the result of this special case. In Table II the numerical values of the terms given in (60) are tabulated, where ΔE(single) is the energy difference given by (60). From the results shown in Table II, we see at once that:

(a) The values of λ are nearly equal to S, that is, λ is a first-order quantity.

(b) The magnitude of ΔE(single) is proportional to S^4 or S^5. Our prediction that it may be a fourth-order quantity is approximately correct.

(c) In the MnO crystal the overlap integral between two orbitals $(3d0)$ and $(2p0)$ is expected to be 0.05–0.1. Therefore, we will take values of S which correspond to $a/a_0 = 7$, 8, and 9.

(d) The contribution from ΔE_4 is rather small.

(e) The most important contribution comes from the term $2(2\lambda S + \lambda^2)[B\langle AB\rangle - (AB|BA)]$; however, this contribution is at most half of the total value.

(f) The contribution from the term $\{2(2\lambda S + \lambda^2) \times (X^2/Y) - \Omega_{AF}\}$ is rather large.

In Table III the numerical values of the terms given in (65) are tabulated, where ΔE(config.) is the energy difference given by (65). It is shown in Table III that:

(a) Under the assumption that the $(p-s)$ transition energy is 3 ev, the large contribution from these higher configurations is rather remarkable.

(b) As for the energy, the single determinant gives very similar results for the many configurations. Thus it seems to be a good approximation, at least for the four-electron problem.

(c) As for the probability of the higher configurations, the single determinant gives very poor results. This is not surprising, because the total energy is determined accurately by the minimum condition, but the wave function is not.

ACKNOWLEDGMENTS

Electron Interaction in Solids. General Formulation

PHILIPPE NOZIÈRES AND DAVID PINES
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received April 10, 1957)

A general Hamiltonian formalism is developed to treat from first principles the motion of electrons in solids, including their mutual Coulomb interaction. By a series of canonical transformations it is shown that under suitable circumstances (which obtain in nearly all solids) plasmons (a quantized collective plasma oscillation of the electron gas) represent a well-defined elementary excitation of the solid. The "existence criterion" for plasmons is found to be a high electronic polarizability. Where plasmons exist, we are able to give a satisfactory description of their properties when the majority of the individual electron oscillator strengths correspond to transitions in which the energy change is large or small with respect to the plasmon energy, $\hbar\omega_p = \hbar(4\pi Ne^2/m)^{\frac{1}{2}}$. After the plasmon modes are separated out, the remaining electron interaction is found to be screened, with a range of the order of the interelectronic spacing. The usefulness of this effective Hamiltonian for the calculation of the electronic energy levels and cohesive energy in solids is discussed briefly.

1. INTRODUCTION

THE interaction between electrons in a free-electron gas has been considered recently in a series of papers.[1] There it was shown that the interaction between the electrons could be treated simply provided one recognizes at the outset the existence of collective behavior in the system, the plasma oscillations induced by the Coulomb interaction. The plasmons (the quanta of the plasma oscillations) possess an energy near $\hbar\omega_p$, where $\omega_p^2 = 4\pi Ne^2/m$, N and m being the electron density and mass. They describe almost completely the long-range part of the Coulomb interaction, corresponding to momentum transfers $< k_c$, where k_c is the maximum wave vector for which the plasmons constitute an independent mode of elementary excitation. Because the plasmons are not usually excited ($\hbar\omega_p$ is greater than E_0, the energy of an electron at the top of the Fermi distribution), the long-range part of the interaction is effectively frozen out. There remains a set of electrons interacting via a short-range interaction of range k_c^{-1}. For electron gases

[1] D. Pines and D. Bohm, Phys. Rev. **85**, 338 (1952), hereafter referred to as BP II; D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953), hereafter referred to as BP III; D. Pines, Phys. Rev. **92**, 626 (1953), hereafter referred to as P IV; D. Pines, Proceedings of the Tenth Solvay Congress, (R. Stoops, Bruxelles, 1955); D. Pines, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1955), Vol. 1, p. 367; hereafter referred to as SSP. A quite similar approach has been developed by D. N. Zubarev, J. Exptl. Theoret. Phys. (U.S.S.R.) **25**, 548, (1953) and by N. N. Bogoliubov and D. N. Zubarev, J. Exptl. Theoret. Phys. (U.S.S.R.) **28**, 129 (1955).

of densities equal to those encountered in metals, the screening radius is sufficiently short that the motion of a one-electron-like excitation appears meaningful. In other words, once the plasmons are separated out, the remaining elementary excitations are of an essentially independent electron character, modified somewhat, to be sure, by the short-range screened interaction.

To what extent does the foregoing picture apply to solids? It is not, at first sight, obvious that it should work even for the alkali metals, which are the most nearly "free-electron-like" solids. The difficulty arises from the periodic potential, which leads to the possibility of two kinds of excitations for a solid described in the usual one-electron model. The first kind are the intraband transitions, which, for an alkali metal, are well-represented by a free-electron gas (in which, perhaps, the mass of an electron is taken to be the effective mass, m^*). The second kind are the interband transitions, which are peculiar to the periodic character of the potential, and which fall quite outside the scope of a free-electron model. It is the interband excitations which one might expect could alter the free-electron plasma treatment.

Actually, in most cases they do not, as was first pointed out by Mott.[2] Mott showed, using a semiclassical argument, that if the majority of the interband transitions correspond to excitation frequencies, ω_{n0}, which are small compared to the plasma frequency, ω_p, then the plasmons will be little affected by the interband transitions. Because the plasma frequency is so high, ($\hbar\omega_p \sim 15$ ev for a wide variety of solids), such a situation is frequently encountered in practice. One might, therefore, hope that electron interaction in such solids could be treated in detail by methods similar to those used for the free-electron gas. In this, and succeeding papers, such a theory of electron interaction in solids is presented.

Our principal aim is to discuss the modification of the ground-state properties and the excitation spectrum of electrons in solids brought about by the electron interaction. In this paper we present the general Hamiltonian formulation of the problem. We develop the conditions under which plasmons constitute a well-defined elementary excitation of the system. Where the conditions are satisfied, we transform to a system in which there are N' independent plasmon modes. ($N' < N$, the number of electrons.) There remains a short-range screened interaction between the electrons, plus a set of N' subsidiary conditions on the electron wave functions.

In the following paper,[3] we consider the response of the electrons in the solid to both static and oscillating test charges. We derive the dielectric constant from first principles, and thereby establish the conditions under which local field corrections are negligible. We

also work out the optical properties and obtain the effective interaction between a selected small group of electrons, taking into account the role played by the subsidiary conditions. In the third paper of the series,[4] we discuss the elementary excitations of electrons in solids. We establish from first principles that under suitable conditions there are, in fact, "one-electron-like" elementary excitations in solids, which have a behavior quite similar to that presupposed in the usual one-electron theory of solids. In a fourth paper[5] we consider the excitation of plasmons in solids, and the role that plasmons play in various solid-state phenomena.

A discussion of the influence of the periodic potential on plasmons has been given previously by Wolff,[6] Kanazawa,[7] and Adams.[8] Wolff uses a Hartree approximation to obtain the dispersion relation; the major defect in his treatment is that one cannot derive the conditions under which such an approximation is applicable. (We return to this question in Sec. 3.) Adams has treated two special cases by a Hamiltonian formalism, but did not apparently recognize either the possibility of a more general treatment, or the fact, already mentioned, that plasmons are so frequently only slightly modified by the periodic potential of the ion cores. Kanazawa has carried out a treatment which is closely related to that given here. Our treatment differs from his in that we emphasize the simplifying role played by a generalized "f-sum" rule for the problem. We also consider critically the validity of the Hamiltonian treatment of the problem, and take into account the influence of the core electrons on the valence-electron excitation modes. We show that the treatment goes through when the majority of the electron excitations (either core or valence) has energies either smaller *or larger* than $\hbar\omega_p$. What is required is a separation of the plasmon mode from the individual electron excitation modes.

In order to illustrate under what circumstances the plasmon approach may be expected to succeed, let us consider several classes of solids in order of increasing complexity. The simplest solids are those for which there is a well-defined separation between the valence electrons and the core electrons. By valence electrons in an atom, we mean those electrons in the external unfilled shell: for instance, the three M electrons of Al. In a solid, this is not always a clean-cut definition. For instance, in Zn, we may wonder whether to treat d electrons as core or valence electrons. There are certain solids for which no such ambiguity arises. Such solids are made up of atoms for which the core and valence states are widely separated in energy. When the atoms are brought together to make the solid, no appreciable

[2] N. F. Mott, Proceedings of the Tenth Solvay Congress, (R. Stoops, Bruxelles, 1955).

[3] P. Nozières and D. Pines, [Phys. Rev. 108, 762 (1958)], following paper, hereafter referred to as NP II.

[4] P. Nozières and D. Pines, Phys. Rev. (to be published), hereafter referred to as NP III.

[5] P. Nozières and D. Pines, Phys. Rev. (to be published), hereafter referred to as NP IV.

[6] P. Wolff, Phys. Rev. 92, 18 (1953).

[7] H. Kanazawa, Progr. Theoret. Phys. (Kyoto) 13, 227 (1955).

[8] E. N. Adams II, Phys. Rev. 98, 947 (1955).

admixing of atomic core and valence states takes place, and so it is clear which are the valence electrons in the solid.

Examples of simple valence solids are the light alkali metals, alkaline earths, and the solids just beyond them; for such solids suppose we first neglect the influence of the core electrons. We shall show that the valence-electron plasmons have an energy large compared to characteristic individual electron intraband or inter-band excitation energies. Hence, they represent well-defined excitation modes of the system. Further, if we want to take into account the influence of the core electrons on the valence plasmons, we can do so by introducing the static core polarizability, since the characteristic core excitation frequencies are very large compared to the plasma frequency. By the time one reaches the heavier alkali metals, it is, in fact, desirable to treat the influence of the core electrons on the valence-electron motion, since the core polarizability becomes appreciable. For such cases, however, one can treat the core electrons on an individual-particle model, since their excitation frequency is very large compared to their plasma frequency.

Insulators for which the band gap is relatively small with respect to the valence-electron plasmon energy represent another comparatively simple example. In such cases, the characteristic valence-electron "individual-particle" excitation energy will be comparable to the gap, so that again the valence plasmons constitute an independent excitation mode. They correspond simply to a high-frequency polarization wave in the solid.

Semiconductors have interesting plasmon properties. Here one may have both valence-electron plasmons and conduction-electron plasmons. A plasma treatment is necessary to describe the valence-electron excitations because the valence plasmons possess energies large compared to the individual particle valence-electron interband excitations. Moreover, because the number of conduction electrons is relatively small, the conduction-electron plasma may be easily decoupled from the individual-particle conduction-electron interband excitations and the valence-electron motion. (The conduction-electron plasma quantum is small with respect to the conduction-electron interband excitations, the valence-plasmon energy, and the band gap.) In such cases, then we deal with two kinds of electrons, but their excitations are distinct.

This is no longer the case when we consider the transition metals, the noble metals, or those immediately beyond them. For the transition metals, it is, of course, obvious that a clear distinction between "s" and "d" electrons is out of the question. Consider the noble metals. When we are dealing with energies of the order of 10 ev (the "s" electron plasmon energy in, say Cu or Ag), the distinction between "valence" s electrons and the "core" d electrons is again meaningless. A similar point may be made for the metals just

beyond them in the periodic table. It is for just this group of electrons that the methods developed in this paper are least useful. They provide, perhaps, a qualitative guide, but that is all.

The present paper is largely formal. The main physical discussions are carried out in the following paper, and in NP III, so that the reader interested primarily in a qualitative picture would be well-advised to read the next two sections, and go on to the aforementioned papers. In Sec. 2 we discuss our basic starting Hamiltonian, and derive and discuss the generalized f-sum rule. In 3 we give a simple analysis of plasmons and electron correlation which is based on a study of the equations of motion of the density fluctuations. One may thereby obtain results, equivalent to Mott's, which provide a useful qualitative guide and orientation for what follows.

In 4 we begin our Hamiltonian treatment by introducing plasmon field coordinates and carrying out a canonical transformation which relates these to the appropriate oscillating electronic variables. We consider the possibilities for an accurate treatment of the plasmon-electron interaction, and estimate k_c, and hence the number of independent plasmon modes one may expect to encounter. In 5 we carry out a further canonical transformation which under suitable circumstances leads to a set of essentially independent plasmons. In 6, 7, 8, and 9, we discuss what those circumstances may be, considering the convergence of a series of such canonical transformations, the influence of the short-range electron interaction on the plasmons, the damping of plasmons by individual electron interband transitions, and the importance of core electron-valence electron exchange and correlation. In 10 we summarize the conclusions we are led to in the course of the paper.

2. GENERAL CONSIDERATIONS

Let us consider a solid in which no appreciable admixing of core and valence states occurs. The valence electrons then move in the potential of the periodic array of nuclei, and of all the core electrons. We assume that this interaction is well described by an average Hartree potential $V(\mathbf{r})$, depending only on the positions $\mathbf{r}_i$ of the valence electrons. We thus neglect exchange and correlations between core and valence electrons, which is quite a sensible approximation for our simple "valence" solids. The core electrons are repelled by the valence electrons through forces whose origin is both statistical (exchange potential) and purely electrostatic (correlation potential). Under this influence they are polarized, and their potential upon valence electrons is modified. Our neglect of core-valence exchange and correlations is equivalent to the assumption that the core polarizability is very small; which is true. Later on, we shall take the core polarizability into account, and we shall give a mathematical expression of the extent to which such an approximation applies.

The Hamiltonian of the valence electrons may then be written

$$H = \sum_i \left\{ \frac{p_i^2}{2m} + V(\mathbf{r}_i) \right\} + \sum_k{}' \frac{2\pi e^2}{k^2} (\rho_k \rho_{-k} - N), \quad (2.1)$$

where N is the number of electrons and ρ_k their density fluctuation

$$\rho_k = \sum_i \exp(-i\mathbf{k} \cdot \mathbf{r}_i).$$

By $V(\mathbf{r}_i)$, we mean the potential of the periodic array of nuclei plus a uniform background of negative charge, giving an over-all neutral charge distribution. We consider the interaction of the electrons with one another and with a uniform background of positive charge, the latter compensating exactly the above negative charge. We thus cancel the term with $k=0$ in the electron interaction; we indicate this by a prime on the $\sum$.[9]

We shall find it convenient in what follows to define all operators involving the electrons in terms of the eigenstates Φ_n of the operator

$$H_0 = \sum_i \left\{ \frac{p_i^2}{2m} + V(\mathbf{r}_i) \right\}. \quad (2.2)$$

H_0 is a sum of one-electron operators. Each term has a complete spectrum of eigenstates φ_n, which are the usual Bloch wave functions. The states Φ_n of the whole system are Slater determinants built with the φ_n. The set of Φ_n provides a complete basis of the space of antisymmetric wave functions. In general, the Φ_n have no resemblance to actual electronic states: they simply yield a convenient representation. However, where the effect of the Coulomb interaction on the electronic motion is small, the Φ_n will be very close to the actual physical eigenstates of the system of individual particles (after the collective modes have been separated out). This is a quite useful case. Where it does not obtain, we run into some difficulty since we express all the physical properties in terms of states Φ_n which we cannot obtain experimentally.

