathrm{d}x = 2\pi\mathrm{i}\left[\tfrac{1}{2}(\text{residue at } x_0) + \sum\text{(residues)}\right]. \tag{VI.20}$$

If the small semicircle were taken in the lower half-plane, the pole at x_0 would be inside the contour, but the sign of the integral around the small semicircle would be reversed, giving the same final result. It follows that the Cauchy principal value is the mean of the two values of the contour integral with the pole at x_0 inside and outside the contour.

As a first example, we may deduce the result proved without the use of contour integration in Appendix III, eq (III.10). Suppose that

$$F(z) = \frac{f(z)}{z-x_0-\mathrm{i}\eta}, \tag{VI.21}$$

where $f(z)$ is analytic on the x-axis, but may have a finite number of poles in the upper half-plane. Then, in the limit as $\eta \to 0$, $F(z)$ has a simple pole at $x = x_0$, with residue $f(x_0)$. If we assume, as above, that the integral of $F(z)$ around a semicircle of radius R in the upper half-plane vanishes in the limit as $R \to \infty$, then it follows from (VI.20) that

$$\lim_{\eta \to 0} \int_{-\infty}^{\infty} \frac{f(x)\,\mathrm{d}x}{x - x_0 - \mathrm{i}\eta} = 2\pi\mathrm{i} \sum \text{(residues)}$$

$$= \mathscr{P} \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} + \pi\mathrm{i} f(x_0), \qquad \text{(VI.22)}$$

where Σ (residues) includes that at $z = x_0 + \mathrm{i}\eta$ (or $z = x_0$, in the limit as $\eta \to 0$). This is one of the results given in eq. (III.10) – the other, with $+\mathrm{i}\eta$ replacing $-\mathrm{i}\eta$, follows immediately since, in this case, Σ (residues) does not contain that at $z = x_0 - \mathrm{i}\eta$.

As a second example of the use of eq. (VI.20), let us derive the result given in eq. (11.87). In the notation of this Appendix, we first wish to evaluate

$$\mathscr{P} \int_{-\infty}^{\infty} \frac{\exp(-\mathrm{i}xt)\,\mathrm{d}x}{x - x_0},$$

where t is a *negative* constant. (*N.B.* in §11.6 we have ϵ instead of z and ϵ_k instead of x_0.) We integrate the function

$$F(z) = \frac{\exp(-\mathrm{i}zt)}{z - x_0} \qquad \text{(VI.23)}$$

around a contour consisting of a semicircle of radius R in the upper half-plane together with that section of the real axis from $x = -R$ to $x = R$, and ultimately let $R \to \infty$. On the semicircle we have $z = R\,\mathrm{e}^{\mathrm{i}\theta}$, so that the integral of $F(z)$ around the semicircle becomes

$$\int_0^{\pi} \frac{\exp[-\mathrm{i}(\cos\theta + \mathrm{i}\sin\theta)Rt]}{R\,\mathrm{e}^{\mathrm{i}\theta} - x_0} \mathrm{i}R\,\mathrm{e}^{\mathrm{i}\theta}\,\mathrm{d}\theta,$$

which tends to zero as $R \to \infty$, since $t < 0$ and $\sin\theta > 0$ for $0 < \theta < \pi$. The function $F(z)$ has no poles inside the contour, but only a simple pole at $z = x_0$, the residue there being $\exp(-\mathrm{i}x_0 t)$. It follows from eq. (VI.20), therefore, that

$$\mathscr{P} \int_{-\infty}^{\infty} \frac{\exp(-\mathrm{i}xt)\,\mathrm{d}x}{x - x_0} = \pi\mathrm{i}\exp(-\mathrm{i}x_0 t), \quad \text{for} \quad t < 0. \qquad \text{(VI.24)}$$

