[1]M. Toda and M. Glicksman, Phys. Rev. 140, A1317 (1965).
[2]D. Ferry and H. Heinrich, Phys. Rev. 169, 670 (1968).
[3]K. Hübner and L. Blossfeld, Phys. Rev. Letters 19, 1282 (1967).
[4]H. Schmidt and D. J. Nelson, Phys Rev. 184, 760 (1969).
[5]W. P. Dumke, Phys. Rev. 167, 783 (1968).
[6]G. A. Baraff, Phys. Rev. 128, 2507 (1962).
[7]W. B. Kunkel, in Proceedings of the Eighth International Conference on Phenomena in Ionized Gases (Springer, Vienna, 1967), p. 177.

Statistical Mechanics of Charged Traps in an Amorphous Semiconductor*†

G. Srinivasan

*Department of Physics and the James Franck Institute,
The University of Chicago, Chicago, Illinois* 60637

(Received 12 February 1971)

The effect of a Coulomb interaction between charged traps in an amorphous semiconductor is investigated within the premises of the Mott-Cohen-Fritzsche-Ovshinsky model. The grand partition function is expressed as a functional integral over a set of Gaussian random fields. The free energy is expressed as a sum of the mean-field result plus fluctuations about the mean field. It is shown that for the system under consideration, the mean field is just the Hartree self-consistent field and that at $T = 0$ °K it represents the exact ground state. It is shown that the fluctuations about the mean field represent correlations in the system. Approximate expressions for the mean occupation number and the renormalized energies of the charges are obtained as well as the renormalized single-particle density of states. The excitation spectrum of single quasiparticles, within any given band, is shown to have a quasi-gap. It is shown that the effect of a Coulomb interaction between the charged traps is to reduce the density of states at the Fermi energy by a factor of 2 below its value in the absence of interactions.

I. INTRODUCTION

The discovery of novel phenomena in amorphous semiconductors has prompted considerable theoretical interest in these materials. The theoretical attempts are aimed towards the understanding of the electronic structure and transport in disordered systems, in particular in covalent amorphous semiconducting alloys. There exists, as yet, no rigorous theory of the electronic structure in disordered systems. There has emerged, however, a basic band model (which was synthesized by Mott[1] out of earlier work) which illustrates the universal features of the electronic structure of disordered materials. This basic band model has been further elaborated and clarified, in the context of amorphous semiconductors, by Cohen, Fritzsche, and Ovshinsky[2] (we will refer to it as the Mott-CFO model). The basic features of their model are displayed in Fig. 1 via a sketch of the density of states. For our purpose it is necessary to briefly review some of the essential assumptions underlying the model. They postulate that in an amorphous semiconductor there exist bands of extended states, that these bands have tails of localized states, and that in sufficiently disordered materials (such as alloys or bad films) these tails overlap in the forbidden gap. They further assume that every localized state has a well-defined parentage, i.e., it is always possible to assign a valence or conduction character to a localized state in the gap. The overlapping of the tails and the fact that there is a finite density of localized states at the Fermi energy has many interesting consequences. Since valence states are electrically neutral when occupied and conduction states electrically neutral when empty, there results a random distribution of localized charged traps throughout the material, positive trapped holes (corresponding to the empty valence-band tail above the Fermi energy E_F) and negative trapped electrons (corresponding to the occupied conduction-band tail below E_F). The overall electrical neutrality of the material is guaranteed through a proper choice of E_F. One believes that there may be as many as 10^{19} of these localized states (per cm^3 eV) at E_F. One anticipates that such a distribution of localized charges will have a significant effect on carrier kinetics. One would like to know, for example, what the effect would be of these localized charges on the carrier concentra-

tion, on the mobility edge and the extent of renor-
malization of the tail in the density of states, etc.
In this paper we investigate the effect of interac-
tion between the charged localized state in the re-
gion of tail overlap. This, we feel, is a necessary
first step before one can ask the general question
as to what is the effect of these charged localized
states on the carrier kinetics.

In Sec. II we define the model within which we
shall be working and clarify it. The question of
configurational averaging of the free energy of
these localized charges is discussed. This point
assumes particular importance since we are deal-
ing with a random system.

In Sec. III we introduce the method of functional
integrals or the method of Gaussian random fields
and formulate the problem of evaluating the grand
partition function in these terms.

In Sec. IV we evaluate the free energy in the
mean-field approximation. We linearize the formal
expression for the free energy with respect to the
mean field and then evaluate its configurational
average. In performing the configurational aver-
age we use Mayer's graph theory and restrict our-
selves to the so-called "ring diagrams."

In Sec. V we obtain an approximate expression
for the renormalized energy of the quasiparticles
and their mean occupation numbers. We then
demonstrate that the mean field is just the gen-
eralized Hartree self-consistent field, and that,
for the particular problem under consideration,
the mean-field approximation represents the exact
ground state at $T = 0\,°$K. As a final result of the
section we show that the single-quasiparticle ex-
citation energy exhibits an energy gap.

In Sec. VI we discuss the fluctuations about the
mean field. We obtain explicit results for the
lowest-order fluctuations. We then prove that the
higher-order fluctuations yield a contribution to
the free energy which is higher order in (kT) and
(e) than the leading term {which comes from fluc-
tuations of $O[(\Phi - \Phi^0)^2]$}.

In Sec. VII we discuss some of the consequences
of Secs. IV and V to the electronic structure and
transport in amorphous semiconductors.

II. MODEL

We shall now turn to the question of specifying
the model with which we will be working and clari-
fying the notation. Lately there have been consid-
erable theoretical attempts towards understanding
the nature of localized states in a disordered sol-
id. While considerable insight has been gained in
this regard,[3] no definitive picture as to the precise
nature of localized states has emerged. For the
purpose of the present calculation we will assume
that the states under consideration are localized
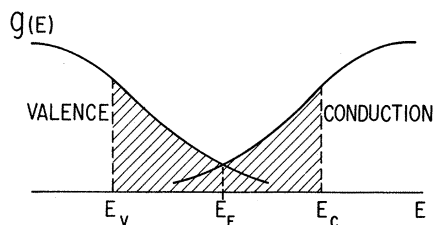over distances of the order of several atomic sep-

FIG. 1. Plot of the density of states in the Mott–CFO
model for a covalent amorphous semiconducting alloy.
E_v and E_c are the mobility edges. The shaded area cor-
responds to localized states.

arations. To be more precise, we will investigate
a situation in which a set of localized states (of
both valence and conduction character) are dis-
persed at random throughout the amorphous semi-
conductor. We shall further assume that these
states are singly occupied. This is one of the as-
sumptions of the Mott-CFO model, but is not nec-
essarily correct. We shall adopt the following
notation for the specification of the system. We
are given a (specific) set of lattice sites, corre-
sponding to the centers of localized states in the
energy gap. These include two species of sites:
valence sites and conduction sites. We shall label
the set of sites $\{A\}$, where

$$\{A\} \to \{i, \sigma\} \ , \quad i \to \vec{R}_i, b_i, E_i^b \ .$$

$\vec{R}_i$ is the coordinate of the center of the localized
state, b_i the species or band index (v for valence
and c for conduction state), E_i^b the energy of a
noninteracting quasiparticle at the site $\vec{R}_i$, and σ
the spin index. Let N^L be the number of localized
states in the gap, consisting of N_v^L valence sites
and N_c^L conduction sites, i.e.,

$$N^L = \int_{E_v}^{E_c} g_v(E)\,dE + \int_{E_v}^{E_c} g_c(E)\,dE \ .$$

Here $g_b(E)$ is the single-quasiparticle density of
states in energy and E_v and E_c are the mobility
edges in the Mott-CFO model (Fig. 1). As was
remarked earlier one of the assumptions of the
Mott-CFO model is that a localized state always
has a well-defined valence or conduction-band
character. As such, an unoccupied valence site
carries a positive charge, and an occupied conduc-
tion site has a negative charge. We shall refer to
these as quasiholes in the valence band and quasi-
electrons in the conduction band. We shall assume
for the potential energy of interaction between
these charged localized states, $\mathcal{V}_{AB}$, the following
form:

$$\mathcal{V}_{AB} = e q_A V_{AB} e q_B / \kappa_0,$$

with

$$V_{AB} = 1/R_{AB} , \quad R_{AB} > R_0$$
$$\qquad = 1/R_0, \quad\ \ R_{AB} < R_0 \ .$$

Here, κ_0 is the phenomenological dielectric constant of the amorphous semiconductor, R_0 is the mean separation between charges, and q_A is the charge characteristic of the site A ($q_A^v = +1$, $q_A^c = -1$). We can therefore write for the total energy of such a random assemblage of interacting charged localized states

$$E(\{n\}) = \sum_A \epsilon_A n_A + \frac{1}{2} \sum_{ij}{}' n_A q_A \hat{e} V_{AB} \hat{e} q_B n_B , \quad (2.1)$$

where $n_A = 0, 1$ is the occupation number of quasi-holes in the valence band and quasielectrons in the conduction band; $\epsilon_A = q_A (\mu - E_A)$ is the energy of the quasiparticle measured from the chemical potential μ and we have defined an effective charge $\hat{e} = e / \kappa_0^{1/2}$. In Eq. (2.1), and in what follows, the summation over the (composite) index A runs over the entire set of localized states. Finally, in order to be able to use some mathematical identities we shall set the diagonal terms V_{AA} in the interaction energy equal to zero.