The one-electron states φ_n may be divided into two categories:

(1) States which correspond to a set of low-lying levels, filled by the core electrons. Such states are forbidden to the valence gas by the exclusion principle. However, we must keep them in order to get a complete set of φ_n. As a consequence, we shall encounter in our calculations matrix elements involving states Φ_n of the valence gas for which some of the "core levels" are occupied. The corresponding terms are spurious, and arise from the fact that we did not treat correctly the indistinguishability of core and valence electrons. This trouble disappears in a better treatment. In fact, the

importance of these spurious terms just measures the validity of our neglect of core-valence exchange.

(2) Proper valence levels which lie above the "core levels." The N lowest valence levels are filled in the ground state Φ_0 of the valence gas, which may involve one band (Na,K,$\cdots$) or several (Be,Al,$\cdots$).

In the following we shall express all our results in terms of the matrix elements of the ρ_k between the ground state Φ_0 and various excited states Φ_n. Since the ρ_k are sums of one-electron operators, and belong to the representation of the translation group with wave vector $\mathbf{k}$, $(\rho_k)_{0n}$ is nonzero only when the state Φ_n is produced from Φ_0 by excitation of only one electron from the level $\varphi_{\kappa,\nu}$ to the level $\varphi_{\kappa+k,\nu'}$ (κ is an arbitrary wave vector inside the first Brillouin zone, and ν, ν' two arbitrary band indices). We could explicitly label the states Φ_n by the four indices κ, k, ν, ν', but then we must take explicitly into account the exclusion principle in the choice of indices. In order to simplify the notation, we shall keep the condensed expression Φ_n.

For every $\mathbf{k}$, the matrix element of ρ_k satisfies an important f-sum rule. To obtain it, let us calculate in two different ways the expression $\{[H_0,\rho_k],\rho_{-k}\}_{00}$. Since we have no exchange terms in H_0, the commutator $[H_0,\rho_k]$ involves only the kinetic energy. We have

$$[H_0,\rho_k]_{mn} = \left\{ \sum_i \frac{\hbar\mathbf{k}}{m} \cdot (\mathbf{p}_i + \tfrac{1}{2}\hbar\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}_i} \right\}_{mn}$$

$$= \hbar\omega_{mn}(\rho_k)_{mn}, \quad (2.3)$$

where $\hbar\omega_{mn} = (E_m - E_n)$. We may now evaluate the double commutator either directly (thereby obtaining a "c" number), or in our particular representation. Equating the two results, we get

$$\sum_n \hbar\omega_{n0}\{ |(\rho_k)_{0n}|^2 + |(\rho_{-k})_{0n}|^2 \} = \hbar^2 k^2 N/m. \quad (2.4)$$

Each of the terms inside the curly brackets yields a total sum which is even in $\mathbf{k}$.[10] Therefore, if we define a set of oscillator strengths $f_{0n}(k)$ as

$$f_{0n}(k) = \frac{2m}{\hbar k^2} \omega_{n0} |(\rho_k)_{0n}|^2, \quad (2.5)$$

they satisfy the usual f-sum rule

$$\sum_n f_{0n}(k) = N. \quad (2.6)$$

The $f_{0n}(k)$ measure the strength of the coupling of the ground state Φ_0 with various excited states Φ_n. We shall express all our results in terms of them.

We now exhibit certain properties of the f_{0n} which will prove useful in later sections. Consider values of

---

[9] As in BP, we assume that the system is enclosed in a box of unit volume: N is therefore the density of the electrons. Since we shall not make explicit use of the fact that the term $k=0$ is excluded, we shall drop the prime in the summation.

[10] Generally, the reflection R belongs to the space group of the crystal. The states Φ_n and $\Phi_m = R\Phi_n$ are degenerate. As a consequence, $R\Phi_0 = \pm\Phi_0$. One then shows easily that $|(\rho_k)_{0n}|^2 = |(\rho_{-k})_{0m}|^2$. Since $\omega_{0n} = \omega_{0m}$, the sum over n of either of the above terms is even in k. The oscillator strengths have the simple property $f_{0n}(\mathbf{k}) = f_{0m}(-\mathbf{k})$.

k much smaller than the Fermi wave vector k_0. Since the state Φ_n corresponds to excitation of one electron from state $\varphi_{\kappa, \nu}$ to state $\varphi_{\kappa+k, \nu'}$, we have two possibilities:

(1) If $\nu'=\nu$, we have an intraband transition. Then $|(\rho_k)_{0n}|^2$ is equal to $[1+O(k^2)]$. (There are no terms in k because of the normalization requirements on the $\varphi_{\kappa, \nu}$.) ω_{n0} is given by a standard Taylor expansion; we find[11]

$$f_{0n}{}^{\text{intra}}(k) = \frac{2m}{\hbar^2 k^2}\left\{\sum_\alpha k_\alpha \frac{\partial E}{\partial \kappa_\alpha} + \frac{1}{2}\sum_{\alpha,\beta} k_\alpha k_\beta \frac{\partial^2 E}{\partial \kappa_\alpha \partial \kappa_\beta} + O(k^3)\right\}_{\kappa, \nu}. \quad (2.7)$$

(2) If $\nu'\neq\nu$, we have an interband transition. Then $(\rho_k)_{0n}$ is of order k, while ω_{0n} is, to lowest order, k independent. From Eq. (2.3), we obtain

$$(\rho_k)_{0n} = \frac{[\sum_i \mathbf{k}\cdot(\mathbf{p}_i+\frac{1}{2}\hbar\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}_i}]_{0n}}{\hbar\omega_{n0}};$$

to the lowest order in k, we find

$$f_{0n}{}^{\text{inter}}(k) = \frac{2}{mk^2}\frac{|(\mathbf{k}\cdot\mathbf{p}_i)_{0n'}|^2}{\hbar\omega_{n'0}}, \quad (2.8)$$

where $\Phi_{n'}$ is the limit of Φ_n when k goes to zero. ($\Phi_{n'}$ denotes a state in which one electron has made a "vertical" transition, going from the level $\varphi_{\kappa, \nu}$ to the level $\varphi_{\kappa, \nu'}$.)

In (2.7) and (2.8), let us now average over the direction of k, and let $|k|$ go to zero. We obtain the following limits:

$$f_{0n}{}^{\text{intra}}(0) = \frac{m}{3\hbar^2}\sum_\alpha \frac{\partial^2 E}{\partial \kappa_\alpha{}^2} = \frac{1}{3}\sum_\alpha \left(\frac{m}{m^*}\right)_{\alpha\alpha},$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.9)$$
$$f_{0n}{}^{\text{inter}}(0) = \frac{2}{3m}\sum_\alpha \frac{|(p_{i\alpha})_{0n'}|^2}{\hbar\omega_{n'0}},$$

where $(1/m^*)_{\alpha\beta}$ is the effective-mass tensor. (For the isotropic case, the intraband oscillator strength is just m/m^*.) These $f_{0n}(0)$ are just the optical oscillator strengths, averaged also over the direction of k. While the optical f_{0n}'s refer to the coupling of electrons with transverse waves, the present ones refer to the coupling with longitudinal waves. From (2.9) we see that for an isotropic solid, both are identical in the limit of low k. For an anisotropic solid, only their averages are equal. The equality suggests an easy experimental approach to the $f_{0n}(k)$ which will enter our calculations.

[11] When summing over n, one gets results even in k. The linear term of (2.7) therefore disappears in the summation. The "useful" part of $f_{0n}(k)$ is the even part, which is regular at $k=0$.

3. SIMPLE ANALYSIS OF PLASMONS AND ELECTRON CORRELATION

Just as was the case for the free-electron gas (see BP II), considerable insight into the role played by the Coulomb interactions may be obtained by a study of the equations of motion of the ρ_k. We may obtain these easily in our Bloch wave representation. The result is

$$\left(\frac{d^2\rho_k}{dt^2}+\omega_p{}^2\rho_k\right)_{mn} = -\omega_{mn}{}^2(\rho_k)_{mn}$$
$$-\sum_{k'\neq k}\frac{4\pi e^2}{mk'^2}\mathbf{k}\cdot\mathbf{k}'(\rho_{k-k'}\rho_{k'})_{mn}. \quad (3.1)$$

The second term on the right-hand side arises from the Coulomb interaction between the electrons. We have separated out the contribution arising when $\mathbf{k}'=\mathbf{k}$, and transposed it to the left-hand side, in the expectation that in so doing we take into account the major effects arising from the Coulomb interaction. The remaining part represents a nonlinear interaction between different density fluctuations. In the case of the free-electron gas, it was shown to be negligible (random phase approximation) and for the time being we also take it to be small. Such an approximation is difficult to justify by a consideration of the equations of motion.[12] One of the desirable features of the Hamiltonian formulation in the following sections is that, by using it, the approximation is shown to be valid in most cases of interest.

In the linear approximation, then, one sees that the motion of the ρ_k (and hence of the electrons) depends on the electron kinetic energy and the periodic potential through the factor $\omega_{mn}{}^2(k)$ (for a given transition, mn), while its dependence on the Coulomb interaction between the electrons appears in the factor $\omega_p{}^2$. A typical value for ω_p in solids is 15 ev. We see at once that if for given transition, $\omega_{mn}{}^2\ll\omega_p{}^2$, then the Coulomb correlations dominate in determining the behavior of $(d^2\rho_k/dt^2)_{mn}$, whereas when $\omega_p{}^2\ll\omega_{mn}{}^2$, the Coulomb correlations are unimportant in determining $(d^2\rho_k/dt^2)_{mn}$.

The general behavior of the electrons in the solid thus depends on the relative size of the characteristic individual electron excitation frequencies, ω_{mn} and ω_p. Two simple cases may appear.

(1) Almost all transitions connecting the ground state Φ_0 to excited states Φ_n with appreciable matrix elements $(\rho_k)_{0n}$ are such that $\omega_{n0}\gg\omega_p$. Because ω_p is so very high, this is a rather idealized case for long-wavelength excitations which may, however, be met in practice with the solid inert gases. Under such circumstances we may always neglect $\omega_p{}^2$ compared to $\omega_{mn}{}^2$;

[12] A similar difficulty occurs when one attempts to derive the dispersion relation for plasmons by linearizing the Hartree equation. [See P. Wolff,[6] or R. A. Ferrell, Phys. Rev. 107, 450 (1957).] The linearization of the Hartree equations may be shown to be completely equivalent to making the random phase approximation.

the characteristic excitation frequencies are essentially unaffected by the Coulomb interaction, and for all phenomena of interest the electron interaction may be treated as a small perturbation. In other words, the binding forces of the electrons to the lattice are much more important than the electrostatic force between the electrons. There will correspondingly be only a small displacement of charges under an external electric field, and hence the solid will possess a low electronic polarizability.

(2) The opposite case occurs when most of the transitions with appreciable $(\rho_k)_{0n}$ are such that $\omega_{n0} \ll \omega_p$. In such a case, (3.1) may be replaced by the simple approximate operator equation

$$d^2\rho_k/dt^2 + \omega_p^2\rho_k = 0. \tag{3.2}$$

The density fluctuation, ρ_k, oscillates approximately at the free-electron plasma frequency, and we clearly have plasmons as the dominant kth excitation mode. Physically this means that the electrostatic interaction between the electrons is far more important than their binding to the lattice. The Coulomb correlations thus act to increase markedly the characteristic excitation frequency of the electrons, and furthermore tend to produce a fairly unique one in the region of ω_p. We shall see that this is a situation frequently encountered in solids. The plasmon energy depends only slightly on the binding in the lattice because the plasma frequency is so high that during a time of interest $(\sim 1/\omega_p)$, the electron essentially does not know it is bound. In such circumstances we expect the electrons to be very mobile as far as any electrostatic perturbation is concerned; they will respond in such a way as to screen out any charge disturbance, and so display a high polarizability.

We should like to emphasize that in the above considerations we have made no distinction based on the conductivity of a given solid. Thus the existence or nonexistence of plasmons does not depend on whether a solid is a metal, semiconductor, or insulator. The relevant question is whether, for a given momentum, the large matrix elements of ρ_k occur for low-frequency or high-frequency transitions. The irrelevance of a distinction based on conductivity follows from the fact that we are asking a very high-frequency question of the electrons in the solid, and distinctions based on energy gaps of only a few electron volts are meaningless at such frequencies.

Let us consider the second class of solids $(\omega_{n0}^2 \ll \omega_p^2)$ in more detail. The screening of a charge disturbance we have mentioned is really of two kinds. The first corresponds to the influence of intraband transitions; the electrons are free to respond spatially to the perturbation, and so alter to the potential of an external charge q from q/r to $(q/r)\exp(-k_c r)$. The second corresponds to interband transitions (the electrons do not respond spatially) which reduce the potential from q/r to $q/\epsilon r$ where ϵ is the dielectric constant. The latter

reduction will again depend on the Fourier component of the test charge, and will be negligible for $k > k_c$. There is a very close connection between the screening and the plasma oscillations between k_c and the maximum wave vector for which the plasmon represents an independent mode of excitation.

To see this, let us write the Coulomb interaction between the electrons as

$$\sum_{k>k_c} \frac{2\pi e^2}{k^2}\rho_{-k}\rho_k + \sum_{k<k_c} \frac{2\pi e^2}{k^2}\rho_{-k}\rho_k, \tag{3.3}$$

where k_c represents the maximum wave-vector for which the plasmon is an independent mode of excitation. Further, the energy in a plasmon quantum is so high that no plasma quanta will be present under ordinary circumstances of temperature excitation. Since the ρ_k oscillate at $\sim\omega_p$, this means that the low-momentum components of the Coulomb interaction are effectively frozen out. They can participate only if sufficient energy to excite a plasmon is supplied. This is just another way of saying that the Coulomb correlations between the electrons (which give rise to the plasmons) are also such that the electron interaction is screened within a distance k_c^{-1}.

We may also remark that this maximum wave vector k_c in a metal will not be expected to be much larger than it is for a free electron gas of the same density. As we have mentioned, the intraband transitions are aptly described on a free-electron picture; for these $\omega_{n0} \sim \mathbf{k} \cdot \mathbf{p}_i/m$. We expect that the plasmon will no longer be an independent mode of excitation when $\omega_{n0} \sim \omega_p$, since here binding effects are competing equally with Coulomb correlations. Hence, $k_c \lesssim \omega_p/v_0$ for metals (v_0 is the velocity of an electron at the top of the Fermi distribution).