Similarly, for $t > 0$, we take a semicircle in the lower half-plane, so that the integral around it again tends to zero as $R \to \infty$, and find that

$$\mathscr{P} \int_{-\infty}^{\infty} \frac{\exp(-\mathrm{i}xt)\,\mathrm{d}x}{x - x_0} = -\pi\mathrm{i}\exp(-\mathrm{i}x_0 t), \quad \text{for} \quad t > 0. \qquad \text{(VI.25)}$$

Appendix VII

WICK'S THEOREM

We have made no explicit use of Wick's theorem in the text, since its results are implicit in the graphical analysis of the perturbation series in ch. 7. However, most texts employ this theorem in order to establish the graphical method, so it may be useful to include a brief account of it here.

We revert to the *interaction picture* and the hole and particle notation introduced in §7.3. The *normal product* $N(AB\ldots)$ of a set of hole and particle, creation and destruction operators $A, B, \ldots$, is defined as the product obtained from $AB\ldots$ by moving all the creation operators to the left and all the destruction operators to the right, including a factor -1 for every interchange of adjacent operators required to bring about this ordering. For example,

$$N[a_i(t_1)b_j^\dagger(t_2)] = -b_j^\dagger(t_2)a_i(t_1), \qquad \text{(VII.1)}$$

$$N[a_i^\dagger(t_1)a_k(t_2)b_j(t_3)b_l^\dagger(t_4)] = a_i^\dagger(t_1)b_l^\dagger(t_4)a_k(t_2)b_j(t_3), \qquad \text{(VII.2)}$$

$$N[b_j(t_1)b_l^\dagger(t_2)a_i^\dagger(t_3)a_k^\dagger(t_4)] = -b_l^\dagger(t_2)a_i^\dagger(t_3)a_k^\dagger(t_4)b_j(t_1). \qquad \text{(VII.3)}$$

It should be noted that *the time arguments are immaterial here*. Also, the creation operators and the destruction operators may be ordered in any way among themselves provided the sign is adjusted accordingly – for example, (VII.2) might equally well be written

$$-b_l^\dagger(t_4)a_i^\dagger(t_1)a_k(t_2)b_j(t_3)$$

or

$$b_l^\dagger(t_4)a_i^\dagger(t_1)b_j(t_3)a_k(t_2).$$

The utility of a normal product lies in the fact that its expectation value is zero in the ground state Φ_0 of the non-interacting system (see §6.2), that is,

$$\langle\Phi_0|N(AB\ldots)|\Phi_0\rangle = 0. \tag{VII.4}$$

This is because Φ_0 contains neither holes nor particles, so that, as shown in eqs. (7.31) and (7.32), both hole and particle destruction operators act upon it to give zero.

The *contraction* or *pairing* $\underbracket{AB}$ of two operators is defined as

$$\underbracket{AB} = T(AB) - N(AB), \tag{VII.5}$$

where T is the time-ordering operator defined in eq. (11.11). For example, if $t_2 > t_1$,

$$\underbracket{a_i(t_1)a_k^\dagger(t_2)} = -a_k^\dagger(t_2)a_i(t_1) + a_k^\dagger(t_2)a_i(t_1) = 0, \tag{VII.6}$$

$$\underbracket{a_k^\dagger(t_2)a_i(t_1)} = a_k^\dagger(t_2)a_i(t_1) - a_k^\dagger(t_2)a_i(t_1) = 0, \tag{VII.7}$$

$$\begin{aligned}\underbracket{a_i(t_2)a_k^\dagger(t_1)} &= a_i(t_2)a_k^\dagger(t_1) + a_k^\dagger(t_1)a_i(t_2)\\ &= (a_ia_k^\dagger + a_k^\dagger a_i)\exp\,[\mathrm{i}(\epsilon_kt_1 - \epsilon_it_2)]\\ &= \delta_{ik}\exp\,[\mathrm{i}(\epsilon_kt_1 - \epsilon_it_2)],\end{aligned} \tag{VII.8}$$

using eqs. (7.34).