We now turn to the question of evaluating the free energy of these charges. Since we have chosen the occupation-number formalism we must work in the grand canonical ensemble. In this representation the central quantity is the grand partition function Z given by

$$Z = \mathrm{Tr} e^{-\beta(H - \mu N)} ,$$

where H is the Hamiltonian of the system, μ the chemical potential, N the total number of particles, and $\beta = 1/kT$. Since we are dealing with a "frozen system" the trace in the above expression is to be taken only over the "internal degrees of freedom." In our present case this refers to summing over all configurations $n_A = 0, 1$ except those which correspond to double occupancy of a site, i.e., we forbid all configurations of the type $n_{\uparrow} = 1$ and $n_{\downarrow} = 1$. Since we have already expressed the quasiparticle energies with respect to the chemical potential μ, we have for the partition function

$$Z(\{R\}, \{\epsilon\})$$

$$= \sum_{\{n\}}{}' e^{-\beta E(\{n\})}$$

$$= \sum_{\{n\}} \exp\left(-\beta \sum_A \epsilon_A n_A - \frac{\beta}{2} \sum_{AB} n_A q_A \hat{e} V_{AB} \hat{e} q_B n_B\right) .$$

$$(2.2)$$

The prime over the summation refers to the neglect of double occupancy of sites. The free energy is given by $-kT \ln Z$. Clearly, the above expression yields the partition function for a "specific" configuration ($\{R\}, \{\epsilon\}$) of the charges. Before obtaining the thermodynamic free energy one should, at some stage, perform an average over all configurations $\{R\}$ and $\{\epsilon\}$. Since we are dealing

with a random sample there is no correlation between the occupation numbers and the specific spatial configuration or the configuration of single-particle energies. Consequently we must calculate the free energy for a specific configuration and then average this over all configurations, i.e., we must calculate $\langle -kT \ln Z(\{R\}, \{\epsilon\}) \rangle$ as opposed to $-kT \ln \langle Z(\{R\}, \{\epsilon\}) \rangle_c$. To follow the latter prescription would imply the neglect of fluctuations in the system and could lead to qualitatively different results. To summarize then, since we are dealing with a random frozen sample, i.e., a specific nonequilibrium distribution of the position and energies of the charges, one must calculate the free energy for each configuration and *then* perform an average over all configurations. As will be seen later this is precisely the point that leads to difficulties in performing the configurational average.

III. METHOD OF GAUSSIAN RANDOM FIELDS

The problem of evaluating the free energy of a system of particles with long-range interactions is encountered in many places, e.g., in electrolyte theory, plasma physics, etc. Systems with long-range forces cannot be handled in the same straightforward manner as systems with short-range forces. It is to be recalled that a simple calculation of the virial coefficients leads to a divergent series[4]; this is the classical analog of the divergence of the thermodynamic perturbation theory in the calculation of the quantum-mechanical partition function.[5] In certain classical problems ingenious methods have been devised to manipulate the divergent series to yield a finite answer,[6] and these have been generalized in a natural fashion to quantum statistical mechanics. These methods, however, are increasingly difficult to work with if one wants to do better than the first approximation. In this paper we shall employ the method of functional integrals to evaluate the partition function. While it is true that the use of this technique (which is closely related to Feynman's path-integral technique) has not yielded any new results in problems where it has been employed, it has the advantage that the mathematics of it least obscures the physics. For the purpose of the present calculation we shall introduce the formalism through a well-known identity in the theory of quadratic forms.

For a set of n (real or complex) variables x_k and for any real symmetric positive-definite matrix $\underline{A}$

$$\exp\left(-\frac{1}{2} \sum_{kl} x_k A_{kl} x_l\right) = (2\pi)^{-n/2} |\underline{A}|^{-1/2} \int_{-\infty}^{+\infty} \cdots \int \prod_i d\,\phi_i$$

$$\times \exp\left(-\frac{1}{2} \sum_{kl} \phi_k A_{kl}^{-1} \phi_l + i \sum_k \phi_k x_k\right), \quad (3.1)$$

where $|\underline{A}|$ is the determinant of $\underline{A}$ and $\underline{A}^{-1}$ is the matrix inverse of $\underline{A}$. This identity may be proved by transforming to variables in which $\underline{A}$ is diagonal.[7] Let

$$P_n(\{\phi\}) \equiv (2\pi)^{-n/2} |\underline{A}|^{-1/2} \exp\left(-\frac{1}{2}\sum_{kl} \phi_k A_{kl}^{-1} \phi_l\right) .$$

(3.2)

It is easy to see that $P_n(\{\phi\}) \geq 0$ and $\int P_n(\{\phi\})\Pi_i d\phi_i = 1$. $P_n(\{\phi\})$ is therefore a multivariate probability distribution. In fact, it can be readily shown that

$$\langle\phi_k\rangle_{P(\phi)} = 0 , \quad \langle\phi_k\phi_l\rangle_{P(\phi)} = A_{kl} ,$$

(3.3)

where $\langle \rangle_{P(\phi)}$ denotes an average with respect to the distribution $P(\{\phi\})$. Thus $\{\phi\}$ constitutes a set of Gaussian random variables with mean zero and dispersion given by the matrix $\underline{A}$. Going back to (3.1) we may write

$$\exp\left(-\frac{1}{2}\sum_{kl} x_k A_{kl} x_l\right) = \left\langle\exp\left(i\sum_k \phi_k x_k\right)\right\rangle_{P(\{\phi\})} .$$

(3.4)

Partition Function

We now turn to the question of evaluating the partition function. What we shall do is to cast the long-range part of Z [Eq. (2.2)] as an integral over a space of Gaussian random variables through the use of the identity (3.1). That is, we shall write

$$\exp\left(-\frac{\beta}{2}\sum_{AB} n_A q_A \hat{e} V_{AB} \hat{e} q_B n_B\right)$$

$$= L\int \delta\phi \exp\left(-\frac{\beta}{2}\sum_{AB} \phi_A V_{AB}^{-1} \phi_B - \beta\sum_A \hat{e}' q_A \phi_A n_A\right) ,$$

(3.5)

where

$$L^{-1} \equiv \int \delta\phi \exp\left(-\frac{\beta}{2}\sum_{AB} \phi_A V_{AB}^{-1} \phi_B\right)$$

(3.6)

and $\hat{e}' = \sqrt{(-1)}\hat{e}$. Substituting (3.5) into (2.2) we get

$$Z = L\int \delta\phi \exp\left(-\frac{\beta}{2}\sum_{AB} \phi_A V_{AB}^{-1} \phi_B\right) \sum_{\{n\}}' \exp\left(-\beta\sum_A (\epsilon_A + \hat{e}' q_A \phi_A) n_A\right) .$$

(3.7)

In (3.7) the prime over $\sum_{\{n\}}$ means that we sum over all configurations $n_A = 0, 1$ except those that correspond to double occupancy of sites, i.e., for any given site index i the allowed configura-

tions are $n_{i\uparrow} = 0$, $n_{i\downarrow} = 0$; $n_{i\uparrow} = 0$, $n_{i\downarrow} = 1$; $n_{i\uparrow} = 1$, $n_{i\downarrow} = 0$; but we exclude the configuration $n_{i\uparrow} = 1$, $n_{i\downarrow} = 1$. This sum over $\{n\}$ can be readily performed, and this yields

$$Z(\{R\}, \{\epsilon\}) = L\int \delta\phi \exp\left(-\frac{\beta}{2}\sum_{ij} \phi_i V_{ij}^{-1} \phi_j\right) \exp\left(\sum_i \ln(1 + 2e^{-\beta(\epsilon_i + \hat{e}' q_i \phi_i)})\right)$$

(3.8)

$$= \left\langle\exp\left(\sum_i \ln(1 + 2e^{-\beta(\epsilon_i + \hat{e}' q_i \phi_i)})\right)\right\rangle_{P(\{\phi\})} .$$

(3.9)

We will soon see that the Gaussian random variable ϕ_i will have the physical interpretation of a Coulomb field at the site i. It may be recalled that the expression inside the angular brackets in Eq. (3.9) is just the partition function for a system of "noninteracting" fermions, each in an external field. We have therefore succeeded in reducing the problem of interacting particles to one of noninteracting particles but with each particle interacting with an external field; we then average this result over all possible configurations of the external fields. It should be noted that the external fields, which are Gaussian random fields, are determined entirely by the interaction matrix $\underline{V}$. It must be emphasized, however, that the simplification achieved through such a decoupling procedure is offset to some extent by the fact that

we now have to evaluate a rather complex functional integral, viz., Eq. (3.8).