We can also see qualitatively that for solids for which $k_c \sim \omega_p/v_0$, the one-electron Bloch functions may be a fairly good approximation. For such solids, k_c^{-1} is of the order of the interparticle spacing, so that the short-range interaction,

$$\sum_{k>k_c} (2\pi e^2/k^2)\rho_k\rho_{-k},$$

would not be expected to alter the electronic wave functions and energy levels appreciably. We return to this question in NP III.

The foregoing considerations do not tell us that high-frequency transitions $(\omega_{n0} \gg \omega_p)$ may not appreciably alter plasmon behavior. To see that, it is necessary to consider the coupling between the different frequencies at which ρ_k may oscillate (ω_{n0} and ω_p). One way to do this is directly analogous to that used in BP II, that is, find the actual operators, X_k, which, within the random phase approximation, do oscillate at a well-defined frequency. Such an approach is sketched in the Ap-

pendix. The resultant plasmon dispersion relation is

$$1 = \frac{4\pi e^2}{m} \sum_n \frac{f_{0n}}{\omega^2 - \omega_{n0}^2}, \quad (3.4)$$

which is just that derived earlier by Mott. A brief inspection of (3.4) shows that high-frequency transitions may reduce the plasmon energy from $\hbar\omega_p$, but they do not destroy the plasmon concept. We shall not go further into the details of this approach at this time, because such questions are better discussed within the framework of a Hamiltonian treatment, which we begin in the following section. We also postpone discussion of the dispersion relation (3.4).

We should like to mention that fast charged particles passing through a solid constitute a splendid probe to determine the behavior of the ρ_k. The interaction of a charged particle (position R_0, charge Ze) with the valence electrons in a solid is

$$4\pi Z e^2 \sum_k \rho_k \frac{e^{i\mathbf{k}\cdot\mathbf{R}_0}}{k^2},$$

so that the energy and momentum transfer from the particle to the valence electrons is completely determined by the ρ_k. Further, the charged particle has sufficient energy to excite a plasmon, so that its characteristic energy losses furnish a true measure of the nature of the valence-electron excitation spectrum.[13] We return to this question in detail in a subsequent paper.

4. INTRODUCTION OF COLLECTIVE COORDINATES

We have seen in the preceding section that there are many solids for which the plasmons may be expected to exist as a well-defined mode of elementary excitation. Furthermore, in such cases, their introduction may be expected to simplify the description of the remaining electronic modes of excitation. For this reason, we wish to begin by introducing the plasmons. We then inquire into the nature of the electron motion after the correlations which give rise to the plasmon have been taken into account.

Our development of a detailed Hamiltonian formulation of the valence-electron interaction problem will closely parallel that given in BP and SSP. (A slightly different version may be found in the article of Kanazawa.[7]) We begin by describing the plasmons explicitly in terms of a set of new conjugate variables, P_k, Q_k. The field variables may be chosen to be Hermitean,[14] so that $P_k = P_{-k}$ and $Q_k = Q_{-k}$. Their

commutation rules are

$$[P_k, Q_{k'}] = -i\hbar\delta_{k, k'}.$$

The P_k and Q_k commute with all electron variables. Because we anticipate that plasmons exist as an independent excitation only up to a maximum wave vector k_c, we wish to introduce only a limited number of plasmon degrees of freedom, N' in number, where $N' = k_c^3/6\pi^2$.

We introduce the plasmons by adding certain terms to our basic Hamiltonian (2.1). The resultant extended Hamiltonian is given by

$$H_{\text{ext}} = H_0 + \sum_{k>k_c} \tfrac{1}{2} M_k^2 \rho_k \rho_{-k}$$
$$+ \sum_{k<k_c} \frac{(M_k\rho_k + P_k)(M_k\rho_{-k} + P_{-k})}{2}$$
$$- \sum_{\text{all } k} \tfrac{1}{2} N M_k^2, \quad (4.1)$$

where M_k is defined by $M_k^2 = 4\pi e^2/k^2$. The added terms describe the inertia of the plasmons, and their coupling with the electrons. We may guarantee that the energy spectrum and the number of independent degrees of freedom of H_{ext} are identical with H_0 by imposing a set of N' subsidiary conditions on the wave function Ψ of the extended system,[15]

$$P_k\Psi = 0 \quad (\text{when } k < k_c). \quad (4.2)$$

For any wave function Ψ satisfying the conditions (4.2), the expectation value of the extended Hamiltonian is the same as the one of the original Hamiltonian. Therefore our procedure does not modify the energy spectrum of the physical system.

Just as was the case for the BP treatment of the free electron gas, for many applications it is not necessary that we work with wave functions which satisfy the subsidiary conditions. It was shown in BP III, and discussed in some detail by Bohm, Huang, and Pines,[16] that with the particular choice of added terms in (4.1), (1) H_{ext} is positive definite, and (2) the subsidiary conditions are automatically satisfied for the ground state of the extended system, so that the ground state is identical with the ground state of H_0. We refer the interested reader to the article of Bohm, Huang, and Pines[16] for a more detailed discussion of the role played by the subsidiary conditions in the free-electron case. In what follows, we shall continue to follow the subsidiary conditions, in order to evaluate their role for the present problem.

We now relate the P_k to the density fluctuations ρ_k, and thereby redescribe the long-range part of the

[13] For a recent discussion of the experimental knowledge of this excitation spectrum see D. Pines, Revs. Modern Phys. **28**, 184 (1956).

[14] Note that in BP and SSP the field variables were taken to be anti-Hermitean.

[15] We use the following notation: φ_n or $\varphi_{\kappa, \nu}$ represents a one-electron level; Φ_n is a Slater determinant for N noninteracting particles, and is a wave function of the original system; Ψ is a wave function of the extended system which has $(3N+N')$ degrees of freedom.

[16] Bohm, Huang, and Pines, Phys. Rev. **107**, 71 (1957), hereafter referred to as BHP.

Coulomb interaction in terms of field variables. This is easily done by the canonical transformation

$$\Psi_{\text{old}} = \exp(iS/\hbar)\Psi_{\text{new}},$$

generated by the Hermitean operator S defined by

$$S = \sum_{k<k_c} \{-M_k\rho_k Q_k\}. \tag{4.3}$$

Since ρ_k commutes with all terms in H but the kinetic energy, we obtain exactly the same result as for free electrons. The new Hamiltonian is

$$H_{\text{ext}} = H_0 + \sum_{k<k_c} \tfrac{1}{2}(P_k P_{-k} + \omega_p^2 Q_k Q_{-k})$$
$$+ H_{\text{int}} + U + H_{\text{sr}} - \sum_{k<k_c} \tfrac{1}{2}M_k^2 N, \tag{4.4}$$

where the terms have the following meanings:

(1) H_{sr} is the short-range part of the Coulomb interaction,

$$H_{\text{sr}} = \sum_{k>k_c} \tfrac{1}{2}M_k^2\{\rho_k\rho_{-k} - N\}. \tag{4.5}$$

(2) H_{int} describes a linear interaction between plasmons and electrons, given by

$$H_{\text{int}} = \sum_{k<k_c} \left\{ -i\frac{M_k}{\hbar}[H_0,\rho_k]Q_k \right\}$$

$$= \sum_{k<k_c} i\frac{M_k}{m} \sum_i \{[\exp(-i\mathbf{k}\cdot\mathbf{r}_i)]\mathbf{k}\cdot(\mathbf{p}_i - \tfrac{1}{2}\hbar\mathbf{k})\}Q_k. \tag{4.6}$$

We remark that H_{int} may be written in the general form

$$H_{\text{int}} = \sum_{k<k_c} V_k Q_k, \tag{4.7}$$

where V_k is an operator depending on electron coordinates and momenta, the representation of which is simply

$$(V_k)_{mn} = -iM_k\omega_{mn}(\rho_k)_{mn}.$$

(3) U describes a nonlinear interaction between plasmons and electrons:

$$U = \sum_{\substack{k,l<k_c \\ k\neq -l}} \left\{ \frac{-M_k M_l}{2m}\rho_{k+l}(\mathbf{k}\cdot\mathbf{l})Q_k Q_l \right\}. \tag{4.8}$$

U is the so-called "random phase" term, because it involves the factor ρ_{k+l}, depending on the phase of all individual particles.

The subsidiary conditions are modified by the canonical transformation, and become[17]

$$(P_k - M_k\rho_k)\Psi = 0 \quad (k<k_c). \tag{4.9}$$

The Hamiltonian (4.4) describes a set of N electrons, interacting with a screened Coulomb interaction of range k_c^{-1}, and a collection of plasmons of maximum momentum k_c. The two systems are coupled through the terms H_{int} and U. The strength of this coupling will essentially determine k_c, since it is only meaningful for us to introduce "reasonably" independent plasmon modes of excitation. We assume that we can treat H_{int} and U independently of H_{sr}, so that we can estimate the strength of the plasmon-electron coupling in the representation in which H_0 is diagonal. Such will not always be the case, as we shall see.

The plasmon-electron coupling brought about by H_{int} may be estimated either through the shift in the energy it produces, or through the admixture of electronic excited states in the wave function of each plasmon. We calculate these quantities by ordinary perturbation theory, a procedure whose consistency we shall then be in a position to establish. The energy shift is

$$\Delta E_{\text{int}} = \sum_{k<k_c} \Delta E_{\text{int}}(k) = -\sum_{k<k_c}\sum_n \frac{|(V_k)_{0n}|^2}{\omega_p + \omega_{n0}}\frac{\hbar}{2\omega_p}.$$

The energy shift $\Delta E_{\text{int}}(k)$ arising from the plasmon k is easily expressed in terms of the oscillator strengths:

$$\Delta E_{\text{int}}(k) = -\frac{\hbar\omega_p}{4}\sum_n \frac{1}{N}f_{0n}(k)\frac{\omega_{n0}}{\omega_{n0}+\omega_p}. \tag{4.10}$$

In order for the plasmon k to be little modified by its interaction with electrons, the energy shift $\Delta E_{\text{int}}(k)$ must be much smaller than the ground state energy $\tfrac{1}{2}\hbar\omega_p$, which yields the condition:

$$\tfrac{1}{2}\sum_n f_{0n}(k)\frac{\omega_{n0}}{\omega_{n0}+\omega_p} \ll N. \tag{4.11}$$

Comparing with the f-sum rule, we see that (4.11) will be satisfied if almost all the oscillator strength corresponds to transitions to excited states Φ_n for which $\omega_{n0}\ll\omega_p$. We shall see that this case occurs frequently among actual solids. On the contrary, if an appreciable amount of oscillator strengths corresponds to transitions for which $\omega_{n0}>\omega_p$, the modification of the plasmon k due to H_{int} will be large. These conclusions do not depend on the intraband or interband character of the transition leading to state Φ_n.[18]

Actually, as we shall see later, the relative magnitude of the energy shift given in (4.10) is a sufficient, but not a necessary, measure of the validity of a perturbation treatment. What is really the crucial test is the admixture of excited states into the ground state

---

[17] Equations (4.4) and (4.9) have a simple physical interpretation. Equation (4.4) is the Hamiltonian of the electrons in interaction with a field of longitudinal photons (which are, in fact, the plasmons). Equation (4.9) is the Poisson equation for longitudinal electromagnetic waves.

[18] For instance, the spurious states Φ_n with some of the core levels occupied by valence electrons, correspond to ω_{n0} negative and much larger than ω_p. Their contribution will be unimportant if they involve only small oscillator strengths, which is generally the case. Apart from these special states, all states Φ_n correspond to $\omega_{n0}>0$.

of each plasmon k. This is measured by the dimensionless quantity

$$\sum_n \frac{|[H_{\text{int}}(k)]_{n0}|^2}{(E_n - E_0)^2} = \sum_n \frac{1}{4N} f_{0n}(k) \frac{\omega_p \omega_{n0}}{(\omega_p + \omega_{n0})^2}. \quad (4.12)$$

If the right-hand member of (4.12) is much smaller than 1, the admixture of excited states is small, and the perturbation treatment converges well. This requires that almost all the oscillator strengths correspond to ω_{n0} much smaller *or much larger* than ω_p. The perturbation method may work, even with a large energy shift $\Delta E_{\text{int}}(k)$. We shall discuss this point in more detail later.

According to either (4.11) or (4.12), the plasmon-electron coupling is k dependent. Hence, although H_{int} may be a weak perturbation for small k, it will strongly couple plasmons and electrons for a sufficiently large value of k. Physically we expect that k_c is determined by the plasmon dynamics, in that it should correspond to the value of k beyond which the plasmon may no longer be regarded as a simple, independent excitation of the system. We shall see that this choice of k_c is essentially equivalent to the "energetic" criterion introduced by BP, namely that it be energetically favorable compared to the Hartree-Fock approximation to introduce the kth plasmon mode.[19]

Let us assume that H_{sr}, H_{int} and U are small perturbations: we then choose the ground state Ψ_0 of the extended system to be the product of Φ_0 (ground state of H_0) by the ground-state wave function of the set of N' independent oscillators. What happens to the energy of the system when we introduce the kth plasmon mode? We replace $(2\pi e^2/k^2)\rho_{-k}\rho_k$ by $\{\frac{1}{2}(P_{-k}P_k + \omega_p^2 Q_{-k}Q_k) + V_k Q_k\}$. The change in energy is given by

$$\delta E(k) = \frac{\hbar\omega_p}{2} - \sum_n \left\{ \frac{2\pi e^2}{k^2} \frac{\omega_{n0}^2}{\omega_p(\omega_p + \omega_{n0})} |(\rho_k)_{0n}|^2 \right\}$$

$$- \sum_n \left\{ \frac{2\pi e^2}{k^2} |(\rho_k)_{0n}|^2 \right\}. \quad (4.13)$$

The first term is the zero-point energy of the new plasmon. The second term is its interaction energy with the electrons (evaluated to second order). The third term represents the energy (evaluated in lowest order) associated with the interaction $(2\pi e^2/k^2)\rho_k\rho_{-k}$ which we have redescribed in terms of the plasmons. We neglected the contribution from U, which we shall show to be very small. It is energetically favorable to introduce the kth plasmon only if $\delta E(k)$ is negative. Ex-

---

[19] It must be pointed out that this procedure is not quite the usual variational procedure of quantum mechanics. We consider a family of extended systems, one per value of N'. Then we select the system which yields the lowest total ground state energy. (While in the usual technique, one minimizes the energy of only one system with respect to different approximate trial wave functions.) This procedure is extensively discussed by Bohm, Huang, and Pines.

pressed in terms of the oscillator strengths, this criterion is

$$\sum_n f_{0n}(k) \frac{\omega_p^2}{\omega_{n0}(\omega_{n0} + \omega_p)} > N. \quad (4.14)$$

If the left-hand member of (4.14) is smaller than N for $k = 0$, a collective treatment does not appear fruitful. On the other hand, if it is larger than N for $k = 0$, we may introduce plasmons up to that k_c for which $\delta E(k) = 0$, or

$$\sum_n f_{0n}(k_c) \frac{\omega_p^2}{\omega_{n0}(\omega_{n0} + \omega_p)} = N. \quad (4.15)$$

We see that (4.14) implies an appreciable amount of transitions such that $\omega_{n0} \ll \omega_p$. In solids with nonfilled bands (metals), this is always realized for low enough k, for the intraband ω_{n0} go to zero with k. Therefore, it is always useful to introduce plasmons for metals. The value of k_c depends of course on the metal we consider.