The contraction of any two operators can be found in the same way and is always a number. It follows, therefore, that

$$\underbracket{AB} = \langle\Phi_0|\underbracket{AB}|\Phi_0\rangle = \langle\Phi_0|T(AB)|\Phi_0\rangle. \tag{VII.9}$$

Applying this to a non-interacting free-electron system, we have

$$\begin{aligned}\underbracket{c_k(t)c_{k'}^\dagger}(t') &= \langle\Phi_0|T\{c_k(t)c_{k'}^\dagger(t')\}|\Phi_0\rangle\\ &= \mathrm{i}G_0(\boldsymbol{k}t,\boldsymbol{k}'t').\end{aligned} \tag{VII.10}$$

The Green function $G(\boldsymbol{k}t,\boldsymbol{k}'t')$ is defined by (11.48) in the Heisenberg picture, but the Heisenberg and interaction pictures are the same for a non-interacting system.

A contraction of one or more pairs of operators in a normal product is taken to mean that the contractions are to be removed from the normal product and plus or minus sign affixed according as the number of interchanges of adjacent operators required to do this is even or odd. Thus

$$\begin{aligned}N(\bcontraction{}{A}{BCD}{E}\bcontraction[0.5ex]{A}{B}{}{C}ABCDEF) &= -\underbracket{AE}N(\underbracket{BC}DF)\\ &= -\underbracket{AE}\,\underbracket{BC}N(DF),\end{aligned} \tag{VII.11}$$

$$\begin{aligned}N(\bcontraction{}{A}{BCD}{E}\bcontraction[0.5ex]{A}{B}{C}{D}ABCDEF) &= -\underbracket{AE}N(\bcontraction{}{B}{C}{D}BCDF)\\ &= \underbracket{AE}\,\underbracket{BD}N(CF).\end{aligned} \tag{VII.12}$$

Finally, *Wick's theorem* states that the time-ordered product of any number of operators is equal to the normal product plus the sum of all normal products in which one or more pairs of operators are contracted in all possible ways: that is,

$$\begin{aligned}T(ABCD\ldots WXYZ) &= N(ABCD\ldots WX\underbracket{YZ})\\ &+N(\underbracket{AB}CD\ldots WXYZ)+\ldots+N(ABCD\ldots WX\underbracket{YZ})\\ &+N(\underbracket{AB}\,\underbracket{CD}\ldots WXYZ)+\ldots+N(ABCD\ldots\underbracket{WX}\,\underbracket{YZ})+\ldots\\ &+N(\underbracket{AB}\,\underbracket{CD}\ldots\underbracket{WX}\,\underbracket{YZ})+N(\bcontraction{}{A}{B}{C}\bcontraction[2ex]{A}{B}{C}{D}ABCD\ldots\underbracket{WX}\,\underbracket{YZ})+\ldots.\end{aligned} \tag{VII.13}$$

We shall not prove this theorem, but deduce from it, together with (VII.4) and the statement leading to (VII.11), that

$$\begin{aligned}\langle\Phi_0|T(ABCD\ldots WXYZ)|\Phi_0\rangle &= \underbracket{AB}\,\underbracket{CD}\ldots\underbracket{WX}\,\underbracket{YZ}\\ &+\bcontraction{}{A}{B}{C}\bcontraction[2ex]{A}{B}{C}{D}ABCD\ldots\underbracket{WX}\,\underbracket{YZ}+\ldots.\end{aligned} \tag{VII.14}$$

In words, the expectation value of a time-ordered product in the state Φ_0 is the sum of all fully-contracted products (i.e. products in which all pairs of operators are contracted). If the two operators in each contracted pair are placed next to each other, a factor -1 must be included for every interchange of adjacent operators required to bring this about: for example,

$$\bcontraction{}{A}{B}{C}\bcontraction[2ex]{A}{B}{C}{D}ABCD\ldots\underbracket{WX}\,\underbracket{YZ} = -\underbracket{AC}\,\underbracket{BD}\ldots\underbracket{WX}\,\underbracket{YZ}.$$

Although, having avoided it in the text, we shall not pursue the matter here, the result expressed in eq. (VII.14) is clearly relevant to the evaluation of the matrix elements in the perturbation series for the ground-state energy, as discussed in ch. 7, or for $G(\boldsymbol{k},t)$, as given in eq. (11.176), and to the graphical representation of these matrix elements.