IV. MEAN-FIELD APPROXIMATION

Since Eq. (3.9) expresses the partition function as an average over all ϕ of $Z(\{\phi\})$ there must be a particular set $\{\phi^0\}$ which will dominate this average. We shall now proceed to determine the set of dominant fields—which we call the "mean field." For example, if we were dealing with the Ising model[8] the mean field will turn out to be just the Weiss molecular field. Let us first define a functional $B[\{\phi\}]$ through the relation

$$Z = L \int \delta\phi\, e^{-B[\{\phi\}]} .$$

(4.1)

Comparison with Eq. (3.8) shows that

$$B[\{\phi\}] = \frac{\beta}{2} \sum_{ij} \phi_i V_{ij}^{-1} \phi_j - \sum_i \ln(1 + 2e^{-\beta(\epsilon_i + \hat{e}'q_i\phi_i)}).$$
(4.2)

The dominant field ϕ_m^0 is obtained by solving the equation

$$(\delta B / \delta \phi_m)\big|_{\phi_m^0} = 0 .$$
(4.3)

We can thus "factor" the partition function into a product of the mean-field result times the fluctuations about the mean field, viz.,

$$Z = Z_{mf} \times Z_{fluct} ,$$
(4.4)

where $Z_{mf} = \exp(-B[\{\phi^0\}])$. In Sec. V we shall evaluate the fluctuations of $O[(\phi - \phi^0)^2]$. Let us now return to the question of evaluating the mean fields. From Eq. (4.2) it follows that

$$\frac{\delta B}{\delta \phi_m} = 0 = \beta \sum_j V_{mj}^{-1} \phi_j + \beta \hat{e}' q_m n_0(\epsilon_m + \hat{e}'q_m \phi_m) ,$$
(4.5)

where

$$n_0(\epsilon) = 1 / (1 + \tfrac{1}{2} e^{\beta \epsilon})$$
(4.6)

is the mean occupation number of charges in the noninteracting system; this differs from the usual Fermi function because we have forbidden double occupancy. The above equation yields a solution for the mean field at the site i

$$\phi_i^0 = - \hat{e}' \sum_m V_{im} q_m n_0(\epsilon_m + \hat{e}'q_m \phi_m^0) .$$
(4.7)

This says that the mean field at any given site is the sum of bare Coulomb potentials due to dressed particles at all other sites. The expression (4.7) is exact and should, in principle, be solved self-consistently. Indeed, we shall show, towards the end of Sec. V, that the mean field is just the generalized self-consistent Hartree field.

We can now write down a formal expression for the free energy in the mean-field approximation, F_{mf}. We have

$$F_{mf} = - kT \ln Z_{mf} = kTB[\{\phi^0\}] .$$
(4.8)

Substituting (4.7) into (4.2) we obtain, after some algebra,

$$F_{mf} = - kT \sum_i \ln(1 + 2e^{-\beta(\epsilon_i + \hat{e}'q_i\phi_i^0)})$$
$$- \tfrac{1}{2} \sum_i \hat{e}' q_i \phi_i^0 n_0(\epsilon_i + \hat{e}'q_i \phi_i^0) .$$
(4.9)

Linearized theory. We shall now proceed to obtain a more explicit expression for the thermodynamic free energy using a linearized version of the above theory. Expanding the two terms in (4.9) in powers of ϕ^0 and keeping only the linear term we obtain for the free energy

$$F_{mf}(\{i\}) = - kT \sum_i \ln(1 + 2e^{-\beta \epsilon_i})$$
$$+ \tfrac{1}{2} \sum_i \hat{e}' q_i \phi_i^0 n_0(\epsilon_i) + O[(\phi^0)^3]$$
(4.10)

{the terms of $O[(\phi^0)^2]$ in the expansion of the two

terms in (4.9) cancel}. We shall soon see that the second term in (4.10), upon averaging, will be of $O[\hat{e}^2 \kappa]$, κ being the inverse screening length. The terms higher order in ϕ^0 involve the higher derivatives of the Fermi function and, upon averaging, will be proportional to the (higher) derivatives of the density of states at the chemical potential. If we define the typical width of the tail of density of states, Δ, through

$$\Delta^{-1} \approx \frac{d \ln g(E)}{dE}\bigg|_\mu ,$$

then the correction due to the higher-order terms will be small if $\Delta > \hat{e}^2 \kappa$. Therefore the linearization of the free energy in the mean field is valid in the limit of (a) low concentration of charges and (b) wide tails for the density of localized states in the gap. The first term is just the free energy of a system of noninteracting fermions F_0 and we shall call the second term, in anticipation, the Hartree term F_H. Before proceeding further we need an expression for the mean field in this linear theory. Let us go back to Eq. (4.5),

$$\sum_j V_{mj}^{-1} \phi_j^0 + \hat{e}'q_m n_0(\epsilon_m + \hat{e}'q_m \phi_m^0) = 0 .$$

Linearizing this with respect to ϕ_m^0 we get

$$\sum_j \left(V_{mj}^{-1} - \hat{e}^2 \frac{\partial n_0(\epsilon_m)}{\partial \epsilon_m} \delta_{mj}\right) \phi_j^0 = - \hat{e}'q_m n_0(\epsilon_m) .$$

Let us define a diagonal matrix $\underline{\Lambda}$ by

$$\underline{\Lambda} \equiv \text{diag}\left\{ - \hat{e}^2 \frac{\partial n_0(\epsilon)}{\partial \epsilon} \right\} .$$
(4.11)

Inverting the above equation we obtain for ϕ_i^0

$$\phi_i^0 = - \hat{e}' \sum_m (\underline{V}^{-1} + \underline{\Lambda})_{im}^{-1} q_m n_0(\epsilon_m) .$$
(4.12)

We now define the "screened Coulomb potential" $\underline{V}^{sc}$ through

$$\underline{V}^{sc} \equiv \underline{V} \cdot \underline{\epsilon}^{-1}$$
$$\equiv (\underline{V}^{-1} + \underline{\Lambda})^{-1}$$
$$= \underline{V} \cdot (\underline{I} + \underline{\Lambda}\,\underline{V})^{-1} .$$
(4.13)

Here $\underline{I}$ is the unit matrix and $\underline{\epsilon} = (\underline{1} + \underline{\Lambda}\,\underline{V})$ is the "dielectric matrix." It is clear, from the definition of $\underline{\Lambda}(\epsilon)$ [Eq. (4.11)], that $\underline{\epsilon}$ is just the Hartree dielectric function. Substituting (4.12) into (4.10) we may write the second term as

$$F_H = \frac{1}{2} \sum_{ij} n_0(\epsilon_i) q_i \hat{e} \, V_{ij}^{sc} \hat{e} q_j n_0(\epsilon_j).$$
(4.14)

This is just the electrostatic energy of bare particles interacting via a shielded potential.

To summarize what we have done up to this point, we have derived formal expressions for the free energy and the mean field and in addition obtained explicit expressions for these within the

framework of a linear theory, viz., linear in the mean fields. All this has been done for a specific configuration ($\{R\}, \{E\}$) of localized states. As we discussed in Sec. II, we must now average the free energy over all configurations.

We define the configurational average of a quantity $\mathcal{O}$ as

$$\langle \mathcal{O} \rangle \equiv \int \mathcal{O} \, \mathcal{P}(\{R\}, \{E\}) d\{R\} d\{E\} , \qquad (4.15)$$

where $\mathcal{P}(\{R\}, \{E\})$ is the probability of finding the configuration ($\{R\}, \{E\}$) from among all configurations. As a first approximation we shall assume that there is no correlation between the positions of the quasiparticles and their energies. We shall further assume that the charges are distributed randomly throughout the sample. This is a fairly good assumption for two reasons: To begin with we are primarily interested in glassy (or disordered) solids; and second, in the kinds of samples we are envisaging, viz., a dilute concentration of charges, the separation between the charges will be large compared to both the mean separation between the charges and the spatial extent of the localized states. Under these circumstances there will be negligible correlations between the positions and energies of the charges. This means that we can write

$$\mathcal{P}(\{R\}, \{E\}) \cong \prod_i^{N^L} p_i(\vec{R}_i) p_i(E_i)$$

$$= \frac{1}{(\Omega)^{N^L}} \prod_i p_i(E_i) ,$$

where Ω is the volume of the sample.

The probability $p(E_i)$ that a quasiparticle will have an energy E_i is just the frequency of occurrence of energy levels in a macroscopic sample, viz., the density of states in energy:

$$p_i(E_i) = p(E_i) = g(E_i) / \int_{E_v}^{E_c} g(E_i) dE_i . \qquad (4.16)$$

We shall now evaluate the configurational average of the mean-field energy [(Eq. (4.10)]. In what follows, we shall restrict ourselves to the case of symmetrical bands. This means that $N_v^L = N_c^L = \frac{1}{2} N^L$ and further $g_v(\mu) = g_c(\mu) \equiv \frac{1}{2} g^{\mathrm{tot}}(\mu)$. Here we have defined $g^{\mathrm{tot}}(\mu)$ as the total density of states at the chemical potential.