When k is equal to k_c, the average value of ω_{n0} is of the order of ω_p. That is precisely when our perturbation treatment of H_{int} loses its reliability. Therefore, Eq. (4.15) defining k_c is only qualitatively correct. In fact, this does not matter, for the cutoff is certainly not sharp: there is no exact k_c. The fact that at $k = k_c$ the electron-plasmon coupling becomes strong indicates that the damping of the plasmons by the electrons is becoming appreciable. Our "energetic" criterion is thus seen to be equivalent to the dynamical criterion that the plasmons represent an independent elementary excitation of the system.

In order to evaluate k_c explicitly, let us neglect the contribution arising from the electron-plasmon interaction. (In so doing, we underestimate k_c.) Equation (4.15) then becomes

$$\frac{1}{2}\hbar\omega_p = \frac{2\pi e^2}{k_c^2} \sum_n |[\rho(k_c)]_{0n}|^2. \quad (4.16)$$

The right-hand side is the sum of two terms: the positive self energy, and the negative exchange energy, both for momentum k_c, We may obtain an approximate value for k_c by using the free-electron ·value for the exchange contribution. (This will be an excellent approximation in both the tight-binding and weak-binding case.) We then obtain just the value determined in SSP,

$$k_c/k_0 = 0.353 r_s^{\frac{1}{2}}, \quad (4.17)$$

where k_0 is the Fermi momentum of a gas of N free electrons per unit volume, and r_s their average spacing, in Bohr radii a_0, defined by

$$N = (\tfrac{4}{3}\pi r_s^3 a_0^3)^{-1}. \quad (4.18)$$

With the inclusion of the free-electron value for the plasmon-electron coupling, k_c is increased slightly to the value $(k_c/k_0) \sim 0.4 r_s^{\frac{1}{2}}$. For densities N usually encountered in solids, this determination of k_c gives a

value of $N'/3N$ much smaller than 1: the collective modes represent a small fraction of the total set of degrees of freedom.

Let us now consider U. Let us first assume that we have taken a k_c such that part of the sphere of radius k_c lies outside the first Brillouin zone. We then may find a pair of vectors $\mathbf{k}$, $\mathbf{l}$, smaller than k_c, such that $\mathbf{k}+\mathbf{l}=\mathbf{K}$ where $\mathbf{K}$ is a vector of the reciprocal lattice. Since ρ_k has a nonzero expectation value for the ground state Φ_0 (corresponding to the periodic variation of the electron density), this means that U may induce transitions involving two plasmons, with no participation of the electrons. This would lead to direct coupling between different plasma modes. Such terms would seriously complicate the calculation. We shall consider only cases for which they do not appear (k_c within the first Brillouin zone). This appears to be the case for almost all solids. As a matter of fact, even if such terms exist, they will be relatively unimportant provided that $(\rho_k)_{00}$ is small; in such a case, the electron wave function is very smooth, and may be well approximated by free-electron wave functions, for which no such trouble arises.

Knowing k_c, we are able to evaluate the contribution to the energy arising from U. A simple perturbation treatment gives

$$\Delta E_U = -\sum_{\substack{k,l<k_c \\ k\neq -l}} \frac{(4\pi e^2)^2}{2m^2} \frac{\hbar^2}{4\omega_p} \frac{(\mathbf{k}\cdot\mathbf{l})^2}{k^2 l^2} \sum_n \frac{|(\rho_{k+l})_{0n}|^2}{\hbar(2\omega_p+\omega_{n0})}. \quad (4.19)$$

Since ω_{n0} is positive, we overestimate ΔE_U if we neglect ω_{n0} in the denominator of (4.19). Furthermore, $\sum_n |(\rho_{k+l})_{0n}|^2$ is certainly smaller than N (since it is equal to N *minus* an exchange term). Therefore ΔE_U is bounded by

$$\Delta E_U < N'(N'/N)(\hbar\omega_p/48). \quad (4.20)$$

N' is always appreciably smaller than N. This, together with the very small geometrical factor, insures that the energy shift per plasmon arising from U is negligible compared to $\hbar\omega_p$. The situation is in fact even better. When $|\mathbf{k}+\mathbf{l}|<k_c$, ρ_{k+l} is subject to the subsidiary condition, which markedly reduces its matrix element. Replacing ρ_{k+l} by P_{k+l}/M_{k+l}, and again calculating ΔE_U, we find

$$\Delta E_U = \frac{\hbar^2 k_c^2}{120m} \frac{N'^2}{N} = \frac{1}{150}\left(\frac{k_c}{k_0}\right) N' \frac{N'}{N}\hbar\omega_p, \quad (4.21)$$

where we have used the relation $(k_c/k_0)=0.4r_s$. Since many of the $|\mathbf{k}+\mathbf{l}|$ are larger than k_c, the actual numerical coefficient lies somewhere between 1/48 and $(1/150)(k_c/k_0)$.[20] In any case, ΔE_U is negligible.

---

[20] Actually the criteria are essentially identical, because the inclusion of exchange reduces the estimate in (4.20) by a factor of k_c/k_0.

5. TRANSFORMATION TO INDEPENDENT PLASMON MODES

In this section we give a more careful treatment of the influence of the electron-plasmon interaction by using a canonical transformation to decouple the plasmons from the electrons. Such a treatment is desirable for several reasons. First, as pointed out in BHP, it is necessary to carry out a systematic approximation procedure, in which the subsidiary conditions are treated to the same order of approximation as the Hamiltonian. Such a procedure is most easily carried out using a canonical transformation. Second, using canonical transformations, it is possible to deal with interactions which give rise to a large frequency shift. The transformation described below is specifically designed to handle such interactions. Finally, the use of a transformation enables us to separate easily the effects of the interaction on the plasmons and on the electrons.

Let us write the Hamiltonian in the following way:

$$H = H_0 + H_{\text{field}} + H_{\text{int}} + H_{\text{sr}} + \sum_{k<k_c} (\omega_p^2 - \omega^2)$$
$$\times (\tfrac{1}{2}Q_k^*Q_k) - \sum_{k<k_c} \tfrac{1}{2}NM_k^2, \quad (5.1)$$

where

$$H_{\text{field}} = \sum_{k<k_c} \tfrac{1}{2}(P_k^*P_k + \omega^2 Q_k^*Q_k).$$

The quantity ω is an unknown parameter, eventually k dependent, which we shall choose later to be the corrected plasmon frequency for momentum k. We want to find a canonical transformation, generated by an operator S', which cancels H_{int} to first order, i.e., such that

$$(i/\hbar)[H_0 + H_{\text{field}}, S'] = -H_{\text{int}}. \quad (5.2)$$

For the time being, let us neglect the terms arising from H_{sr}, such as $(i/\hbar)[H_{\text{sr}}, S']$. We discuss such terms in Sec. 7. One verifies easily that the second-order terms arising from $(H_0 + H_{\text{field}})$ just cancel one-half of the first-order term arising from H_{int}. The total second-order term is, therefore,

$$(i/2\hbar)[H_{\text{int}}, S']. \quad (5.3)$$

This will contain terms quadratic in plasmon variables, and terms independent of plasmon variables. We shall choose ω in such a way that the quadratic terms cancel the term $\sum_{k<k_c}(\omega_p^2 - \omega^2)(\tfrac{1}{2}Q_\kappa Q_k)$ in (5.1). It is this procedure which enables us to improve on conventional perturbation theory techniques and treat large energy shifts.

In order to realize our program, let us choose the following generator S':

$$S' = \sum_{k<k_c} (A_k P_k^* + B_k Q_k), \quad (5.4)$$

where A_k and B_k are unknown operators depending on electrons coordinates and momenta. Simple algebra

shows that (5.2) is satisfied if A_k and B_k have the following representation:

$$(A_k)_{mn} = M_k \left(\frac{-i\omega_{mn}}{\omega^2 - \omega_{mn}^2} \right) (\rho_k)_{mn},$$

$$(B_k)_{mn} = M_k \left(\frac{-\omega_{mn}^2}{\omega^2 - \omega_{mn}^2} \right) (\rho_k)_{mn}. \quad (5.5)$$

This definition works as long as $\omega_{mn}^2 \neq \omega^2$. The corresponding transitions (from Φ_m to Φ_n) are virtual, and give rise to the energy shift. On the contrary, $(A_k)_{mn}$ and $(B_k)_{mn}$ are infinite when $\omega_{mn}^2 = \omega^2$: such transitions, conserving the energy, are real, and contribute to the finite lifetime of the state Φ_m. Unless we use a Wigner-Weisskopf treatment, the corresponding part of H_{int} has to be handled separately, by a time-dependent perturbation method. To do so, let us split all electron operators X into two parts

$$X = \bar{X} + \tilde{X},$$

with $\bar{X}$ defined by

$$\begin{aligned} \bar{X}_{mn} &= X_{mn} \quad \text{if} \quad |(\omega - |\omega_{mn}|)| > \eta, \\ \bar{X}_{mn} &= 0 \quad \text{if} \quad |(\omega - |\omega_{mn}|)| < \eta. \end{aligned} \quad (5.6)$$

We may expect that η is related to the lifetime for plasmon decay. We discuss later its choice in some detail. We now replace the generator S' by its "principal part" $\bar{S}'$:

$$\bar{S}' = \sum_{k<k_c} \{\bar{A}_k P_{-k} + \bar{B}_k Q_k\}.$$

To first order, the canonical transformation then cancels the principal part $\bar{H}_{\text{int}}$ of the plasmon-electron interaction, leaving unchanged the "scattering part" $\tilde{H}_{\text{int}}$.

We now proceed to look at the second-order terms: they are now given by

$$\frac{i}{2\hbar}[\bar{H}_{\text{int}}, \bar{S}'] + \frac{i}{\hbar}[\tilde{H}_{\text{int}}, \bar{S}']. \quad (5.7)$$

Using 5.4, we may write them as

$$\frac{i}{4\hbar} \sum_{k,l<k_c} ([V_k + \tilde{V}_k, \bar{A}_l]\{Q_k, P_{-l}\} + [Q_k, P_{-l}]$$

$$\times \{V_k + \tilde{V}_k, \bar{A}_l\} + 2Q_k Q_l [V_k + \tilde{V}_k, \bar{B}_l]), \quad (5.8)$$

where the square brackets denote commutators and the curly ones anticommutators. Let us consider first the term independent of plasmon variables, which is

$$-\tfrac{1}{4} \sum_{k<k_c} [(V_k + \tilde{V}_k)\bar{A}_{-k} + \bar{A}_{-k}(V_k + \tilde{V}_k)] = H_{\text{rp}}. \quad (5.9)$$

H_{rp} is the screened long-range interaction of electrons which arises through the virtual exchange of plasmons. A typical term of H_{rp} is the product of two "one-electron" operators, such as $V_k \bar{A}_{-k}$, whose matrix
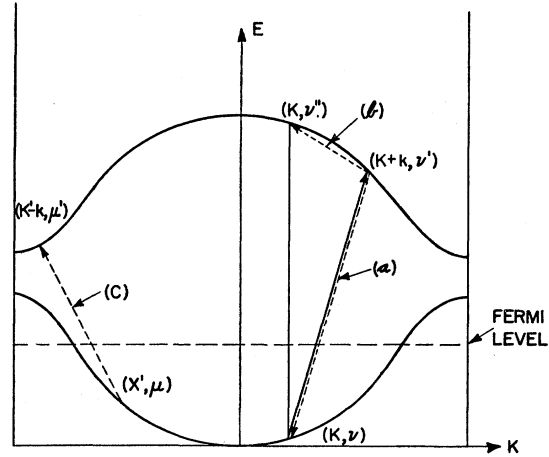

FIG. 1. The different states n for which

$$(V_k A_{-k})_{0n} = \sum_p (V_k)_{0p} (A_{-k})_{pn}$$

is not zero. In state p, one electron has jumped from (κ,ν) to $(\kappa+k, \nu')$. There are three possibilities for state n: (a) the electron $(\kappa+k, \nu')$ goes back to (κ,ν) restoring Φ_0; (b) the electron $(\kappa+k, \nu')$ goes to (κ,ν'') with $\nu'' \neq \nu$. Then Φ_n has only one electron excited "vertically"; (c) another electron is excited from (κ',μ) to $(\kappa'-k, \mu')$. Then Φ_n has two electrons excited.

elements are

$$(V_k \bar{A}_{-k})_{0n} = \sum_p (V_k)_{0p} (\bar{A}_{-k})_{pn}.$$

The states Φ_n which give nonzero matrix elements may be separated into three categories:

(a) The ground state Φ_0 itself. [Fig. 1(a).] This corresponds to the expectation value of $V_k \bar{A}_{-k}$, and gives rise to the screened exchange energy and self-energy.

(b) The states Φ_n with only one electron excited. In order to conserve the momentum, Φ_n must correspond to a vertical transition, from the level $\varphi_{\kappa,\nu}$ to the level $\varphi_{\kappa,\nu''}$ [Fig. 1(b)]. These matrix elements of H_{rp} are equivalent to those of a one-electron periodic potential. They just give rise to the screened long-range exchange potential.

(c) Finally, there are excited states Φ_n for which two electrons are excited, with opposite momentum change [Fig. 1(c)]. This is the real "correlation" part of H_{rp}.