PROBLEMS AND EXERCISES

Chapter 1

1.1. From the expression for ΔE given in eq. (1.76), deduce that the third-order energy correction in the Rayleigh–Schrödinger perturbation theory is

$$\Delta E^{(3)} = \sum_{m,n=1}^{\infty} \frac{\langle \Phi_0|H'|\Phi_m\rangle\langle\Phi_m|H'|\Phi_n\rangle\langle\Phi_n|H'|\Phi_0\rangle}{(E_0-E_m)(E_0-E_n)}$$

$$-\Delta E^{(1)} \sum_{n=1}^{\infty} \frac{|\langle\Phi_0|H'|\Phi_n\rangle|^2}{(E_0-E_n)^2}.$$

(*N.B.* In the second term, ΔE has been replaced by the first-order approximation $\Delta E^{(1)}$.)

1.2. Using eqs. (1.59), (1.63) and (1.68), show that

$$\Psi_0 = C\Phi_0 + \frac{R}{E-H_0}H'\Psi_0,$$

and deduce that

$$\Psi_0 = C\sum_{n=0}^{\infty}\left(\frac{R}{E-H_0}H'\right)^n\Phi_0$$

and

$$\Delta E = \sum_{n=0}^{\infty}\left\langle\Phi_0\middle|H'\left(\frac{R}{E-H_0}H'\right)^n\middle|\Phi_0\right\rangle.$$

(This is the Brillouin–Wigner perturbation series. In this case, unlike the Rayleigh–Schrödinger theory, the perturbed energy E appears in the denominators of all terms on the right-hand side except the first.)

1.3. Show that the third-order approximation to ΔE in the Brillouin–Wigner theory is the same as that in the Rayleigh–Schrödinger theory, namely, $\Delta E^{(1)}+\Delta E^{(2)}+\Delta E^{(3)}$. (*N.B.* The energy denominators in the former theory must be expanded binomially, retaining only those terms of appropriate order.)

Chapter 2

2.1. If Φ stands for Φ^4_{1234}, i.e.

$$\frac{1}{\sqrt{4!}}\sum_P (-1)^p P\phi_1(x_1)\phi_2(x_2)\phi_3(x_3)\phi_4(x_4),$$

verify that

(1) $c_3\Phi = \Phi^3_{124}$,
(2) $c_3^\dagger\Phi = 0$,
(3) $c_5\Phi = 0$,
(4) $c_5^\dagger\Phi = \Phi^5_{12345}$,
(5) $c_3^\dagger c_3\Phi = \Phi$,
(6) $(c_2c_4+c_4c_2)\Phi = 0$,
(7) $(c_1c_5^\dagger + c_5^\dagger c_1)\Phi = 0$,
(8) $c_3c_2c_1\Phi = \Phi^1_4 = \phi_4(x_1)$.

2.2. Express all the determinantal functions appearing in the previous question in terms of occupation numbers.

2.3. Prove that $\{c_l^\dagger,c_k^\dagger\} = 0$, using the definition of $c_k^\dagger$ given in eq. (2.31).

2.4. Prove that $[c_k^\dagger c_k, c_k] = -c_k$ and $[c_k^\dagger c_k, c_k^\dagger] = c_k^\dagger$.

2.5. Use eqs. (2.126) and (2.128) to re-derive eqs. (2.84) and (2.95).

2.6. Show that

$$\psi^\dagger(x)\Phi^N_{123\ldots N}(x_1,x_2,\ldots,x_N)$$

$$=\frac{1}{[(N+1)!]^{1/2}}\begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \ldots\ \phi_1(x_{N+1}) \\ \vdots & \vdots & \vdots \\ \phi_N(x_1) & \phi_N(x_2) & \ldots\ \phi_N(x_{N+1}) \\ \delta(x-x_1) & \delta(x-x_2) & \ldots\ \delta(x-x_N) \end{vmatrix},$$

and hence verify that "$\psi^\dagger(x)$ creates an electron at x".