The averaging of the first term in Eq. (4.10) is straightforward and yields

$$\langle F_0(\{R\}, \{\epsilon\}) \rangle = -2kT \int_{E_v}^{E_c} \ln(1 + 2 e^{-\beta \epsilon_i}) g(E_i) dE_i . \qquad (4.17)$$

The averaging of the second term, Eq. (4.14), is more complicated. We have

$$\langle F_H \rangle = \frac{1}{2} \sum_{ij} \langle n_0(\epsilon_i) q_i \hat{e} V_{ij}^{\mathrm{sc}} \hat{e} q_j n_0(\epsilon_j) \rangle$$

$$= \frac{1}{2} \sum_{ij} \langle n_0(\epsilon_i) q_i \hat{e} [\, \underline{V}(\underline{I} + \underline{\Lambda} \, \underline{V})^{-1}]_{ij} \hat{e} q_j n_0(\epsilon_j) \rangle . \qquad (4.18)$$

The difficulty arises on account of the fact that the "dielectric matrix," $\underline{\epsilon} = [\underline{I} + \underline{\Lambda}(\epsilon) \underline{V}(R, R')]$, is explicitly configuration dependent. The procedure we adopt is the following: We shall formally expand the inverse dielectric function in powers of $\underline{\Lambda} \, \underline{V}$ and average term by term, i.e., we shall write

$$\langle F_H \rangle = \frac{\hat{e}^2}{2} \sum_{ij} \langle n_0(\epsilon_i) q_i [\, \underline{V} - \underline{V} \underline{\Lambda} \, \underline{V} + \underline{V} \underline{\Lambda} \, \underline{V} \underline{\Lambda} \, \underline{V} \cdots]_{ij} q_j n_0(\epsilon_j) \rangle . \qquad (4.19)$$

Upon averaging we would discover that *each* term in the expansion diverges. This is, of course, the well-known divergence associated with the long-range nature of the Coulomb force. It is, therefore, hopeless to try to obtain a meaningful result from a finite number of terms. It is, however, natural to make a *partial summation of the most divergent contribution to all orders in* $\underline{V}$ and see if this yields a finite answer. It may be recalled that this was precisely the procedure that Mayer developed in the context of electrolyte theory.[6,9] To be more specific we shall, following Mayer, represent the various terms of the expansion (4.19) by *graphs*. One represents each charge by a dot and each factor V_{ij} by a line between dots i and j. With each dot (or vertex) will be associated either the function $n_0(\epsilon)$ or $\Lambda(\epsilon)$. We then have a one-to-one correspondence between a term in (4.19) and its graph. Consider, for example,

the fifth term in (4.19), viz.,

$$\frac{1}{2} \hat{e}^2 \sum \langle n_0(\epsilon_i) q_i V_{ij} \Lambda_j V_{jk} \Lambda_k V_{kl} \Lambda_l V_{lm} \Lambda_m V_{mn} q_n n_0(\epsilon_n) \rangle .$$

In Fig. 2 we have drawn the graphs that correspond to the various terms in the above expression. We shall define the *ring* diagrams (or simple cycles) as the least connected irreducible graphs in any order. The first few ring diagrams are represented in Fig. 3. The remainder of the calculation proceeds in complete analogy with the corresponding steps in electrolyte theory and we shall therefore omit the details of the calculation. The line of argument is as follows: One argues that in any given order "for weak charges and small particle density, the most important contribution comes from the ring diagrams." In fact, one can show that the fact that each one of the ring diagrams diverges in a system with Coulomb forces is the basis for the possibility that in a partial summa-
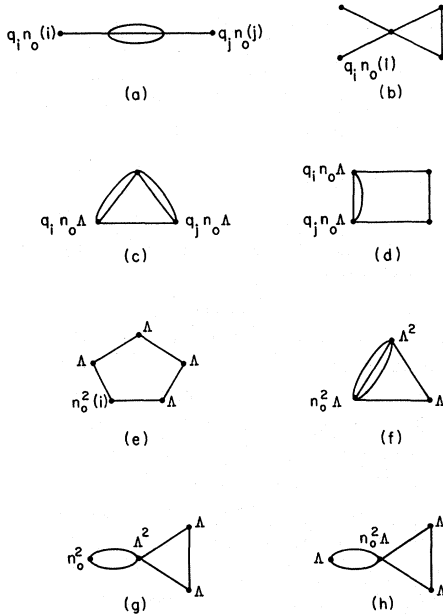
FIG. 2. Some of the graphs that correspond to the term $\frac{1}{2}(e^2) \sum_{j} n_0(i) q_i (\underline{V} \underline{\Lambda} \underline{V} \underline{\Lambda} \underline{V} \underline{\Lambda} \underline{V})_{ij} q_j n_0(\epsilon_j)$ in the expansion (4.19) of F_H.

tion of these the singularities will cancel. In order to do this, one arranges the series in a manner first shown by Mayer and sums it. It can be shown that the contribution due to ring diagrams of all orders to the Hartree energy (4.19) may be written as

$$\langle F_H;\ \text{rings}\rangle = \frac{\hat{e}}{2} N^L \langle n_0^2(\epsilon) \rangle \sum_{n \geq 1} (-1)^n \left(\frac{N^L \langle \Lambda \rangle}{\Omega}\right)^n q_n(0,\,\Omega)\ ,$$

(4.20)

where

$$q_n(0,\,\Omega) = \int_\Omega \cdots \int v_{i1} v_{12} \cdots v_{ni}\, d\tau^n\ .$$

The above series can be summed with the use of Fourier-transform technique and we obtain

$$\langle F_H;\ \text{rings}\rangle \cong \tfrac{1}{2} \hat{e}^2 N^L \langle n_0^2(\epsilon)\rangle (-\kappa)$$

$$= -\hat{e}^2 \kappa \int_{E_v}^{E_c} n_0^2(E) g(E) dE\ ,$$

(4.21)

where κ, defined as the "inverse screening length," is given by

$$\kappa^2 \equiv 4\pi (N^L/\Omega)\langle \Lambda(\epsilon)\rangle$$

(4.22)

and we have substituted for $\langle n_0^2(\epsilon)\rangle$. We recall that $\Lambda(\epsilon) = -\hat{e}^2\, \partial n_0(\epsilon)/\partial \epsilon$. At $T = 0\ {}^\circ\text{K}$, $-\partial n_0(\epsilon)/\partial \epsilon$ is a δ-function at the chemical potential, so that we can write (4.22) as

$$\kappa^2(T=0) = (4\pi \hat{e}^2/\Omega) g^{\text{tot}}(\mu)\ .$$

(4.23)

(We have written the result in terms of the total density of states at the chemical potential.) We

recognize this, of course, as just the inverse Thomas-Fermi screening length. To summarize then, we have evaluated the configurational average of the free energy in the mean-field approximation by summing the contribution due to all ring diagrams, and we find

$$\langle F_{\text{mf}};\ \text{rings}\rangle \cong -2kT \int_{E_v}^{E_c} \ln(1 + 2\,e^{-\beta \epsilon_i}) g(E_i) dE_i$$

$$-\hat{e}^2 \kappa \int_{E_v}^{E_c} n_0^2(E_i) g(E_i) dE_i\ .$$

(4.24)

In Appendix A we shall estimate the contributions of some of the other diagrams in Fig. 2 and show that either they are vanishingly small or of higher order in the electronic charge (we have seen that the ring diagrams give a contribution of $O[e^3]$).

V. MEAN OCCUPATION NUMBER AND ENERGY GAP

We shall now find an expression for the mean occupation number of charges $\bar{n}_l$ at a given site l. We recall that

$$\bar{n}_l \equiv \frac{\partial F[\{R\},\{\epsilon\}]}{\partial \epsilon_l}\ .$$

(5.1)

It therefore follows from (4.9) that[10]

$$\bar{n}_l = n_0(\epsilon_l + \hat{e}' q_l\, \phi_l^0)$$

$$+ \frac{1}{2} \sum_i n_0(\epsilon_i + \hat{e}' q_i\, \phi_i^0) \frac{\partial}{\partial \epsilon_l} (\hat{e}' q_i\, \phi_i^0)$$

$$- \frac{1}{2} \sum_i (\hat{e}' q_i\, \phi_i^0) \frac{\partial}{\partial \epsilon_l} n_0(\epsilon_i + \hat{e}' q_i\, \phi_i^0)\ .$$

(5.2)

We recall that

$$\phi_i^0 = -\hat{e}' \sum_j V_{ij} q_j n_0(\epsilon_j + \hat{e}' q_j\, \phi_j^0)\ .$$

If we substitute this into (5.1) we find that the last two terms cancel, so that

$$\bar{n}_l = n_0(\epsilon_l + \hat{e}' q_l\, \phi_l^0)\ ,$$

(5.3)

where $n_0(x)$ is the Fermi function appropriate to this problem, Eq. (4.6). We can now write for the mean field at site l

$$\phi_l^0 = -\hat{e}' \sum_j V_{lj} q_j \bar{n}_j\ .$$

(5.4)

Let us define V_l^S, the self-consistent potential energy at the site l, through

$$V_l^S = \sum_j \hat{e}^2 q_l q_j V_{lj} \bar{n}_j\ .$$

(5.5)

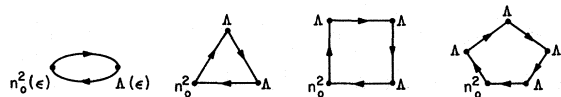This, now, has a clear-cut physical interpretation,

FIG. 3. Ring diagrams that arise in the expansion (4.19). These are the most singular graphs in each order. They correspond to a charge interacting with its own polarization cloud.