Later in this section we shall discuss these various contributions to H_{rp}. Let us now consider the terms quadratic in plasmon variables. In (5.8), let us select first the terms for which $l = -k$:

$$\frac{i}{4\hbar} \sum_{k<k_c} ((Q_k P_k + P_k Q_k)[V_k + \tilde{V}_k, \bar{A}_{-k}]$$

$$+ 2Q_k Q_{-k} [V_k + \tilde{V}_k, \bar{B}_{-k}]). \quad (5.10)$$

We again separate (5.10) into three kinds of terms:

(a) A part involving the expectation value of the electron operators, together with the complete plasmon operators. $[V_k + \tilde{V}_k, \bar{A}_{-k}]_{00}$ turns out to be zero; therefore there are no cross terms in $Q_k P_k$. Expressing $[V_k + \tilde{V}_k, \bar{B}_{-k}]_{00}$ in terms of the oscillator strengths,

we finally obtain

$$\sum_{k<k_c}\frac{\omega_p{}^2}{N}\left\{\sum_n \bar{f}_{0n}(k)\frac{\omega_{n0}{}^2}{\omega^2-\omega_{n0}{}^2}\right\}\frac{Q_kQ_{-k}}{2}. \qquad (5.11)$$

We choose ω in such a way that these terms (5.11) just cancel the left-over terms

$$\sum_{k<k_c}\tfrac{1}{2}(\omega_p{}^2-\omega^2)Q_kQ_{-k}.$$

This leads to the following dispersion relation for the frequency $\omega(k)$ of the plasmon k

$$\sum_n \bar{f}_{0n}(k)\frac{\omega_p{}^2}{\omega^2(k)-\omega_{n0}{}^2}=N \qquad (5.12)$$

[apart from a term involving $\bar{f}_{0n}(k)$, which is small (see footnote 23)].

(b) A part involving the expectation value of the field operators, together with the off-diagonal part of the electron operators. These electron operators, being commutators, are "one-electron" operators since they connect the ground state Φ_0 with excited states Φ_n in which only one electron is excited. The corresponding contribution may be written

$$\sum_{k<k_c}\frac{i}{4\omega}[V_k+\tilde{V}_k, \bar{B}_{-k}]_{\text{off-diagonal}}. \qquad (5.13)$$

The terms have the same structure as the terms (b) of H_{rp}; we shall combine them later.

(c) A part off-diagonal with respect to both electrons and plasmons. This part connects the ground state of the system with states where two plasmons and one electronic excitation are present. Therefore, the terms have exactly the same structure as the "random phase" terms ($\mathbf{l}\neq-\mathbf{k}$) which we shall consider next. But, rather than depending on two wave vectors, $\mathbf{k}$ and $\mathbf{l}$, the terms depend only on one, $\mathbf{k}$. Their contribution to the energy will be $1/N$, smaller than the contribution of the random phase part, and therefore negligible.

The last contribution to the second-order terms is the random phase part ($\mathbf{k}+\mathbf{l}\neq0$) of the quadratic terms. Here again, we assume that ($\mathbf{k}+\mathbf{l}$) cannot be equal to a vector $\mathbf{k}$ of the reciprocal lattice (as in U, this would introduce a direct coupling between plasmons). Let us call this term U', and evaluate its matrix element for the transition emitting one plasmon $\mathbf{k}$, one plasmon $\mathbf{l}$, and exciting the electron gas from state Φ_0 to state Φ_n. This is given by

$$U_{0n}'(k,l)=\frac{M_kM_l}{4\omega}\sum_p\left\{-\frac{\omega_{0p}\omega_{pn}}{\omega-\omega_{pn}}[(\rho_k+\bar{\rho}_k)_{0p}(\bar{\rho}_l)_{pn}\right.$$
$$+(\rho_l+\bar{\rho}_l)_{0p}(\bar{\rho}_k)_{pn}]-\frac{\omega_{0p}\omega_{pn}}{\omega+\omega_{0p}}[(\rho_k+\bar{\rho}_k)_{pn}(\bar{\rho}_l)_{0p}$$
$$\left.+(\rho_l+\bar{\rho}_l)_{pn}(\bar{\rho}_k)_{0p}]\right\}. \qquad (5.14)$$

[In order to shorten the writing, we have neglected the plasmon dispersion and assumed $\omega(k)=\omega(l)=\omega$.]

Let us compare (5.14) with the matrix element of U for the same transition:

$$U_{0n}(k,l)=\frac{M_kM_l}{4\omega}\sum_p\{\omega_{0p}[(\rho_k)_{0p}(\rho_l)_{pn}+(\rho_l)_{0p}(\rho_k)_{pn}]$$
$$-\omega_{pn}[(\rho_k)_{0p}(\rho_l)_{pn}+(\rho_l)_{0p}(\rho_k)_{pn}]\}.$$

In order to compare $U_{0n}'(k,l)$ with $U_{0n}(k,l)$, let us pick the first term of each. One goes from U_{0n} to U_{0n}' by the replacements

$$(\rho_k)_{0p}\to(\rho_k+\bar{\rho}_k)_{0p},$$

$$(\rho_l)_{pn}\to(\bar{\rho}_l)_{pn}\frac{\omega_{pn}}{\omega_{pn}-\omega}.$$

The correction $(\bar{\rho}_k)_{0p}$ is negligible in a qualitative study. (It could only change the energy shift due to U' by a factor smaller than 2.) The important modification comes from the factor $\omega_{pn}/(\omega_{pn}-\omega)$. We shall see later that the treatment is convergent only if the matrix elements $(\bar{\rho}_k)_{pn}$ are small whenever ω_{pn} is of the order of magnitude of ω (for instance, between $\tfrac{1}{2}\omega$ and 2ω). Then $\omega_{pn}/(\omega_{pn}-\omega)$ is smaller than 1 for most of the transitions, and U' gives a smaller energy shift than U. We shall therefore neglect it in all subsequent calculations. It is important to point out that this procedure is justified even if a large number of the ω_{pn} are much larger than ω (in which case there is a large energy shift due to H_{int}). This is another way to state that a perturbation treatment may work, even with a large energy shift.

So far, we have only looked upon the new terms in the Hamiltonian. The set of subsidiary conditions is also modified. The term ρ_k gives rise to first-order terms involving both plasmon and electron coordinates. The off-diagonal part of the electron factor plays the same role as the random phase term in the Hamiltonian. We may therefore neglect it and replace the electron operators by their expectation values. The subsidiary condition then becomes, to first order,

$$\left(M_k\rho_k-P_k+\frac{1}{N}\sum_n\frac{\bar{f}_{0n}(k)\omega_p{}^2}{\omega^2-\omega_{n0}{}^2}P_k-B_k\right)\Psi=0$$

$$\text{for all } k<k_c.$$

Because of the dispersion relation (5.12), the terms in P_k cancel. This insures the self-consistency of our method: when we eliminate the plasmon-electron coupling from the Hamiltonian, we automatically eliminate it from the subsidiary conditions. We furthermore remark that the second-order terms arising from ρ_k cancel one half of the term B_k, so that the final subsidiary condition is actually

$$(M_k\rho_k-\tfrac{1}{2}B_k)\Psi=0 \quad \text{for all } k<k_c. \qquad (5.15)$$

We now have completely performed our canonical transformation up to second order. Let us summarize the results. First, we note that plasmons and electrons move independently. The plasmons are free oscillators, whose frequency is given by the dispersion relation (5.12). The electrons have the Hamiltonian

$$H_{\rm el}=H_0+H_{\rm sr}-\sum_{k<k_c}\tfrac{1}{2}NM_k{}^2$$

$$+E_{\rm exch}{}^{\rm sc}+V_{\rm exch}{}^{\rm sc}+V_{\rm corr}{}^{\rm sc}, \quad (5.16)$$

where the terms have the following meanings:

(a) $E_{\rm exch}{}^{\rm sc}$ is the diagonal part of $H_{\rm rp}$ and represents the screened long-range exchange energy (including the self-energy).

$$E_{\rm exch}{}^{\rm sc}=-\tfrac{1}{4}\hbar\omega_p\sum_{k<k_c}\sum_n \bar{f}_{n0}(k)\frac{\omega_p\omega_{n0}}{\omega^2-\omega_{n0}{}^2}. \quad (5.17)$$

(b) $V_{\rm exch}{}^{\rm sc}$ is the screened long-range exchange potential, which can only induce one-electron transitions. It arises partly from $H_{\rm rp}$ and partly from (5.13). Its matrix elements are given by

$$(V_{\rm exch}{}^{\rm sc})_{0n}=\sum_{k<k_c}\frac{M_k{}^2}{4\omega}\sum_p\Big\{(\rho_k+\bar{\rho}_k)_{0p}(\bar{\rho}_{-k})_{pn}\frac{\omega_{0p}\omega_{pn}}{\omega+\omega_{pn}}$$

$$+(\bar{\rho}_{-k})_{0p}(\rho_k+\bar{\rho}_k)_{pn}\frac{\omega_{0p}\omega_{pn}}{\omega-\omega_{0p}}\Big\}, \quad (5.18)$$

where the state Φ_n has one electron excited [otherwise $(V_{\rm exch}{}^{\rm sc})_{0n}=0$].

(c) Finally, $V_{\rm corr}{}^{\rm sc}$ is the screened long-range correlation potential, involving the momenta and coordinates of electron pairs. It relates Φ_0 only to states Φ_n where two electrons are excited. The matrix element is

$$(V_{\rm corr}{}^{\rm sc})_{0n}=\sum_{k<k_c}\tfrac{1}{4}M_k{}^2\sum_p\Big\{(\rho_k+\bar{\rho}_k)_{0p}(\bar{\rho}_{-k})_{pn}\frac{\omega_{0p}\omega_{pn}}{\omega^2-\omega_{pn}{}^2}$$

$$+(\bar{\rho}_{-k})_{0p}(\rho_k+\bar{\rho}_k)_{pn}\frac{\omega_{0p}\omega_{pn}}{\omega^2-\omega_{0p}{}^2}\Big\}. \quad (5.19)$$

Furthermore, the electronic wave functions must satisfy the subsidiary conditions (5.15).

Let us first discuss the dispersion relation (5.12). If, for a given k, ω_{n0} is on the average small compared to ω, then ω is very close to ω_p, the free-electron plasma frequency (and this irrespective of the intraband or interband nature of the transition leading to Φ_n). If, on the contrary, many $f_{0n}(k)$ correspond to $\omega_{n0}\gg\omega$, the frequency shift is large. For instance, let us assume that in a metal all interband ω_{n0} are much larger than ω. Then, only the intraband part is important in (5.12), and in the limit of low k the frequency ω is given by

$$\omega^2=4\pi Ne^2/m^*.$$

In some cases (for instance, in semiconductors) (5.12)

may have several roots. In what follows, we assume that we are dealing only with the highest frequency solution which is reasonably close to ω_p. The other roots will be discussed in NP III, as will the detailed application of (5.12) to actual solids.[21]

We now turn to the average screened exchange energy, given in (5.17). Let us compare it with the corresponding nonscreened term (including also the self-energy). This is given by

$$E_{\rm exch}=\tfrac{1}{4}\hbar\omega_p\sum_{k<k_c}\sum_n f_{n0}(k)\frac{\omega_p}{\omega_{n0}}, \quad (5.20)$$

(obtained by a simple perturbation treatment). Both $E_{\rm exch}$ and $E_{\rm exch}{}^{\rm sc}$ are the sum of contributions from each excited state Φ_n. We then see that the screening amounts to the following:

The contribution from high-frequency transitions $(\omega_{n0}\gg\omega)$ is practically unchanged.

On the contrary, the contribution from low-frequency transitions $(\omega_{n0}\ll\omega)$ is reduced by a factor $(-\omega_{n0}{}^2/\omega^2)$.

Finally, the transitions for which $\omega_{n0}\sim\omega$ give an enhanced contribution, due to the resonant exchange of virtual plasmons.[22]

The fact that the screening disappears for frequencies ω_{n0} larger than ω is easily understandable: the "relaxation" time needed by the electron gas to screen any disturbance is of the order of ω^{-1} (since ω describes the inertia of electrons when responding to an electric field). When we consider processes faster than ω^{-1}, the electron gas has no time to respond to the excitation, and there is no screening. This may be summarized as follows: those components of the electronic motion for which $\omega_{n0}\gg\omega$ keep a mainly individual behavior [as shown in (5.20)]. On the other hand, those components for which $\omega_{n0}\ll\omega$ display very little individual character [see (5.20)], but on the contrary are frozen into collective oscillations [as shown in the dispersion (5.12), where only low-energy components contribute appreciably].

It is interesting to compare the part of the energy shift carried out by electrons, $E_{\rm exch}{}^{\rm sc}$, with the part taken by plasmons: $\tfrac{1}{2}\sum_{k<k_c}\hbar(\omega-\omega_p)$. Let us assume that $(\omega-\omega_p)\ll\omega_p$: then we can write

$$\tfrac{1}{2}\hbar(\omega-\omega_p)=(\hbar/4\omega)(\omega^2-\omega_p{}^2).$$

Using Eq. (5.11), we obtain the part of the energy shift carried out by plasmons:

$$\Delta E_{\rm pl}=\tfrac{1}{4}\hbar\omega_p\sum_{k<k_c}\frac{1}{N}\sum_n \bar{f}_{0n}(k)\frac{\omega_{n0}{}^2}{\omega^2-\omega_{n0}{}^2}.$$

[21] At first sight, it would appear that there are as many solutions as there are ω_{n0}, each solution corresponding to a frequency lying between successive ω_{n0}. Since we are using principal parts, such solutions are automatically discarded. Our use of principal parts amounts to taking into account the damping of the plasmons, which automatically removes the spurious roots.

[22] Except, of course, when ω_{n0} lies between $\omega-\eta$ and $\omega+\eta$. Then the screened energy component is zero. Inside this range, the electronic transitions do not properly contribute to an exchange energy, but rather to a broadening of the plasmon line. That is why we left them out of the canonical transformation.

The total energy shift due to H_{int} is

$$\Delta E_{int} = \Delta E_{pl} + E_{exch}{}^{sc} = -\tfrac{1}{4}\hbar\omega_p \sum_{k<k_c} \frac{1}{N} \sum_n \bar{f}_{0n}(k)$$

$$\times \frac{\omega_p\omega_{n0} - \omega_{n0}{}^2}{\omega^2 - \omega_{n0}{}^2}. \quad (5.21)$$

If we replace ω by ω_p in the bracket (which is consistent with our previous approximation), we see that ΔE_{int} is just the same as the one found in (4.10), through an elementary perturbation approach.[23] Both ΔE_{pl} and $E_{exch}{}^{sc}$ are sums of terms corresponding to all possible electron virtual transitions. Let us call $\Delta E_{pl}(k,n)$ and $E_{exch}{}^{sc}(k,n)$ such terms. From (5.21), we see at once that

$$\Delta E_{pl}(k,n)/E_{exch}{}^{sc}(k,n) = -\omega_{n0}/\omega_p. \quad (5.22)$$

The high-energy transitions give, therefore, an energy shift almost entirely carried by plasmons, while the energy shift due to low-energy transitions is mostly taken by electrons.