2.7. Show that the operator

$$\nu = \int \psi^\dagger(x)\psi(x)\,\mathrm{d}x$$

represents the total number of particles, and that

$$\psi(x)\nu = (\nu+1)\psi(x).$$

[*Hint:* let $f(x_i) = 1$ in H_0.]

2.8. With H_0 and H' given by (2.66) and (2.68), respectively, prove that

$$\sum_k c_k^\dagger[H_0,c_k] = -H_0$$

and

$$\sum_k c_k^\dagger[H',c_k] = -2H',$$

and hence that the energy of a state Ψ is given by

$$\tfrac{1}{2}\langle\Psi|H_0|\Psi\rangle - \tfrac{1}{2}\sum_k \langle\Psi|c_k^\dagger[H,c_k]|\Psi\rangle.$$

(See §11.10 for an application of this.)

Chapter 3

3.1. With $E(\eta)$ given by eq. (3.15), prove that

$$\left(\frac{\mathrm{d}^2E}{\mathrm{d}\eta^2}\right)_{\eta=0} = 2[\langle\Phi_q|H|\Phi_q\rangle - \langle\Phi_0|H|\Phi_0\rangle],$$

and hence that $E(0)$ is a minimum (assuming that

$$\langle\Phi_q|H|\Phi_q\rangle > \langle\Phi_0|H|\Phi_0\rangle).$$

3.2. Using eq. (3.25), prove directly that

$$\langle q|F_\mathrm{X}|p\rangle = \langle p|F_\mathrm{X}|q\rangle^*,$$

i.e. that $F_\mathrm{X}(\boldsymbol{r})$ is Hermitian.

3.3. Derive the expression for ϵ_k given in eq. (3.45).

3.4. Write down the Hartree–Fock equation, similar to (3.23), for the ground state of a ferromagnetic system, in which the spins of all the electrons are parallel. Apply this to a free-electron gas and show that the average Hartree–Fock energy per electron in this case is

$$\frac{3.51}{r_\mathrm{s}^2} - \frac{1.154}{r_\mathrm{s}}\ \mathrm{ry}.$$

3.5. Using the method of §3.6, show that the contribution to the second-order energy correction $\Delta E^{(2)}_{\uparrow\uparrow}$, from the interaction of electrons with parallel spins, is

$$\Delta E^{(2)}_{\uparrow\downarrow} + \frac{16\pi^2e^4m}{\hbar^2\Omega^2}\sum_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{k}} \frac{1}{k^2(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})^2\boldsymbol{k}\cdot(\boldsymbol{p}-\boldsymbol{q}+\boldsymbol{k})},$$

where $p,q < k_\mathrm{F}$ and $|\boldsymbol{p}+\boldsymbol{k}|,|\boldsymbol{q}-\boldsymbol{k}| > k_\mathrm{F}$. Show also that the second term does not diverge.

3.6. Assuming a delta function interaction

$$v(\boldsymbol{r}_1,\boldsymbol{r}_2) = \lambda\delta(\boldsymbol{r}_1-\boldsymbol{r}_2),$$

with λ constant, instead of the Coulomb interaction in §3.4, show that

$$H' = \frac{\lambda}{2\Omega}\sum_{\substack{k,p,q\\ \sigma,\sigma'}} c^\dagger_{p+k,\sigma}\, c^\dagger_{q-k,\sigma'}\, c_{q\sigma'}\, c_{p\sigma},$$

and that the total exchange energy becomes $-3N\lambda/16\pi r_\mathrm{s}^3$. Does $\Delta E^{(2)}$ diverge in this case?