G. SRINIVASAN

viz., V_l^S is the total electrostatic energy at site l due to all other sites. In fact, it is just the Hartree self-consistent potential energy. We can rewrite the mean occupation number at site l [Eq. (5.3)] as

$$\bar{n}_l = n_0(\epsilon_l + V_l^S) . \qquad (5.6)$$

This form shows that for this problem the mean field is just the Hartree self-consistent field. As can be readily seen from Eqs. (5.4) and (5.6), the above expression represents the mean occupation number of charges at site l for a specific configuration of the positions and energies of the other charges. Therefore, to get a thermodynamically meaningful quantity we must average this over all spatial configurations and over the energies of all other charges. Let us, therefore, define the average mean occupation number $\langle \bar{n}_l \rangle$ as

$$\langle \bar{n}_l \rangle \equiv \langle n_0(\epsilon_l + \hat{e}' q_l \phi_l^0) \rangle_{\neq \epsilon_l} , \qquad (5.7)$$

where $\langle \rangle_{\neq \epsilon_l}$ signifies the fact that we average over all variables $\{R\}$ and $\{\epsilon\}$ except ϵ_l. In principle one should expand $n_0(\epsilon_l + \hat{e}' q_l \phi_l^0)$ in powers of ϕ_l^0 and average term by term. Since this is very difficult to do we shall, instead, write

$$\langle \bar{n}_l \rangle \cong n_0(\epsilon_l + \langle \hat{e}' q_l \phi_l^0 \rangle) . \qquad (5.8)$$

Such a procedure implies that we have set $\langle (\phi^0)^n \rangle = \langle \phi^0 \rangle^n$, that is, we have assumed that the distribution of ϕ^0 in various configurations is sharply peaked. Actually the fluctuations about this value can be appreciable but we shall ignore them.[11] The question then reduces to evaluating $\langle \hat{e}' q_l \phi_l^0 \rangle_{\neq \epsilon_l}$. We can do this explicitly only for the case of the linearized self-consistent field [Eq. (4.12)]. Since this is almost identical to the averaging of the Hartree term in the free energy we shall simply give the result one obtains upon summing the ring diagrams. We get

$$\langle \hat{e}' q_l \phi_l^0 \rangle_{\neq \epsilon_l} \approx - \hat{e}^2 \kappa n_0(\epsilon_l) . \qquad (5.9)$$

Substituting this into (5.8) we get

$$\langle \bar{n}_l \rangle \approx n_0 [\epsilon_l - \hat{e}^2 \kappa n_0(\epsilon_l)] . \qquad (5.10)$$

At this point one can perhaps make an ansatz about an improved and a more symmetrical form for $\langle \bar{n}_l \rangle$ by setting $n_0(\epsilon_l) \to n_0(\tilde{E}_l)$ in the argument of the right-hand side of the above equation, where $\tilde{E}_l \equiv \epsilon_l - \hat{e}^2 \kappa n_0(\epsilon_l)$ is the renormalized energy of the charge at site l, i.e., one writes

$$\langle \bar{n}_l \rangle = n_0 [\epsilon_l - \hat{e}^2 \kappa n_0(\tilde{E}_l)] . \qquad (5.10')$$

The above expression for $\tilde{E}_l$ says that if a localized state is occupied, i.e., $\epsilon_l < 0$, then its energy is reduced (on the average) by an amount $\hat{e}^2 \kappa$. We can now go back and interpret the Hartree term in the free energy. We saw that

$$\langle F_H; \text{rings} \rangle = - \hat{e}^2 \kappa \int_{E_v}^{E_c} n_0^2(E_i) g(E_i) dE_i .$$

At $T = 0\,°K$, $n_0(E_i^v)$, for example, is a unit step function that is zero for $E_i^v < \mu$ and unity for $E_i^v > \mu$, and therefore so is $n_0^2(E_i^v)$. Hence

$$\int n_0^2(E_i^v) g(E_i^v) dE_i^v$$

is just the number of quasiholes in the valence band in the "noninteracting system." Since the system as a whole is electrically neutral we may rewrite $\langle F_H \rangle$ as $- \frac{1}{2} \hat{e}^2 \kappa N_0^{\text{ch}}$ where N_0^{ch} is the total number of charges in the noninteracting system. This is simply the number of unoccupied valence states plus the number of occupied conduction states (or in the quasiparticle language introduced in Sec. II, the number of occupied localized states). Since we have already shown that if a localized state is occupied then its energy is lowered due to interactions, it follows that the effect of interactions is to lower the energy of the ground state of the noninteracting system by an amount $\frac{1}{2} \hat{e}^2 \kappa$ for every localized state of the system that is occupied (the factor of $\frac{1}{2}$ corrects for the double counting of the contribution of a charge to the total electrostatic energy).

Let us now ask the question of how the mean-field result (or equivalently the Hartree approximation) relates to the exact ground state at $T = 0\,°K$. Since our Hamiltonian is diagonal in the occupation number representation it follows that there is a configuration $\{n_i^0\}$ corresponding to the ground state. Once this is known, the energy of a single-quasiparticle excitation can be determined exactly. This allows us to test the supposed ground state for consistency. For the ground state, $n_i \equiv n_i^0 = 0$ or 1. Consider all single excited states of energy E_l^* in which only one n_l differs from n_l^0. We can write for the ground-state energy

$$E_G = \sum_i \epsilon_i n_i^0 + \frac{1}{2} \sum_{ij} n_i^0 q_i \hat{e} V_{ij} \hat{e} q_j n_j^0 . \qquad (5.11)$$

The statement that E_G is the ground state implies that

$$E_l^* - E_G = (\epsilon_l + \sum_j q_l \hat{e} V_{lj} \hat{e} q_j n_j^0)(n_l - n_l^0) > 0$$

$$\text{for all } l. \qquad (5.12)$$

Now, we have

$$n_l = 1 \quad \text{if } n_l^0 = 0 ,$$

$$n_l = 0 \quad \text{if } n_l^0 = 1 .$$

Therefore, we have

$$n_l - n_l^0 = +1 \quad \text{if } n_l^0 = 0$$

$$= -1 \quad \text{if } n_l^0 = 1 . \qquad (5.13)$$

This means that $(E_l^* - E_G) > 0$ for all l if

$n_l^0 = 0$ when $\epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_j^0 > 0$,

$n_l^0 = 1$ when $\epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_j^0 < 0$. $\qquad$ (5.14)

But this is precisely the set of equations satisfied in the mean-field (or Hartree) approximation at $T = 0\,°K$:

$$n_l^0 = \lim_{T \to 0} f\left(\epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_j\right), \qquad (5.15)$$

where f is the appropriate Fermi function [in our case the function n_0 given by Eq. (4.6)]. We thus see that for the particular problem under consideration the Hartree approximation gives the exact ground state at $T = 0\,°K$. This means that the correlation energy must be zero at $T = 0\,°K$. This is in contrast to the case of a system with dynamics (an electron gas, for example) where one has a finite correlation energy at absolute zero. In the formalism employed in the present work, correlations are expressed in terms of fluctuations about the mean field. We shall see in Sec. VI that the free energy due to fluctuations does, indeed, go to zero as $T \to 0$, in accord with the observation made above. It would be interesting to show, independently, that the fluctuations about the mean field do, in fact, correspond to correlations in the system. In Appendix B we shall derive the lowest-order correlation energy using the generalized self-consistent field theory[12] and show that it is identical to the lowest-order fluctuation term.

Finally let us address ourselves to the discussion of single-quasiparticle excitations. Let us define $|\mathcal{E}_l|$ as the excitation energy of a quasiparticle at site l, where $\mathcal{E}_l = (E_l^* - E_G)/(n_l - n_l^0)$. It follows from Eq. (5.12) that

$\mathcal{E}_l = (\epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_j^0)$.

Clearly, this is explicitly dependent on the specific configuration $(\{R\}, \{\epsilon\})$. Hence we must average this over all configurations except ϵ_l, i.e.,

$\langle \mathcal{E}_l \rangle = \langle (\epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_j^0) \rangle_{\neq \epsilon_l}$. $\qquad$ (5.16)

This can be written as

$\langle \mathcal{E}_l \rangle = \lim_{T \to 0} \left\langle \epsilon_l + \hat{e}^2 \sum_j q_l q_j V_{lj} n_0 \left(\epsilon_j + \hat{e}^2 \sum_k q_j q_k V_{jk} n_k\right) \right\rangle$.

This can be rewritten to read

$\langle \mathcal{E}_l \rangle = \lim_{T \to 0} \langle \epsilon_l + \hat{e}' q_l \phi_l^0 \rangle_{\neq \epsilon_l}$ $\qquad$ (5.17)

[see Eq. (5.4)]. We have already performed this average in the context of our discussion of the mean occupation number [Eq. (5.9)]. We therefore get

$\langle \mathcal{E}_l \rangle \approx \lim_{T \to 0} [\epsilon_l - \hat{e}^2 \kappa n_0(\epsilon_l)]$. $\qquad$ (5.18)

This equation has a very interesting interpretation. It says that, at $T = 0\,°K$, if a localized state l is not occupied [i.e., $\epsilon_l > 0$ and hence $n_0(\epsilon_l) = 0$]
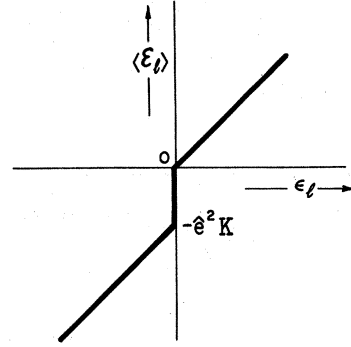

FIG. 4. We have plotted the average excitation energy of a single quasiparticle at site l, $\langle \mathcal{E}_l \rangle$, vs the single-particle energy of that site. It shows that there is a gap in the excitation spectrum corresponding to removing a charge from its correlation hole.

then the excitation energy of a single quasiparticle at that site is simply the single-particle energy of that site ϵ_l. However, if the site l were occupied [$\epsilon_l < 0$ and hence $n_0(\epsilon_l) = 1$] then the excitation energy is given by

$$|\langle \mathcal{E}_l \rangle| = |\epsilon_l - \hat{e}^2 \kappa| . \qquad (5.19)$$

This is not a surprising result. It says that originally if a site was not occupied then it does not polarize the medium in its immediate vicinity and hence $\mathcal{E}_l = \epsilon_l$. If the site were occupied, however, then it creates a polarization cloud surrounding it and with which it interacts. The effect of this, as was already mentioned, is to lower the energy of that site. So that in order to excite a quasiparticle at that site we have to remove it from its "correlation hole." These features are illustrated in Fig. 4 where we have plotted $\langle \mathcal{E}_l \rangle$ vs ϵ_l. It clearly shows that there is a gap of approximate value $\hat{e}^2 \kappa$ in the excitation spectrum of single quasiparticles. In Sec. VII we shall return to this point and discuss the consequences of this for hopping conduction.