Let us now consider the exchange potential $V_{exch}{}^{sc}$. That is the one-electron long-range exchange potential which must be included in an Hartree-Fock treatment. Let us compare (5.18) with the corresponding expression for the usual "nonscreened" exchange potential[24]:

$$(V_{exch})_{0n} = \sum_{k<k_c} \tfrac{1}{4}M_k{}^2 \sum_p \{ (\rho_k)_{0p}(\rho_{-k})_{pn}$$

$$+ (\rho_{-k})_{0p}(\rho_k)_{pn}\}. \quad (5.23)$$

Let us assume again that η is small, and forget about the operators $\bar{\rho}_k$ (it being understood that we always take the principal part of any divergent term). We then see that (5.18) differs from (5.23) by factors of the type $\omega_{0p}\omega_{pn}/\omega(\omega+\omega_{pn})$. If the average ω_{mn} is much smaller than ω, such factors reduce drastically $V_{exch}{}^{sc}$ compared with V_{exch}, by a factor of order $\langle\omega_{0n}{}^2/\omega^2\rangle_{Av}$. On the other hand, if both ω_{0p} and ω_{pn} are much larger than ω, it would look as if $V_{exch}{}^{sc}$ is much larger than V_{exch}. This is an illusion, for $(V_{exch}{}^{sc})_{0n}$ may also be written in the following way[25]:

$$(V_{exch}{}^{sc})_{0n} = \sum_{k<k_c} \tfrac{1}{4}M_k{}^2 \sum_p \left\{ (\rho_{-k})_{0p}(\rho_k)_{pn}\frac{\omega_{pn}}{\omega-\omega_{0p}} \right.$$

$$\left. - (\rho_k)_{0p}(\rho_{-k})_{pn}\frac{\omega_{0p}}{\omega+\omega_{pn}} \right\}. \quad (5.24)$$

---

[23] In (4.10) there appeared $f_{0n}(k)$ and not $\bar{f}_{0n}(k)$. In fact, we shall see that the treatment only converges if we can choose $\eta\ll\omega$. Then $f_{0n}(k)$ and $\bar{f}_{0n}(k)$ are practically equivalent in a term regular at $\omega_{n0}=\omega$.

[24] The equivalence of (5.23) with the usual expressions of the exchange potential is readily verified expressing explicitly the states p and n in terms of one electron levels $\varphi_{k,\nu}$.

[25] One may write the factor $\omega_{0p}\omega_{pn}/[\omega(\omega+\omega_{pn})]$ as $(\omega_{0p}/\omega) - \omega_{0p}/(\omega+\omega_{pn})$. The first term, when collected in (5.18), gives rise to the expression $[(H_0,\rho_k),\rho_{-k}]_{0n}$, which is zero since the double commutator is a c number. This establishes the equivalence of (5.18) and (5.24).

In this expression, it is obvious that $(V_{exch}{}^{sc})_{0n}$ is of the same order as $(V_{exch})_{0n}$ when the average ω_{mn} is larger than ω. This potential $V_{exch}{}^{sc}$ should be dealt with by a self-consistent method. In the next section, we shall discuss the influence it may have on the electron motion.

Finally, let us consider the screened correlation potential $V_{corr}{}^{sc}$, given in (5.19). The screening appears in the factors $\omega_{0p}\omega_{pn}/(\omega^2-\omega_{pn}{}^2)$. For very weakly bound electrons $(\langle\omega_{0n}{}^2\rangle_{Av}\ll\omega^2)$, this amounts to a large reduction of the original Coulomb potential, by a factor $\langle\omega_{0n}{}^2/\omega^2\rangle_{Av}$, while for tightly bound electrons $(\langle\omega_{0n}{}^2\rangle_{Av}\gg\omega^2)$, $V_{corr}{}^{sc}$ is of the same order of magnitude as V_{corr}. This screened correlation potential may be treated by a perturbation method (the divergence at low k being removed by the screening). In fact, we shall see later that the perturbation series does not converge, and that a more involved treatment is required.

6. DISCUSSION OF THE CONVERGENCE OF THE PERTURBATION TREATMENT

Let us now consider the convergence of our treatment. Trouble may arise from two kinds of matrix elements in H_{int}. The first kind are those for which $\omega_{n0}>\omega$, which give rise to a large frequency shift, and to a screened interaction of the same order of magnitude as the original Coulomb interaction. The simple perturbation approach of Sec. 4 in fact suggests that these terms do not affect the convergence. The second kind are those for which $\omega_{n0}\sim\omega$, which give rise to real transitions and are therefore related to the plasmon line width. We got rid of part of them by a cutoff of width 2η in the spectrum of excited states Φ_n. This leads to the question: is it actually possible to find an η such that the treatment converges well? It is fairly obvious that if only a negligible amount of the oscillator strengths correspond to frequencies $\omega_{n0}\gtrsim\omega$, such terms will be unimportant, and the method will work easily. What is interesting is the question of whether our results are valid under more general conditions, and if so, what these conditions are. In order to decide the matter, we need a "convergence criterion." Let us therefore consider the higher order terms.

The simplest ones are those arising from

$$\Delta H_{field} = \sum_{k<k_c} \tfrac{1}{2}(\omega_p{}^2-\omega^2)Q_kQ_{-k}.$$

Since this term is canceled by the second-order terms arising from H_{int}, we must be careful about the numerical coefficients of the expansion. The third- and fourth-order terms turn out to be

$$\frac{i}{3\hbar}[\Delta H_{field},\bar{S}'] + \frac{1}{4}\left(\frac{i}{\hbar}\right)^2[[\Delta H_{field},\bar{S}'],\bar{S}']. \quad (6.1)$$

The third-order term is an extra plasmon-electron

interaction which we may write

$$\Delta H_{\text{int}} = - \sum_{k < k_c} \tfrac{1}{3}(\omega_p{}^2 - \omega^2)\bar{A}_k Q_k. \qquad (6.2)$$

The matrix element $(\Delta H_{\text{int}})_{0n}$ is smaller than the original $(H_{\text{int}})_{0n}$ by a factor

$$\tfrac{1}{3}\langle(\omega_p{}^2 - \omega^2)/(\omega^2 - \omega_{n0}{}^2)\rangle_{\text{Av}}.$$

If ω is close to ω_p, this factor is always small, and the treatment is certainly convergent. On the other hand, if $(\omega_p - \omega) \sim \omega$ (an appreciable f_{0n} going with $\omega_{n0} \gg \omega$) the factor is roughly of order $\omega^2/(\omega^2 - \omega_{n0}{}^2)$. In this case, for transitions for which ω_{n0} is much smaller than ω, $(\Delta H_{\text{int}})_{0n}$ is about as big as $(H_{\text{int}})_{0n}$. But in just such a case we know that these low-frequency matrix elements essentially do not contribute to the frequency shift; therefore, it does not matter whether $(\Delta H_{\text{int}})_{0n}$ is large or not. On the other hand, $(\Delta H_{\text{int}})_{0n}$ is much smaller than $(H_{\text{int}})_{0n}$ when ω_{n0} is much larger than ω. Consequently, ΔH_{int} is negligible in determining the frequency shift if most of the oscillator strengths correspond to ω_{n0} much larger or much smaller than ω. We may express that by saying that the coupling constant $\Delta\alpha$ corresponding to ΔH_{int} is much smaller than the one α arising from H_{int}.

The next order terms are of two kinds. The first arise from the action of $\bar{S}'$ on ΔH_{int}: they are the fourth-order terms appearing in (6.1), and are of order $\alpha\Delta\alpha$. The second kind occurs when one makes a special perturbation treatment to eliminate ΔH_{int} (with a canonical transformation generated by $\Delta\bar{S}'$). This will yield a higher order term $(i/2\hbar)[\Delta H_{\text{int}}, \Delta\bar{S}']$, of order $\Delta\alpha^2$. If $\Delta\alpha \ll \alpha$, the second contribution is negligible. We therefore only consider the first kind of terms of order $\alpha\Delta\alpha$, whose structure is quite similar to that of the second-order terms. These include, first, an extra contribution to H_{rp}, independent of plasmon variables, which differ from H_{rp} [Eq. (5.9)] only by the replacement of V_k by ΔV_k:

$$(\Delta V_k)_{mn} = -\tfrac{1}{2}\left(\frac{\omega_p{}^2 - \omega^2}{\omega^2 - \omega_{mn}{}^2}\right)(V_k)_{mn}. \qquad (6.3)$$

In addition there are extra terms quadratic in plasmon variables, with a factor which depends on the electron coordinates. As with the second-order terms, we neglect the part of this term which is off diagonal with respect to the electrons. We are then left with a term $\sum_{k < k_c} \tfrac{1}{2}\Delta\omega^2 Q_k Q_{-k}$, where the additional frequency shift $\Delta\omega^2$ is given by

$$\Delta\omega^2 = -\tfrac{1}{2}(\omega_p{}^2 - \omega^2)\frac{1}{N}\sum_n \bar{f}_{0n}(k)\frac{\omega_{n0}{}^2\omega_p{}^2}{(\omega^2 - \omega_{n0}{}^2)^2}. \qquad (6.4)$$

Let us consider the extra screened Coulomb interaction (for the case in which ω is appreciably different from ω_p). From (6.3), we see that the high-energy components of ΔV_k are much smaller than the corresponding ones of V_k. On the contrary, when $\omega_{mn} \ll \omega$

the correction $(\Delta V_k)_{mn}$ is comparable to $(V_k)_{mn}$. Our perturbation determination of the low-energy components of H_{rp} is not convergent. The question now arises: are these low-energy components of H_{rp} of any importance? We know from the free electron case that their role is mainly to introduce the slight long-range correlations necessary in order to satisfy the subsidiary conditions. Therefore, our perturbation treatment does not describe well the mechanism by which the electrons satisfy the subsidiary conditions, but we expect it to be accurate as far as the total energy is concerned.

In any event, the behavior of the plasmons depends only on average properties of the electrons. The only convergence requirement for them is that $\Delta\omega^2/\omega^2$ be much smaller than 1. When $\omega \sim \omega_p$, the requirement is approximately

$$\sum_n \bar{f}_{0n}(k)\frac{\omega^2\omega_{n0}{}^2}{(\omega^2 - \omega_{n0}{}^2)^2} \ll N, \qquad (6.5)$$

the convergence criterion for which we were looking. If ω is much smaller than ω_p, Eq. (6.5) is not sufficiently restrictive, and we must use (6.4). It is obvious that the treatment will diverge only if an appreciable amount of the $f_{0n}(k)$ correspond to frequencies ω_{n0} of the order of ω.

The same criterion may be obtained quickly from a study of the higher order terms in the subsidiary conditions. In order to have convergence, the amount of field variable brought in by the second-order terms must be much smaller than the original P_k term with which we started. When taking the average over electronic coordinates, we find again the condition (6.5).

Up to now, we have considered only the higher order terms arising from ΔH_{field}. There is also a contribution arising from H_{rp}. The only easy way to calculate it is to replace all electronic commutators by their expectation value (which is equivalent to a Hartree approximation). One then finds another extra electron-plasmon interaction, which has the same structure and order of magnitude as ΔH_{int}. The previous discussion still applies. But it is difficult to decide whether this "extended random phase approximation" is a sensible one. Let us advance a simple tentative physical argument: the terms arising from H_{rp} will be negilgible if H_{rp} itself is negligible compared to H_0 and H_{field} (in other words, if H_{rp} acts like a small perturbation on the motion of the system). An easy, if not very precise, way to test this consists in calculating the expectation values of these three operators.

$$|H_0|_{00} \simeq N(E_c + \tfrac{3}{5}E_0),$$

$$|H_{\text{field}}|_{00} = \sum_{k < k_c} \tfrac{1}{2}\hbar\omega,$$

$$|H_{\text{rp}}|_{00} = -\tfrac{1}{4}\hbar\omega_p \sum_{k < k_c} \frac{1}{N}\sum_n f_{n0}(k)\frac{\omega_{n0}\omega_p}{\omega^2 - \omega_{n0}{}^2}, \qquad (6.6)$$

where E_c represents the energy of the bottom of the

valence band, and E_0 is the total width of the occupied bands. (We used the coefficient $\frac{3}{5}$ occurring for free electrons, but the order of magnitude is still right for other solids.) Since ω and ω_p are generally of the same order of magnitude, we may replace ω_p by ω in $|H_{\rm rp}|_{00}$. We then see that $|H_{\rm rp}|_{00}$ will be much smaller than $|H_{\rm field}|_{00}$ if most of the oscillator strengths correspond to ω_{n0} either much larger or much smaller than ω. In this case, we do not expect the higher order terms arising from $H_{\rm rp}$ to influence appreciably the plasmon behavior.

In many solids, the average energy of an electron, $(E_c + \frac{3}{5}E_0)$, turns out to be of the order of the plasmon energy $\hbar\omega$. Then $|H_0|_{00}$ is of order $(N/N')|H_{\rm field}|_{00}$. In such a case, the long-range screened Coulomb energy is very small compared to the one electron energy: we expect $H_{\rm rp}$ to perturb only slightly the electron motion. The higher order terms arising from $H_{\rm rp}$ also do not influence the electron behavior. The fact that $H_{\rm rp}$ is a small perturbation on the electrons suggests that in a Hartree-Fock treatment, one might neglect $V_{\rm exch}{}^{\rm sc}$, or at most treat it as a perturbation.

To summarize this study, let us say that our perturbation method works satisfactorily if the large majority of the $f_{0n}(k)$ correspond to very large or very small ω_{n0}. Especially, it still works when many ω_{n0} are very large (giving a large frequency shift).[26] In this case, the low-energy matrix elements of $H_{\rm rp}$ converge slowly: it does not matter, for their influence on the motion of the system is very small. This confirms the results of our simple treatment of $H_{\rm int}$ carried out in Sec. 4. We furthermore have shown that when these convergence conditions are realized, $H_{\rm rp}$ is only a small perturbation on the electron motion.

One may wonder why the perturbation method is good when there is a large frequency shift. In this case, there are many electron transitions for which $\omega_{n0} \gg \omega$. Our approximations amount to neglecting the effect of the plasmons on these high-energy individual modes, since the plasmons are frozen at such high frequencies, and cannot follow the electron motion. This is just an adiabatic approximation, of the type used by Bardeen and Pines[27] in their collective treatment of the electron-phonon coupling. It is the counterpart of the low-energy approximation ($\omega_{n0} \ll \omega$), in which the electrons cannot follow the plasmons (as for instance in the free electron plasma discussed in BP III).