Chapter 4

4.1. With S defined by eq. (4.25), prove the following commutation relations:

$[P_k,S] = -\mathrm{i}\hbar M_k\rho_k,$

$[\boldsymbol{r}_i,S] = 0, \quad [\rho_k,S] = 0,$

$[[P_k,S],S] = 0, \quad [Q_k,S] = 0,$

$[\boldsymbol{p}_i,S] = -\hbar \sum_{k<k_\mathrm{c}} M_kQ_k\boldsymbol{k}\exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i),$

$[[\boldsymbol{p}_i,S],S] = 0.$

4.2. Using eqs. (4.24) and (4.29), and expanding $\exp(\mathrm{i}S/\hbar)$ in power series, show that

$$O_\mathrm{new} = O + \frac{\mathrm{i}}{\hbar}[O,S] - \frac{1}{2\hbar^2}[[O,S],S] + \ldots.$$

4.3. From the results of the previous two problems, derive eqs. (4.31), (4.32), (4.33), (4.35) and (4.38).

4.4. Give reasons for saying that, for the practical utility of the plasma theory, k_c must be neither too large nor too small.

4.5. Apply the theory of §4.3 and §4.4 to the ground state of a free-electron gas, assumed to be ferromagnetic (see Problem 3.4). Show that the value of k_c obtained in this case is $2^{2/3}$ times that given in eq. (4.82).

In view of the fact that all electrons have parallel spins, what may one deduce about the short-range correlation energy?

Chapter 5

5.1. Verify the formula for $(\partial/\partial t)(AB\Psi)$, given in the footnote on p. 95, when

$$A = \left(\frac{\partial}{\partial x} + t\right), \quad B = x\frac{\partial}{\partial t} \quad \text{and} \quad \Psi = xt^2.$$

5.2. Since $H = H_0 + H'$, why is it not permissible to reduce eq. (5.37) to the form

$$U(t,t') = \exp\left[\mathrm{i}H'(t'-t)\right]?$$

[*Hint:* compare the expansions of $\mathrm{e}^A\mathrm{e}^B$ and e^{A+B} when A and B are non-commuting operators.]

Chapter 6

6.1. If $H' = g$, a constant, eq. (6.16) gives immediately the obvious result $\Delta E = g$. Obtain the same result from eqs. (6.22) and (6.23).

6.2. Show that eq. (6.32) leads to the Rayleigh–Schrödinger expression for $\Delta E^{(3)}$ given in Problem 1.1.

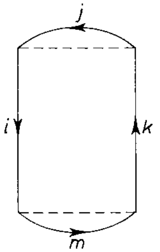
Chapter 7

7.1. Express the vacuum–vacuum matrix elements, for example $\langle\Phi_0|c_i^\dagger c_j^\dagger c_j c_i|\Phi_0\rangle$, in eqs. (7.42), (7.46), (7.51) and (7.55) in terms of particle and hole operators such as $a_i^\dagger, a_i, b_i^\dagger, b_i$.

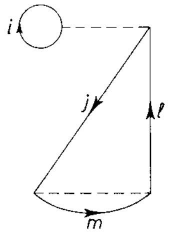
7.2. Verify eq. (7.50) using eqs. (6.36) and (1.48).

7.3. Verify the rule for calculating vacuum–vacuum matrix elements (p. 138) for the graphs on the opposite page.

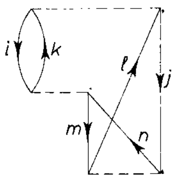
7.4. Although the techniques described in this book are intended mainly for use in connection with electron gases such as occur in metals, most of them can, in principle at least, be applied to any system of electrons. Consider, for example, the application of the work of ch. 7 to a helium atom (see WM, p. 120), taking the interaction between the two electrons as a perturbation and the one-electron functions to be the hydrogenic wave functions of a single electron in the field of the helium nucleus. (Note that, in the ground state of the unperturbed system, both electrons are in 1s hydrogenic states, with opposite spins. There is no Fermi surface, and only two 'hole' states are possible.)