VI. FLUCTUATIONS

Let us now turn to the question of evaluating the fluctuations about the mean field. We saw earlier that the partition function can be factored into a product of the mean-field result times the fluctuations about the mean field. It follows from Eq. (4.4) that

$$Z_{\text{fluct}} = L \int \delta \phi$$

$$\times \exp\left(-\frac{1}{2} \sum_{ij} \bar{\phi}_i \left[\frac{\delta^2 B}{\delta \phi_i \delta \phi_j}\right]_{\phi^0} \bar{\phi}_j + O(\bar{\phi}^3)\right), \qquad (6.1)$$

where $B[\{\phi\}]$ is given by Eq. (4.2) and $\{\bar{\phi}\} \equiv \{\phi - \phi^0\}$. We have

$$\delta^2 B/\delta\phi_i \delta\phi_j = \beta [V_{ij}^{-1} - \hat{e}^2 n_0^{(1)}(\epsilon_i + \hat{e}'q_i \phi_i)], \quad (6.2)$$

where

$$n_0^{(p)}(x) \equiv \frac{\partial^p n_0(x)}{\partial x^p}.$$

Let us define a matrix

$$\tilde{\underline{\Lambda}} \equiv \mathrm{diag}\{-\hat{e}^2 n_0^{(1)}(\epsilon + \hat{e}'q\phi)\}. \quad (6.3)$$

We can now rewrite (6.1) as

$$Z_{\mathrm{fluct}} = \int \delta\bar{\phi} \exp\left(-\frac{\beta}{2}\sum_{ij}\bar{\phi}_i(\underline{V}^{-1} + \tilde{\underline{\Lambda}})_{ij}\bar{\phi}_j + O(\bar{\phi}^3)\right)\bigg/\int \delta\bar{\phi}\exp\left(-\frac{\beta}{2}\sum_{ij}\bar{\phi}_i V_{ij}^{-1}\bar{\phi}_j\right). \quad (6.4)$$

If we ignore terms of $O(\bar{\phi}^3)$ the above integral can be readily performed and we get

$$Z_{\mathrm{fluct}} = (\det[\underline{I} + \tilde{\underline{\Lambda}}\,\underline{V}])^{-1/2}. \quad (6.5)$$

Therefore the contribution to the free-energy due to lowest-order fluctuations is

$$\begin{aligned}
F_{\mathrm{fluct}} &= -kT\ln Z_{\mathrm{fluct}}\\
&= \tfrac{1}{2}kT\ln\left|\underline{I} + \tilde{\underline{\Lambda}}\,\underline{V}\right|\\
&= \tfrac{1}{2}kT\ln\left|\tilde{\underline{\epsilon}}\right|. \quad (6.6)
\end{aligned}$$

We observe that the dielectric function $\tilde{\underline{\epsilon}}$ that enters Eq. (6.6) describes the "redistribution" of dressed particles. This is, of course, as it should be since we are describing fluctuations "about" the mean field.

Next we have to perform the configurational average of $\ln|\tilde{\underline{\epsilon}}|$. As a first approximation we shall, in describing the fluctuations, ignore the local fields. That is, we shall replace $\Lambda(\epsilon + \hat{e}'q\phi^0)$ by $\Lambda(\epsilon)$ in (6.6). Also, we can write $\ln|\underline{A}| = \mathrm{Tr}\ln\underline{A}$. We have, therefore,

$$\langle F_{\mathrm{fluct}}\rangle = \tfrac{1}{2}kT\langle\mathrm{Tr}\ln(\underline{I} + \underline{\Lambda}\,\underline{V})\rangle. \quad (6.7)$$

The way we perform the average is as follows. We formally expand $\ln(\underline{I} + \underline{\Lambda}\,\underline{V})$ in powers of $\underline{\Lambda}\,\underline{V}$, viz.,

$$\langle\mathrm{Tr}\ln(\underline{I} + \underline{\Lambda}\,\underline{V})\rangle = \langle\mathrm{Tr}\underline{\Lambda}\,\underline{V}\rangle - \tfrac{1}{2}\langle\mathrm{Tr}(\underline{\Lambda}\,\underline{V})(\underline{\Lambda}\,\underline{V})\rangle + \cdots. \quad (6.8)$$

We then classify the various terms in the above expansion by graphs. We have illustrated in Fig. 5 the various diagrams that correspond to the fifth term in the expansion. We argue, once again, that the most singular contribution in any order

comes from the ring diagrams. These may be summed to yield

$$\langle\mathrm{Tr}\ln(\underline{I} + \underline{\Lambda}\,\underline{V}); \mathrm{rings}\rangle = -(\Omega/6\pi)\kappa^3. \quad (6.9)$$

Once again κ is the inverse screening length given by Eq. (4.22). We have, finally,

$$\langle F_{\mathrm{fluct}}; \mathrm{rings}\rangle = -kT(\Omega/12\pi)\kappa^3. \quad (6.10)$$

It may be recalled that this is precisely the form of the lowest-order contribution to the equation of state of an ionic solution due to long-range forces. There the relevant screening length is the Debye-Hückel screening length. In our case, however, κ is the inverse Thomas-Fermi screening length and as such has a temperature-independent limit as $T\to 0$, so that $F_{\mathrm{fluct}}\to 0$ as $T\to 0$, as anticipated.

Before closing this section we shall estimate the orders of the higher-order fluctuations about the mean field. We saw that

$$Z_F = L\int\delta\phi\exp\left(-\frac{1}{2}\sum_{ij}\frac{\delta^2 B}{\delta\phi_i^0\,\delta\phi_j^0}\bar{\phi}_i\bar{\phi}_j\right)$$

$$\times\exp\left(-\sum_{p\geq 3}\frac{1}{p!}\sum_{ij\cdots p}\frac{\delta^p B}{\delta\phi_i^0\cdots\delta\phi_p^0}\bar{\phi}_i\cdots\bar{\phi}_p\right), \quad (6.11)$$

where $B[\{\phi\}]$ is given by Eq. (4.2) and $\delta^2 B/\delta\phi_i\delta\phi_j$ is given by Eq. (6.2). Also, we have for the pth variational derivative of the functional $B[\phi]$

$$\frac{\delta^p B}{\delta\phi_i\delta\phi_j\cdots\delta\phi_p} = \beta(\hat{e}'q_i)^p n_0^{(p)}$$

$$\times(\epsilon_i + \hat{e}'q_i\phi_i)\delta_{ip}\delta_{jp}\cdots\delta_{(p-1)p}. \quad (6.12)$$

Let us introduce a set of new variables $\chi \equiv \bar{\phi}/(kT)^{1/2}$. In terms of these we can write (6.11) as

$$Z_F = \int\delta\chi\exp\left(-\frac{1}{2}\sum_{ij}\chi_i(\underline{V}^{-1} + \underline{\Lambda})_{ij}\chi_j\right)\exp\left(-\sum_i\sum_{p\geq 3}\frac{1}{p!}[\hat{e}']^p[kT]^{p/2-1}q_i^p n_0^{(p-1)}(i)\chi^p\right)\bigg/\int\delta\chi\exp\left(-\frac{1}{2}\sum_{ij}\chi_i V_{ij}^{-1}\chi_j\right). \quad (6.13)$$

One can already anticipate, from the structure of the above equations, that if we retained the second exponential factor in the integrand of the numerator then we would get a result which would be higher

order in both kT and e. To see this more explicitly, we shall expand the second exponential factor in powers of the exponent, i.e., we can write (6.13) as

$$Z_F = \left| \underline{I} + \tilde{\underline{\Lambda}} \, \underline{V} \right|^{-1/2} L' \int \delta\chi \, \exp\left(-\frac{1}{2} \sum_{ij} \chi_i (\underline{V}^{-1} + \tilde{\underline{\Lambda}})_{ij} \chi_j \right)$$

$$\times \left(1 - \sum_k \frac{(\hat{e}' q_k)^3}{3!} (kT)^{1/2} n_0^{(2)}(E_k) \chi^3 - \sum_k \frac{(\hat{e}' q_k)^4}{4!} (kT) n_0^{(3)}(E_k) \chi^4 + O(\chi^5) \right) , \quad (6.14)$$

where $E_k \equiv \epsilon_k + \hat{e}' q_k \phi_k^0$ and

$$(L')^{-1} \equiv \int \delta\chi \, \exp\left(-\frac{1}{2} \sum_{ij} \chi_i (\underline{V}^{-1} + \tilde{\underline{\Lambda}})_{ij} \chi_j \right). \quad (6.15)$$