7. EFFECT OF SHORT-RANGE CORRELATIONS ON THE PLASMONS

In the last two sections, we completely neglected the terms arising from $H_{\rm sr}$ in the canonical transformation

---

[26] It is important to point out that there may exist plasmons for which the frequency shift $(\omega - \omega_p)$ is very large. The existence criterion (4.14) implies that the *harmonic* average of ω_{n0} must be smaller than ω. This may be realized with only a few f_{0n} corresponding to *very* low ω_{n0}: then the *arithmetic* average, which determines $(\omega - \omega_p)$, may be much larger than ω. That is, for instance, what happens for the plasmons of conduction electrons in semiconductors.

[27] J. Bardeen and D. Pines, Phys. Rev. **99**, 1140 (1955).

generated by $\bar{S}'$. The structure of such terms is quite similar to those appearing in U, or the higher order terms arising from U, in that they give rise to coupling between different Fourier components of the Coulomb interaction. In U, a typical term involves $Q_k Q_l \rho_{k+l}$, where l is $< k_c$, while the corresponding term from $H_{\rm sr}$ involves $Q_k \rho_l \rho_{k-l}$, where $l > k_c$. The terms from $H_{\rm sr}$ are, thus, much larger, because the density fluctuations which appear are not screened. In addition, their domain is not restricted by $N' < N$. Hence, if we extend the notion of the random-phase approximation to cover all coupling between density fluctuations of different momenta, the random-phase approximation will fail because of the coupling arising from $H_{\rm sr}$ rather than that of U. The first-order terms arising from $H_{\rm sr}$ are

$$\frac{i}{\hbar} \sum_{k' > k_c} \tfrac{1}{2} M_{k'}{}^2 \{ \rho_{k'}[\rho_{-k'}, \bar{S}'] + [\rho_{-k'}, \bar{S}']\rho_{k'} \}. \quad (7.1)$$

The term (7.1) corresponds to an electron-plasmon interaction in which two electrons scatter, with emission or absorption of a plasmon. In the general case, it is extremely difficult to evaluate its effect. However, let us consider the special simple case in which almost all ω_{n0} are much smaller than ω. Then, we may expand $(\bar{S}')_{0n}$ in powers of ω_{n0}/ω.

$$\bar{S}_{0n}' = \sum_{k < k_c} M_k \left\{ \frac{i\omega_{n0}}{\omega^2} P_{-k} - \frac{\omega_{n0}{}^2}{\omega^2} Q_k + \frac{i\omega_{n0}{}^3}{\omega^4} + P_{-k} + \cdots \right\}.$$

From this, it is trivial to go back to the operator form of $\bar{S}'$,

$$\bar{S}' = -\sum_{k < k_c} M_k \left\{ \frac{i[H_0, \rho_k]}{\hbar\omega^2} P_{-k} \right. $$
$$\left. + \frac{[H_0, [H_0, \rho_k]]}{\hbar^2\omega^2} Q_k + \cdots \right\}. \quad (7.2)$$

Since we know H_0, we may evaluate explicitly the successive commutators in terms of the potential $V(\mathbf{r}_i)$,

$$[H_0, \rho_k] = -\sum_i \left\{ e^{-i\mathbf{k}\cdot\mathbf{r}_i} \frac{\hbar\mathbf{k}}{m} \cdot (\mathbf{p}_i - \tfrac{1}{2}\hbar\mathbf{k}) \right\},$$

$$[H_0, [H_0, \rho_k]] = \sum_i \left\{ e^{-i\mathbf{k}\cdot\mathbf{r}_i} \left[\frac{\hbar\mathbf{k}}{m} \cdot (\mathbf{p}_i - \tfrac{1}{2}\hbar\mathbf{k}) \right]^2 \right. \quad (7.3)$$
$$\left. - \frac{i\hbar^2}{m} e^{-i\mathbf{k}\cdot\mathbf{r}_i} (\mathbf{k}\cdot\boldsymbol{\nabla}V(\mathbf{r}_i)) \right\}.$$

In order to calculate the term (7.1), we take the commutator of $\rho_{k'}$ with the successive terms of S'. The leading terms are independent of the periodic potential,

$V(\mathbf{r}_i)$. The first term of the expansion is

$$H_{\text{int}}^{(1)} = \sum_{k'>k_c} \sum_{k<k_c} M_k M_{k'}{}^2 \frac{\mathbf{k}\cdot\mathbf{k}'}{m\omega^2} \rho_{k'}\rho_{(-k-k')} P_{-k}. \quad (7.4)$$

Changing $\mathbf{k}'$ into $(-\mathbf{k}'-\mathbf{k})$, one sees that (7.4) is actually of order k. The second term of H_{int} is seen by direct inspection to be also of order k. The extra electron-plasmon interaction consists therefore of (a) two terms of order k, giving rise to a shift of ω proportional to k^2, and negligible for low enough k's; (b) higher order terms, independent of k, but smaller than the original H_{int} by a factor $\langle\omega_{0n}{}^2/\omega^2\rangle_{\text{Av}}$. In the limit of very weak plasmon electron coupling, $\langle\omega_{0n}{}^2/\omega^2\rangle_{\text{Av}}\ll1$, the frequency shift arising from these terms is certainly negligible. This result comes from the fact that in such a case H_{sr} is a relatively small perturbation on the electron motion. When making the canonical transformation, the terms arising from H_{sr} are therefore negligible compared to those arising from H_0. The free electron plasma affords an example of such a situation. Let us remark that when k becomes large (of order k_c), the effect of H_{sr} may become important. But then the whole treatment loses its validity.

This conclusion is true only if most of the ω_{n0} are very small, i.e., if the frequency shift $(\omega-\omega_p)$ is small. In the case in which many ω_{n0} are much larger than ω, the plasmons may be very strongly influenced by the short-range electron collisions. In its present shape, our treatment then fails. We may overcome this difficulty by changing our representation. Rather than taking the eigenstates Φ_n of H_0 as a basis, let us choose the eigenstates Φ_M of (H_0+H_{sr}). These states Φ_M are many-body wave functions, where the short-range electron correlations are rigorously taken into account. One might object to this representation, saying that it involves many-body eigenstates, which we do not know how to calculate. In fact, these eigenstates Φ_M are just the true, physical ones. We saw in the preceding section that we quite generally expected H_{rp} to have a negligible influence on the electron motion. Then the final electron Hamiltonian is simply (H_0+H_{sr}), and our Φ_M describe the actual wave functions of the system. The oscillator strengths f_{0M}, the energy differences ω_{0M}, are just those which we can obtain from experiment. Although it does not allow easy theoretical evaluation, the representation in terms of the Φ_M is best fitted for comparison with experiment.

In this representation, the calculations go along exactly as in the former one, but for the fact that we no longer have to worry about H_{sr}. The only condition imposed on the representation is that it give rise to an f-sum rule. This requires that ρ_k commute with the potential $(V+H_{\text{sr}})$ which is obviously true. (It would no longer be so if we replaced H_{sr} by an approximate Hartree-Fock potential.) All results are formally the same as in the preceding sections. Their validity is

subject only to the following two conditions:

Condition of existence: $\displaystyle\sum_M \frac{f_{0M}(k)\omega^2}{\omega_{M0}(\omega+\omega_{M0})} > N.$

$$(7.5)$$

Condition of convergence: $\displaystyle\sum_M \langle f_{0M}\rangle_{\text{Av}}(k)\frac{\omega^2\omega_{M0}{}^2}{\omega^2-\omega_{M0}{}^2}\ll N.$

Such a treatment is rigorous. To be able to handle it, we must approximate the Φ_M. One may, for instance, try to treat H_{sr} by a Hartree-Fock method. The f-sum rule is then only approximate. The validity of such a procedure has to be studied in each particular case. When H_{sr} happens to be a small perturbation (when ω_{n0} is much smaller than ω, for example), the Φ_M are very close to our old Φ_n (eigenstates of H_0 alone). It is then much simpler to keep the one-electron representation in terms of the Φ_n, and, eventually, to treat the effect of short-range collisions by perturbation methods.

8. DAMPING OF PLASMONS

In this section, we consider the damping of plasmons and calculate the line width. We also discuss the condition of convergence of the treatment, and study the validity of the cutoff η which we introduced to get rid of the infinite terms in S'.

For free electrons, there is no direct damping due to H_{int}, because for the BP choice of k_c all ω_{n0} are smaller than ω. As was pointed by BP, the damping arises mainly through short-range collisions of two electrons. One may evaluate the corresponding lifetime, and one finds it very long for long-wavelength plasmons. In actual solids, this effect will still be present. If we use a many-body representation in terms of the eigenstates Φ_M of (H_0+H_{sr}), this will appear as a special case of the direct first-order transitions arising from H_{int}. When the effect of H_{sr} on the electronic motion is small, we prefer to work in the representation in terms of the one-electron eigenfunctions Φ_n of H_0 alone. The effect of the short-range collisions must then be treated separately but we expect it to be small, since H_{sr} is relatively a small perturbation. Therefore, in both cases, we have only to consider the direct first-order damping arising from H_{int}.

The major part of the damping arises from the term $\tilde{H}_{\text{int}}$ which we left untouched during the perturbation treatment. Standard time-dependent perturbation theory gives the probability of having a plasmon k decay from the first excited state to the ground state, with excitation of the electron gas from state Φ_0 to any state Φ_n.[28]

$$W_k = \tfrac{1}{2}\pi\omega_p{}^2 \frac{1}{N} \sum_n f_{0n}(k)\delta(\omega-\omega_{0n}). \quad (8.1)$$

[28] In deriving (8.1), we assume that we do not run into any difficulty by using a description of "one electron" states of high excitation energy $(\omega_{n0}\sim\omega)$. Difficulties may in fact arise through H_{rp} and the subsidiary conditions for such highly excited states.

In such a process, a collective plasmon decays into an individual-particle excitation of the same momentum. While a plasmon is an organized electrostatic wave of polarization, which involves simultaneously many electrons, an individual particle excitation involves only the change of wave function of one electron. (If we use a many-body representation, incorporating H_{sr}, we do not excite one electron, but rather a small group of them, localized within a radius k_c^{-1}. Such an excited state is, nevertheless, quite different from a plasmon.) The decay of the plasmon corresponds to a process in which a given electron (or a small group of them), rather than propagating the electrostatic wave by a slight displacement, suddenly absorbs the whole energy of the wave with a drastic change of its wave function. The long-range order is destroyed, and the collective plasmon disappears.

In order to express simply all the preceding results, let us study more closely the distribution of excited states Φ_n for a given value of the momentum change k. In a one-electron representation, Φ_n corresponds to excitation of an electron from level $\varphi_{\kappa, \nu}$ to level $\varphi_{\kappa+k, \nu'}$. The frequency ω_{n0} will then form bands, labelled by the pair (ν, ν'), with ω_{n0} depending on κ within each band. This distribution of ω_{n0} will have the same structure in a many-body solution. Let us denote by $\partial n_{\nu\nu'}(\kappa)/\partial\Omega$ the density of states Φ_n in the band (ν, ν') at point κ, and per unit ω_{n0}. To a given value Ω of ω_{n0} will correspond a finite number of states Φ_n [several from the same band if $\omega_{n0}(\kappa)$ has extrema]. Let us call these states $\Phi_{\Omega, \alpha}$, where α labels the different degenerate states. All our results may be expressed in terms of the kernel

$$g(\Omega,k) = \frac{1}{N}\sum_\alpha f_{\Omega,\alpha}(k)\left(\frac{dn}{d\Omega}\right)_{\Omega,\alpha},$$

which may be written in more condensed form

$$g(\Omega,k) = \frac{1}{N}\sum_n f_{0n}(k)\delta(\omega_{n0}-\Omega). \quad (8.2)$$

[$g(\Omega,k)$ is piecewise continuous with respect to Ω.] For instance, we have:

$$\int_0^\infty g(\Omega,k)d\Omega = 1, \qquad f\text{-sum rule} \quad (8.3)$$

$$\int_0^\infty g(\Omega,k)\frac{\omega_p^2}{\Omega(\omega_p+\Omega)}d\Omega > 1, \quad \text{existence criterion} \quad (8.4)$$

$$\int_0^{\omega-\eta}\frac{g(\Omega,k)d\Omega}{\omega^2-\Omega^2} + \int_{\omega+\eta}^\infty\frac{g(\Omega,k)d\Omega}{\omega^2-\Omega^2} = \frac{1}{\omega_p^2},$$

dispersion relation (8.5)

_____

In the absence of a completely satisfactory treatment, we propose to assume that the resolution of the foregoing difficulties will yield a result close to (8.1). We return to this question in NP III.

$$\int_0^{\omega-\eta} g(\Omega,k)\frac{\omega^2\Omega^2}{(\omega^2-\Omega^2)^2}d\Omega + \int_{\omega+\eta}^\infty g(\Omega,k)\frac{\omega^2\Omega^2}{(\omega^2-\Omega^2)^2}d\Omega < 1,$$

convergence criterion. (8.6)

The relative width of the line is $1/\omega\tau = W/\omega$, which is simply expressed as

$$\frac{1}{\omega\tau} = \frac{\pi}{2}\frac{\omega_p^2}{\omega}g(\omega). \quad (8.7)$$

We are now able to discuss the convergence of the perturbation treatment in the range $\omega_{n0}\sim\omega$, and to determine the suitable cutoff width, η. The quantity η must satisfy the two requirements:

(a) It must be small compared to ω. Otherwise, $\tilde{H}_{int}$ would give rise, not only to a line broadening, but also to an appreciable shift of the plasmon frequency, ω.

(b) It must be such that the convergence criterion (8.6) be satisfied.

The main contribution to the integral (8.6) comes from Ω close to ω. We therefore calculate the integral replacing $g(\Omega)$ by $g(\omega)$. Using (8.8), we find easily that the convergence criterion is roughly

$$\eta\tau \gg \omega^2/(\pi\omega_p^2). \quad (8.8)$$

Generally, ω and ω_p are of the same order of magnitude. Then the treatment is convergent if

$$1/\tau \ll \eta \ll \omega. \quad (8.9)$$

We must therefore choose η larger than the line width $1/\tau$. This is physically obvious: the divergences arise from real transitions. To get rid of them, we must choose η such as to cut off most of the plasma line "area," and therefore take η larger than the line width.

From (8.9), we draw the general conclusion that the treatment presented in this paper works only for narrow lines. In fact, when the line is broad, it becomes difficult to define its center, and a precise frequency ω is not meaningful. For each characterization of plasmons (energy-loss experiments, etc.), it is then necessary to carry out a special Wigner-Weisskopf treatment in order to get the observed spectrum. This will be done in a later publication.

To summarize the conditions of validity of our treatment, let us say that the method works if $\Omega g(\Omega)$ is much smaller than 1 when Ω is of order of magnitude ω. This insures both a narrow line and a satisfactory convergence of the frequency-shift calculation.