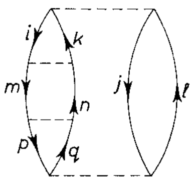
(a)



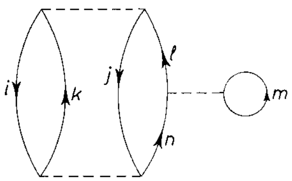
(b)



(c)



(d)



(e)

Chapter 8

8.1. Extend the demonstration of the linked graph theorem given in §8.1 to include graphs of the fourth order.

Chapter 9

9.1. Derive the expression for $\Delta E^{(4)}_{\text{ring}}$ given in eq. (9.41), using fig. 9.6.

9.2. In the derivation of eq. (9.47), prove that

$$\int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3 \int_0^\infty \mathrm{d}t_1 \int_{-t_1}^0 \exp\left[-t_1(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_1)\right.$$

$$\left. + t_2(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_2) - (t_1+t_2)(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_3)\right]\mathrm{d}t_2$$

$$= \int \mathrm{d}\boldsymbol{p}_1 \int \mathrm{d}\boldsymbol{p}_2 \int \mathrm{d}\boldsymbol{p}_3 \int_0^\infty \mathrm{d}t_1 \int_0^\infty \exp\left[-t_1(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_1)\right.$$

$$\left. - t_2(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_2) - (t_1+t_2)(\tfrac{1}{2}k^2 + \boldsymbol{k}\cdot\boldsymbol{p}_3)\right]\mathrm{d}t_2.$$

[*Hint:* first invert the order of the t_1,t_2 integrations and then make the transformation $x = -t_2$, $y = t_1 + t_2$, remembering that the regions of integration for p_1,p_2,p_3 are all the same, namely, the Fermi sphere.]

9.3. Draw graphs of the correlation energy against r_s, between $r_s = 0$ and $r_s = 6$ Bohr units, according to the expressions of (1) Wigner, eq. (4.98), (2) Gell-Mann and Brueckner, eq. (9.68), (3) Bohm and Pines (or Nozières and Pines), eq. (9.69), and (4) the amended form of the latter, eq. (9.71). Explain the differences among the four graphs. (Neglect the terms $O(r_s)$.)

Chapter 10

10.1. Substitute the expression (10.55) for $G(\boldsymbol{r}; \epsilon + i\eta)$ in eq. (10.52), and show that it leads to the expression (10.56) for $G(\boldsymbol{k}; \epsilon + i\eta)$.

10.2. Show that eq. (10.59) leads to the correct density of states for a free electron, as given in eq. (10.50).

10.3. Apply the methods of ch. 10 to a *one-dimensional* system. In particular, verify that the analogues of eqs. (10.46) and (10.59) give the correct density of states

$$\mathcal{N}(\epsilon) = \frac{L}{\pi}\left(\frac{m}{2\epsilon}\right)^{1/2}$$

for a particle in a one-dimensional box of length L.

Chapter 11

11.1. Verify that $\{c_i(t), c_j^\dagger(t')\} = \delta_{ij}$ only if $t = t'$.

11.2. Give physical interpretations of $G(\boldsymbol{k},\boldsymbol{k}',t)$ and $G(\boldsymbol{r},\boldsymbol{r}',t)$ when $t < 0$.