The determinant in front of the integral in (6.14) results from a change of normalization. The first term in the curly brackets in (6.14) corresponds, of course, to the neglect of terms of $O(\chi^3)$ in Z_{fluct} and yields the result derived before [Eq. (5.5)]. The first nonvanishing correction to this would come from the χ^4 term in (6.14). Upon integrating this gives

$$-kT(e^4/4!)\sum_k n_0^{(3)}(E_k)\left[(\underline{V}^{-1} + \Lambda)_{kk}^{-1} \right]^2 . \quad (6.16)$$

Recalling that $(\underline{V}^{-1} + \Lambda)_{ii}^{-1}$ is just the screened potential V_{ii}^{sc}, one can write

$$Z_F = \left| \underline{I} + \underline{\Lambda} \, \underline{V} \right|^{-1/2}$$

$$\times \left(1 - kT \frac{e^4}{4!} \sum_k n_0^{(3)}(E_k)(V_{kk}^{\text{sc}})^2 + O[(kT)^2] \right). \quad (6.17)$$

Therefore we have

$$F_{\text{fluct}} = -kT \ln Z_{\text{fluct}}$$

$$= \frac{kT}{2} \ln \left| \underline{I} + \tilde{\underline{\Lambda}} \, \underline{V} \right|$$

$$+ (kT)^2 \frac{e^4}{4!} \sum_k n_0^{(3)}(E_k)(V_{kk}^{\text{sc}})^2 + O[(kT)^3] . \quad (6.18)$$

The configurational average of the second term will be hard to perform, but one can anticipate that the leading term in $\langle (V_{kk}^{\text{sc}})^2 \rangle$ will be $\sim \kappa^2$, i.e., of $O[e^2]$ and

$$\langle n_0^{(3)}(E_k) \rangle \to \left. \frac{d^2 g}{dE^2} \right|_\mu \quad \text{as } T \to 0 .$$

So that, in terms of the width of the tail of the density of states Δ introduced in Sec. IV

$$\left(\Delta^{-1} \approx \left. \frac{d \ln g}{dE} \right|_\mu \right) ,$$
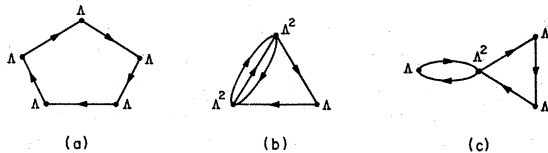

(a)   (b)   (c)

FIG. 5. Various graphs corresponding to the fifth term in the expansion of $\ln(\underline{I} + \Delta V)$ in powers of ΔV. The ring diagram (a) yields the dominant contribution.

we may write

$$\langle F_{\text{fluct}} \rangle \approx -kT\Omega\kappa^3$$

$$\times \left(1 + \frac{kT}{\Delta^2} O[e^3] + \cdots + \left(\frac{kT}{\Delta^2}\right)^{N-1} O[e^{3(N-1)}] \right) .$$

$$(6.19)$$

This means that for $kT < \Delta^2$ and for low concentration of charges the fluctuations of $O[(\phi - \phi^0)^3]$ can be ignored.

As may be seen from (6.6) the contribution to the free energy from fluctuations of $O[(\phi - \phi^0)^2]$ has in it a term proportional to kT. This would imply that at $T = 0\,°K$ there is a finite entropy of the system, i.e.,

$$S_{\text{fluct}} = -\frac{\partial F_{\text{fluct}}}{\partial T} \to \text{const} \quad \text{as } T \to 0 .$$

We know, however, that since there is no macroscopic degeneracy in our system there can be no net entropy at $T = 0\,°K$. This means that the term linear in kT in F_{fluct} must be cancelled by a corresponding term in F_{mf}. To prove that such a cancellation does indeed occur one would have to extract the term linear in kT in F_{mf} in the full nonlinear theory. At this point we have not been able to demonstrate this.

VII. DISCUSSION

Let us next turn to some of the consequences of the results of Sec. V on electronic structure and transport in amorphous semiconducting alloys. We saw that the effect of interaction between the charges is to lower the energy of the charged sites (there is no effect due to interactions on the neutral sites). This has an interesting effect on the single-particle density of states of the noninteracting system (Fig. 1). Consider, for example, the conduction-band tail. At $T = 0\,°K$ all states above the chemical potential are unoccupied (and hence neutral) and the density of states is unaffected by the interactions. On the other hand at $T = 0\,°K$ all states below μ are occupied and their energies are lowered by an amount $\hat{e}^2 \kappa$. In Fig. 6 we have displayed the renormalized density of states for the conduction-band tail. It shows clearly that in order to excite an electron from an occupied conduction state to an empty conduction state it will cost at least $\hat{e}^2 \kappa$ in energy. We have also displayed, in Fig. 6, the occupation number of charges of the interacting system, $\bar{n}(E)$, as a function of the re-
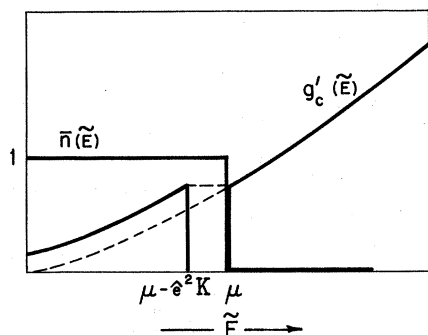
FIG. 6. We have displayed the mean occupation number $\bar{n}(\tilde{E})$ of quasielectrons in the conduction-band tail as a function of their renormalized energy. We have also shown the renormalized single-quasiparticle density of states for the electrons in the conduction-band tail. The effect of Coulomb interaction is to introduce a gap in the density of states.

normalized energy of the localized states. It is once again an unit step function (at $T = 0\,°K$) as in the case of the noninteracting system. There are, however, "no available states" in the interval $\mu - \hat{e}^2\kappa < \tilde{E} < \mu$. It may be recalled that such a well-defined gap resulted as a consequence of our earlier assumption that the distribution of ϕ^0 in various configurations is sharply peaked. More specifically, such an assumption implied that the energy of every charged site was lowered by precisely the same amount, viz., $\hat{e}^2\kappa$. We know that in a random system the distribution of ϕ^0 will have a certain width. The effect of the statistical fluctuations would be to smear out the average energy gap. In order to be able to say precisely how the fluctuations will modify (say) Fig. 6 one would have to develop a probabilistic theory of the self-consistent field. We know, however, that the renormalized density of states will have a discontinuity at the chemical potential. It is perhaps reasonable to anticipate that after the fluctuations have been taken into effect the renormalized density of states
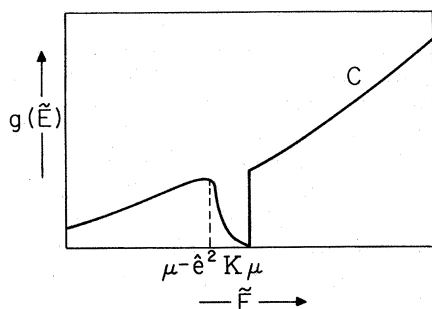

FIG. 7. We have displayed an ansatz of how the density of states would look if one did a probabilistic theory of the self-consistent field.
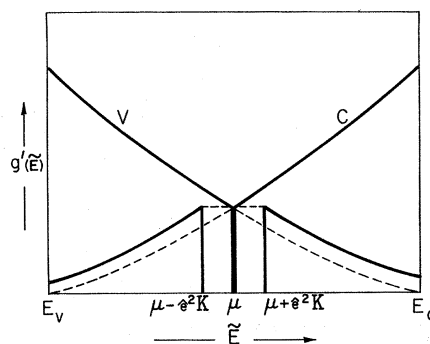

FIG. 8. Renormalized density of states corresponding to the valence- and conduction-band tails in the Mott-CFO model.

would look as shown in Fig. 7. In Fig. 8 we have displayed the complete renormalized density of states. As may be seen from Fig. 8, whereas the distribution of charged traps is bimodal in character with peaks separated by approximately $2\hat{e}^2\kappa$, there is no gap from the top of the occupied region of the valence-band tail which ends at the Fermi energy to the unoccupied region of the conduction-band tail which begins at the Fermi energy. The total density of states at the Fermi energy is reduced by a factor of 2 below its value in the absence of interactions.

The ac conductivity of all amorphous semiconductors exhibits a dependence on frequency proportional to ω^S where the power S is between 0.7 and 1.0, and becomes less and less temperature dependent at the higher frequencies. These two facts have been used as evidence that the ac conductivity is due to phonon-assisted hopping of electrons between the localized states near the Fermi level.[13] It turns out that one needs a very large density of states at the Fermi level to explain the experimental results on the basis of hopping conduction. We have seen that the effect of interaction between the charged states in the gap is to reduce the total density of states at the Fermi energy by a factor of 2. Also, according to what we have said it will cost $\hat{e}^2\kappa$ in energy to remove an electron from a charged conduction state and put it in an empty conduction state. In fact, this activation energy $\Delta E \equiv \hat{e}^2\kappa$ is directly related to the total density of states at the Fermi level $[\kappa \propto g^{1/2}(\mu)]$. In Fig. 9 we have plotted this activation energy vs $g^{tot}(\mu)$. As may be seen from it, a large density of states at the Fermi level can produce an energy gap there (in the above sense) as large as 100 °K. It would, therefore, be interesting to reconsider the theoretical discussion of hopping conduction in the light of these remarks.