9. INFLUENCE OF THE ION CORE ON THE PLASMONS

We are now able to discuss our neglect of the core-valence exchange and correlation effects. The core electrons are characterized by very large excitation frequencies $(\omega_{n0}\gg\omega)$. If we want to include them in the plasma treatment, we must therefore use the "many-body" representation which we just defined. Further-

more, in order to treat correctly the exchange between core and valence electrons, we must treat in the same way all the electrons, and perform the same canonical transformations on core and valence coordinates. We therefore write the initial Hamiltonian in the following way:

$$H = \mathcal{3C}_0 + \sum_{k<k_c} \tfrac{1}{2} M_k^2 R_k R_{-k}, \qquad (9.1)$$

where R_k is the density fluctuation of all electrons

$$R_k = (\rho_k)_{\text{valence}} + (\rho_k)_{\text{core}},$$

and where $\mathcal{3C}_0$ includes the kinetic energy of all electrons, the potential of the periodic array of nuclei, and the short-range interaction between core-core, core-valence, and valence-valence electrons.

One then introduces the plasmons as if all electrons were free, and obtains a frequency Ω_p given by

$$\Omega_p^2 = 4\pi N_{\text{tot}} e^2/m, \qquad (9.2)$$

where N_{tot} is the total number of electrons. This is far too large, but is drastically reduced when one eliminates the plasmon-electron interaction. Let us call χ_μ the eigenstates of $\mathcal{3C}_0$ (which are many-body wave functions describing all the electrons), and denote by $f_{0\mu}$ and $\omega_{0\mu}$ the corresponding oscillator strengths and energy differences. The f-sum rule is now

$$\sum_\mu f_{0\mu}(k) = N_{\text{tot}}. \qquad (9.3)$$

The dispersion relation then becomes

$$\sum_\mu \frac{f_{0\mu}(k)}{\omega^2 - \omega_{0\mu}^2} = \frac{N_{\text{tot}}}{\Omega_p^2} = \frac{m}{4\pi e^2}. \qquad (9.4)$$

Since we are using a many-body representation, we can no longer speak of pure "core" excitation or pure "valence" excitation. However, the $\omega_{\mu 0}$ will always fall into classes—a class of very high $\omega_{\mu 0}$ corresponding to excitation of a core electron with a small admixture of valence excitation, and a class of reasonably small $\omega_{\mu 0}$ corresponding to the reverse situation, mainly valence excitation. To the extent that we may separate these two categories, we may write the dispersion relation as follows:

$$\frac{4\pi e^2}{m} \left\{ \sum_{\mu(\text{val})} \frac{f_{0\mu}(k)}{\omega^2 - \omega_{0\mu}^2} + \sum_{\mu(\text{core})} \frac{f_{0\mu}(k)}{\omega^2 - \omega_{0\mu}^2} \right\} = 1, \qquad (9.5)$$

where the indices "core" and "val" refer to the above two classes of states χ_μ.

In such a treatment, the spurious transitions of valence electrons to occupied core levels disappear, since we take rigorously into account the exclusion principle. [In fact, the corresponding terms in (9.5) are canceled by the terms corresponding to transition of a core electron to an occupied valence level.] Neglecting the core-valence exchange and correlation amounts roughly to neglecting the "core" terms in (9.5).

This will be a good approximation if

$$\frac{4\pi e^2}{m} \sum_{\mu(\text{core})} \frac{f_{0\mu}(k)}{\omega^2 - \omega_{0\mu}^2} \ll 1. \qquad (9.6)$$

In NP II, we shall see that (9.6) means that the core has a very small polarizability; this result confirms our qualitative discussion of Sec. 2.

The present treatment of the effect of the core on plasmons yields the right results. But it is unsatisfactory in several ways. First, it seems unphysical to introduce such a large plasma frequency as Ω_p, and to cancel most of it in a latter stage. Second, we obtain a screened interaction H_{rp} which has a very unsymmetrical form. We expect the long-range core-core Coulomb interaction to be unscreened so that it should appear in the form

$$\sum_{k<k_c} \tfrac{1}{2} M_k^2 (\rho_k)_{\text{core}} (\rho_{-k})_{\text{core}}. \qquad (9.7)$$

On the contrary, we obtain an H_{rp} which is roughly

$$- \sum_{k<k_c} \tfrac{1}{4} M_k^2 \{ \alpha_k \beta_{-k} + \beta_{-k} \alpha_k \},$$

$$(\alpha_k)_{mn} = [(\rho_k)_{\text{core}}]_{mn} / \omega_{mn}, \qquad (9.8)$$

$$(\beta_{-k})_{mn} = [(\rho_{-k})_{\text{core}}]_{mn} \times \omega_{mn}.$$

In fact, the expressions (9.7) and (9.8) are equivalent. (They have the same expectation value, and the differences in the off-diagonal terms do not matter, for tightly bound electrons are not sensitive to long-range correlations.) However, this is again an esthetically unpleasant result.

We need therefore a treatment satisfying the following requirements:

(a) Use of the representation where $\mathcal{3C}_0$ is diagonal (in order to treat correctly the short-range correlations).

(b) Introduction at the first stage of plasmons reasonably close to the actual ones.

This may be obtained with a "symmetrized" treatment, where one introduces directly the true field variables, previously obtained from a study of the equations of motion. This method is mathematically more elegant, but more formal, and leads to a complicated discussion of the approximations. It will be published elsewhere.

10. CONCLUSION

Let us summarize what we have accomplished in this paper. First, and most important, by our series of canonical transformations we have shown that there may exist a set of N' plasmons which, under suitable circumstances, represent well-defined elementary excitations of electrons in solids. We have discussed in considerable detail what are the conditions for the existence of the plasmon as an independent excitation in solids. The first condition (our "existence" criterion)

is that the electrons be sufficiently free to take part in an electrostatic excitation. In other words, we require a highly polarizable gas, which means that a sizeable fraction of oscillator strengths correspond to transitions in which $\omega_{n0} \ll \omega_p$. Our dispersion relation also shows that only the low-energy oscillator strengths are important in the determination of the plasmon energy.

The plasmons are not completely independent excitations, since a plasmon may decay into an individual electronic excitation. In Sec. 8 we have derived an expression for this damping, which confirms the result obtained by Wolff and Kanazawa. However, we should like to emphasize that such a result depends on the assumption that neither the subsidiary condition, nor H_{rp}, alter markedly the individual electronic excited states for which $\omega_{n0} \sim \omega_p$. This question is still not clear; we return to it in NP III.

We have developed a satisfactory description of the plasmons only if our series of canonical transformations is convergent. Our "convergence" criterion is that the majority of the oscillator strengths correspond to ω_{n0} much larger or much smaller than ω_p. This criterion also guarantees a narrow plasma line, which, however, may be shifted very much from the free electron one. Where both the "existence" and "convergence" criteria are satisfied, then, we expect to find well-defined plasmons, with a lifetime which is long compared to $1/\omega$.

We have stated that we expect to find both of these criteria satisfied in a wide variety of solids. It is of course desirable to have an experimental check on this point. We postpone the comparison of our theoretical expectations with experiment to NP III, because it is desirable to introduce the concept of the dielectric constant before carrying out the comparison.

Where our existence criterion, but not our convergence criterion, may be satisfied (as in the transition and noble metals), we still expect to find plasmons. The plasmon "lines" will be broad and the energy substantially shifted from $\hbar\omega_p$. Under these circumstances, we are simply not able to describe their properties accurately.

What of the electrons? The electrons are found to interact through a screened Coulomb interaction. They are further subject to a set of subsidiary conditions which reduce the number of independent electronic degrees of freedom from $3N$ to $(3N-N')$. The subsidiary conditions may also alter somewhat the excitation spectrum of the individual electrons. They do not, however, influence the calculation of the ground state energy of the solid, since for this state they are automatically satisfied.

Thus, for a solid for which plasmons are a well-defined elementary excitation, we may expect to be able to calculate the cohesive energy, and the position of the energy levels, for instance, by solving the Hamiltonian, $H_0 + H_{sr} + H_{rp}$. We may further anticipate that H_{rp} will not influence the energy levels appreciably; its low-energy matrix elements are strongly screened,

while the high-energy matrix elements, which are not, will not be effective. (Tightly bound electrons are not sensitive to a long-range interaction.)

We accordingly propose $H_0 + H_{sr}$ as the basic electron Hamiltonian in "well-defined plasmon" solids. (Li, Na, Be, Mg, Al, Si are a few examples of such solids.) One still has a complicated many-body problem. However, the simplification arising from the appearance of H_{sr} rather than the full Coulomb interaction in the Hamiltonian may be considerable. For instance, in doing a Hartree-Fock calculation, the exchange contribution will be markedly reduced in magnitude. One might accordingly hope that calculations which have displayed extreme sensitivity to the form chosen for the exchange potential[29] will lose that sensitivity when the proper screened exchange potential is introduced. A sensible approximation for the above solids might be to carry through a Hartree-Fock calculation for $H_0 + H_{sr}$, then combine it with the free electron value for the correlation energy to get the cohesive energy and energy levels for the system. (A variational choice of k_c could be made, unless an experimental value[30] is available, in which case the latter would, of course, be preferable.) A calculation of just this kind has been carried out recently by Heine[31] for Al; in this fashion he has achieved considerable success in obtaining the experimentally observed properties of the electrons.

We saw in Sec. 7 that our calculation was best carried out in a representation where $(H_0 + H_{sr})$ was diagonal, since then we automatically include correctly the correction to plasmon behavior arising from H_{sr}. When most of the ω_{mn} are much smaller than ω_p (very weakly bound electrons), we can also use a representation in which H_0 alone is diagonal, since we are then able to justify the neglect of the terms arising from H_{sr}. Our ability to do this is essentially limited by the extent to which H_{sr} represents a relatively small perturbation on the electron motion. H_{sr} may give a large shift of the total energy, and our treatment using the eigenstates of H_0 alone may still succeed. As far as its success is concerned, we require only that H_{sr} change slightly the matrix elements $(\rho_k)_{n0}$ and the energy *differences* ω_{n0}. When these quantities are strongly affected by H_{sr}, we must switch to the many-body representation, and try to approximate the influence of H_{sr} on the $(\rho_k)_{n0}$ and the ω_{n0} in some other way.

As long as k_c is smaller than any vector $\mathbf{K}$ of the reciprocal lattice, the screening does not affect the Fourier components of wave vector $\mathbf{K}$ of the Coulomb interaction. We know that such components give rise to the usual Hartree potential. Therefore, the screening only affects the exchange potential. If, however, k_c lies outside the first Brillouin zone, a new situation arises.

---

[29] See, for instance, J. Callaway, Phys. Rev. **97**, 933 (1955).
[30] H. Watanabe, J. Phys. Soc. Japan **11**, 112 (1956); for a discussion of yet another method of determining k_c, see R. A. Ferrell, Phys. Rev. **107**, 450 (1957).
[31] V. Heine, Proc. Roy. Soc. (London) (to be published).

We have seen that in such a case there is a direct coupling between plasmons of different momenta. It moreover leads to a physically absurd Hartree potential: while the component of wavelength $\mathbf{K}$ of the electron interaction is screened, the component V_K of the periodic potential of the nuclei is not. Since it is physically obvious that the electrons shield the nuclei as well as the electrons, the result is incorrect. For the variational choice of k_c discussed in Sec. 4, it is marginal whether such a situation will arise in practice, even in multivalent solids for which it is most likely. It is not difficult to take into account the direct coupling between plasmons of different momenta, and such a calculation will be outlined in NP IV. Finally, we remark that if V_K is small, no difficulties at all will arise, since one then has essentially the free electron case.

In conclusion, we remark that in this paper we have determined the conditions under which plasmons can be regarded as well-defined physical entities. No problem arises with the subsidiary conditions, as far as the plasmons are concerned, because the plasmons do not appear in them. The ground-state energy of the electron system is likewise unaffected by the subsidiary conditions. At this stage we have not clarified the nature of the remaining electron excitations; we consider the problem in the next two papers of the series.

APPENDIX

We want to find an operator X_k which describes a physical variable of wavelength k, oscillating at a well-defined frequency ω. We must therefore have

$$\dot{X}_k = (i/\hbar)[H, X_k] = i\omega X_k. \qquad (A1)$$

Any operator describing a longitudinal motion of wavelength k must be deduced from ρ_k, which has the same translational properties. Let us therefore try

$$(X_k)_{mn} = g_{mn}(\rho_k)_{mn}, \qquad (A2)$$

where g_{mn} is an unknown quantity, depending on both states Φ_n and Φ_m.

If we put this value of X_k in (A1), we get a set of equations for g_{mn}, which are too complicated to be solved rigorously:

$$\hbar(\omega - \omega_{mn})g_{mn}(\rho_k)_{mn} = \sum_{k'} \frac{2\pi e^2}{k'^2}\{[\rho_{k'}, X_k]\rho_{-k'} + \rho_{-k'}[\rho_{k'}, X_k]\}_{mn}. \qquad (A3)$$

If, however, we make the random phase approximation, and replace the commutator $[\rho_{k'}, X_k]$, by its expection value, Eqs. (A3) are considerably simplified, and become

$$\hbar(\omega - \omega_{mn})g_{mn}(\rho_k)_{mn}$$
$$= \frac{4\pi e^2}{k^2}(\rho_k)_{mn}\sum_p\{|(\rho_k)_{0p}|^2(g_{p0} - g_{0p})\}. \qquad (A4)$$

(A4) is trivially solved and gives the following result:

$$1 = \frac{4\pi e^2}{k^2\hbar}\sum_n \frac{2\omega_{n0}|(\rho_k)_{0n}|^2}{\omega^2 - \omega_{n0}^2} = \frac{4\pi e^2}{m}\sum_n \frac{f_{0n}(k)}{\omega^2 - \omega_{n0}^2}, \qquad (A5)$$

$$g_{mn} = c/(\omega - \omega_{mn}), \qquad (A6)$$

where c is an arbitrary constant.

(A5) is just the dispersion equation for the plasmon frequency, proposed by Mott, which we obtain in Sec. 5 by another method. (A6) gives g_{mn}, and therefore determines the oscillating variable X_k. If ω_{mn} is much smaller than ω, we see the X_k is roughly proportional to ρ_k, which confirms the result of Sec. 3. If, however, a large part of the ω_{mn} are larger than ω, X_k differs appreciably from ρ_k. This determination of the true X_k constitutes in fact the starting point of the symmetrized treatment mentioned at the end of Sec. 9.

Let us remark that this determination of the dispersion equation is valid for any longitudinal mode. If there are several modes of wavelength k (analogous for instance to the optical and acoustical longitudinal phonons), their frequencies must correspond to different roots of the *unique* dispersion equation (A5). In NP III, we shall verify this property for the two kinds of plasmons occurring in an impurity semiconductor.