11.3. From the definitions (11.104) and (11.105), prove that

$$\int_0^\infty [A(\boldsymbol{k},\epsilon) + B(\boldsymbol{k},\epsilon)]\, d\epsilon = 1.$$

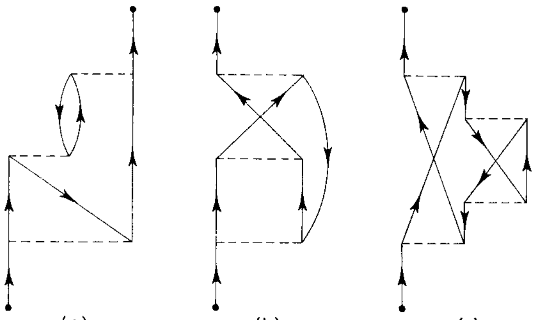
[*Hint:* $\sum_n |\langle\Psi_n^{N+1}|c_k^\dagger|\Psi_0\rangle|^2 = \langle\Psi_0|c_k c_k^\dagger|\Psi_0\rangle$.]

11.4. Using the Lehmann representation of $G(\boldsymbol{k},\epsilon)$ and the relation to be proved in the previous problem, show that $G(\boldsymbol{k},\epsilon) \to \epsilon^{-1}$ as $\epsilon \to \infty$.

11.5. Verify that the expression on the right-hand side of eq. (11.173) can equally well be written $T\{Sc_{\boldsymbol{kl}}(t)c_{\boldsymbol{kl}}^\dagger(0)\}$.

11.6. Show that eq. (11.174) is also valid when $t < 0$.

11.7. Draw the following graphs in the manner of those shown in figs. 11.3(d), (e) and (f):


(a) (b) (c)

11.8. If the delta function $\delta(x)$ in the expressions (11.110), (11.111) for the spectral density functions of a non-interacting gas is replaced by the more 'spread-out' function

$$f(x) = \begin{cases} 1/2c, & -c < x < c, \\ 0, & \text{otherwise}, \end{cases}$$

which becomes the delta function in the limit as $c \to 0$, show that the resulting $G(\boldsymbol{k},t)$ obtained from eq. (11.106) is simply the non-interacting $G_0(\boldsymbol{k},t)$ multiplied by $\sin ct/ct$. Initially, therefore, the modulus of $G(\boldsymbol{k},t)$ decreases with time, which is what is expected of a 'quasi-particle' with a finite lifetime. (Substitution of a more realistic function, peaked at $x = 0$, for $\delta(x)$ results in an exponential decay factor, but the mathematics is much more complicated.)

11.9. Give a physical interpretation of the two-particle Green function, defined in eq. (11.193), similar to that for the one-particle Green function given in §11.4.

Bibliography

The following is a short, and by no means exhaustive, list of books, of varying degrees of difficulty, which may be used to augment the present text.

Abrikosov, A. A., Gorkov, L. P. and Dzyaloshinski, I. E., 1963, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Engelwood Cliffs, New Jersey).

Brout, R. and Carruthers, P., 1963, *Lectures on the Many-Electron Problem* (Interscience, New York).

Kittel, C., 1963, *Quantum Theory of Solids* (John Wiley, New York).

March, N. H., Young, W. H. and Sampanthar, S., 1967, *The Many-Body Problem in Quantum Mechanics* (Cambridge University Press, London).

Mattuck, R. D., 1967, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill, New York).

Nozières, P., 1964, *Theory of Interacting Fermi Systems* (Benjamin, New York).

Pines, D., 1961, *The Many-Body Problem* (Benjamin, New York).

Pines, D., 1963, *Elementary Excitations in Solids* (Benjamin, New York).

Raimes, S., 1961, *The Wave Mechanics of Electrons in Metals* (North-Holland, Amsterdam).

Roman, P., 1965, *Advanced Quantum Theory* (Addison-Wesley, Reading, Mass.).

Schultz, T. D., 1964, *Quantum Field Theory and the Many-Body Problem* (Gordon and Breach, New York).

Thouless, D. J., 1961, *The Quantum Mechanics of Many-Body Systems* (Academic Press, New York).

Ziman, J. M., 1969, *Elements of Advanced Quantum Theory* (Cambridge University Press, London).

INDEX