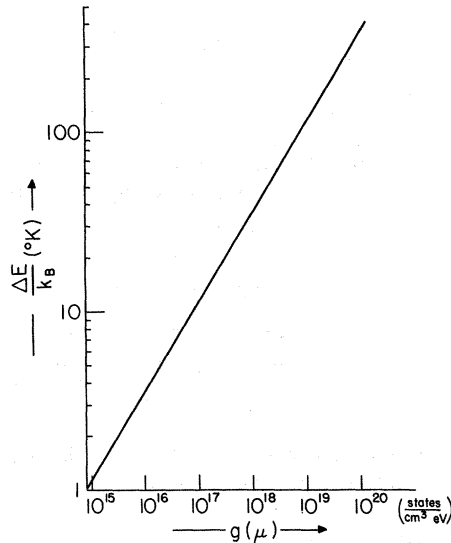In all of the preceding discussions we have as-

FIG. 9. Plot of the energy gap $\Delta E \equiv \hat{e}^2 \kappa$ (expressed in °K) vs the density of states at the Fermi level. This gap may be interpreted as representing an activation energy for hopping conduction between localized states near the Fermi level. We have assumed a value of 15 for the static dielectric constant of the amorphous semiconductor.

sumed that the chemical potential μ is unaffected by interactions. In general, of course, it will be and the correct chemical potential should be determined by the relation

$$\frac{\partial \langle F \rangle}{\partial \mu} = N^+ - N^- \, ,$$

where N^+ and N^- are the total number of positive and negative charges, respectively, in the interacting system. Since the system as a whole is electrically neutral the chemical potential should be determined from the relation $\partial \langle F \rangle / \partial \mu = 0$. Since we have dealt with symmetrical bands the chemical potential is unaffected by interaction between the charges.

VIII. SUMMARY

Within the premises of the Mott-CFO model for a covalent amorphous semiconducting alloy we have investigated the effect of interaction between the charged traps in the mobility gap. We expressed the grand partition function as a functional integral over a set of Gaussian random fields. We evaluated this integral in the mean-field approximation and showed that the mean field corresponds to the generalized Hartree self-consistent field. We then demonstrated that for the system under consideration the Hartree approximation represented the exact ground state at $T = 0$ °K. We then derived an approximate expression for the mean occupation number and the renormalized band structure. It was found that the effect of Coulomb interaction

between the charged states was to render the distribution of charged traps in energy bimodal in character with peaks separated by approximately $2e^2\kappa$ and to reduce the density of states at the Fermi energy by a factor of 2 below its value in the noninteracting system.

There is convincing evidence that the localized states in an amorphous semiconductor are associated with potential fluctuations.[14] It would be very interesting to investigate what effect the local fields, which we saw result from self-consistent interaction between the localized charges, would have on the mobile electrons. That is, one should study the motion of electrons in a system with fluctuating potential plus randomly distributed Coulomb fields.

APPENDIX A

In Sec. IV we performed the configurational average of the Hartree term in the free energy. We expanded the inverse dielectric matrix, $(\underline{I} + \underline{\Lambda}\,\underline{V})^{-1}$, in powers of $\underline{\Lambda}\,\underline{V}$ [see Eq. (4.19)] and averaged term by term. We then argued that in each order in V_{ij} the most singular contribution comes from the ring diagrams, and we summed the contributions of ring diagrams of all orders. This yielded a result of $O(e^3)$. Let us now estimate the orders of some of the other diagrams. Consider, for example, Fig. 2(a). It will correspond to a term

$$\sum_{i12j} \langle n_0(\epsilon_i)q_i \rangle \langle n_0(\epsilon_j)q_j \rangle \langle \Lambda_1^2 \rangle \langle \Lambda_2^2 \rangle \langle V_{i2}V_{12}^2 V_{2j} \rangle \, .$$

We know that

$$\langle n_0(\epsilon_i) \rangle = \int_{E_v}^{E_c} n_0(E_i)g(E_i)dE_i \Big/ \int_{E_v}^{E_c} g(E_i)dE_i \, .$$

Suppose i were a valence site, we would then have

$$\langle n_0(\epsilon_i^v) \rangle = N_0^+ / N_v^L \, ,$$

where N_0^+ is the number of quasiholes in the valence-band tail, in the noninteracting system. Therefore it is easy to see that, because of the factor q_i, the summation over i would give a factor $(N_0^+ - N_0^-)$; and since electrical neutrality is preserved in the noninteracting system this would vanish. Thus diagrams with loose ends $[q_i n_0(\epsilon_i)]$ would vanish due to electrical neutrality.

Since we know that the divergences that occur in the averaging procedure are due to the long-range nature of the Coulomb force, one can anti-

cipate that diagrams with single V_{ij} bonds will dominate the averaging over those with multiple bonds.[9] In fact, the most dominant of these are the simple cycles or the rings. The next order contributions will come from diagrams of the type shown in Figs. 2(g) and 2(h). We have summed all such diagrams to all orders and we find that they give a contribution to the free energy of $O[e^6]$. We give the explicit result for the contribution to the free energy from diagrams of the type given in Fig. 2(g) (for example):

$$\langle F \rangle = -\frac{4\pi}{\Omega} e^6 \int_{E_v}^{E_c} n_0^2(E_i) g(E_i) dE_i \int_{E_v}^{E_c} \left(-\frac{\partial n_0}{\partial \epsilon_i}\right)^2 g(E_i) dE_i ,$$

i.e., F is of $O[e^6]$. We recall that the rings gave a contribution of $O[e^3]$.

APPENDIX B

We shall now derive the lowest-order correlation energy for the system we have been considering, using the dielectric formulation of Cohen.[12] We shall adopt the notation used in the reference cited above and we shall suppress the details. The calculation follows very closely Sec. IV of the above reference. The essential point of the generalized self-consistent field theory[12] is that while calculating the total energy of a system of interacting charges one parametrizes the energy by some parameter ξ (different from e^2) in such a way that for $\xi = 0$ the parametrized energy reduces to that of a problem which is closer to the real problem than the noninteracting system, and for $\xi = 1$ becomes the actual energy.

We have for the total energy of our system

$$E(\{n\}) = \sum_i (\epsilon_i + V_i^{ext}) n_i + \frac{1}{2} \sum_{ij} \hat{e}^2 q_i q_j V_{ij} n_i n_j . \tag{B1}$$

We have introduced a set of external fields V_i^{ext} with a view to employing the dielectric formulation. To obtain the free energy one starts with the expression

$$U = \sum_i (\epsilon_i + V_i^{ext}) \langle n_i \rangle + \frac{1}{2} \sum_{ij} \hat{e}^2 q_i q_j V_{ij} \langle n_i \rangle \langle n_j \rangle$$
$$+ \frac{1}{2} \xi \sum_{ij} \hat{e}^2 q_i q_j V_{ij} (\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle) \tag{B2}$$

for the average total energy in thermal equilibrium, where $\langle \mathcal{O} \rangle = \mathrm{Tr}(\mathcal{O}P)$, P being the density matrix. The free energy is given by

$$F = U + kT[\mathrm{Tr}(P \ln P)]. \tag{B3}$$

Applying the variational principle $\delta F = 0$ subject to $\delta \mathrm{Tr} P = 0$ to (B2) and (B3) we obtain for P

$$P = e^{-\beta H_S}/\mathrm{Tr}e^{-\beta H_S} , \tag{B4}$$

where

$$H_S = \sum_i [\epsilon_i + V_i^{ext} + (1-\xi) V_i^s] n_i$$

$$+ \frac{1}{2} \xi \sum_{ij} \hat{e}^2 q_i q_j V_{ij} n_i n_j , \tag{B5}$$

$$V_i^S = \sum_j \hat{e}^2 q_i q_j V_{ij} \langle n_j \rangle . \tag{B6}$$

V_i^S is the total self-consistent potential energy at site i. Because of the stationarity of F with respect to P,

$$\frac{dF}{d\xi} = \frac{1}{2} \sum_{ij} \hat{e}^2 q_i q_j V_{ij} (\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle). \tag{B7}$$

In order to get F we have to find $\langle n_i \rangle$ and $\langle n_i n_j \rangle$, and integrate (B7) with respect to ξ. We have

$$\langle n_i \rangle = \mathrm{Tr} n_i e^{-\beta H_S}/\mathrm{Tr} e^{-\beta H_S} . \tag{B8}$$

Taking the variational derivative of this with respect to the external field we get

$$-kT \frac{\delta \langle n_i \rangle}{\delta V_j^{ext}} = \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle$$
$$+ (1-\xi) \sum_{kl} \hat{e}^2 q_k q_l V_{kl} \frac{\delta \langle n_i \rangle}{\delta V_j^{ext}} (\langle n_i n_k \rangle - \langle n_i \rangle \langle n_k \rangle). \tag{B9}$$

Let us define

$$C_{ij} \equiv \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle , \tag{B10}$$

$$K_{ij} \equiv \delta \langle n_i \rangle / \delta V_j^{ext} . \tag{B11}$$

Therefore (B9) reads (written as a matrix equation)

$$-kT \underline{K} = \underline{C}[\underline{I} + (1-\xi)\underline{V}\underline{K}], \tag{B12}$$

where we have redefined $V_{kl} = \hat{e}^2 q_k q_l V_{kl}$ and therefore

$$\underline{C} = -kT \underline{K}[\underline{I} + (1-\xi)\underline{V}\underline{K}]^{-1} . \tag{B13}$$

Hence we have

$$\frac{dF}{d\xi} = \frac{1}{2} \mathrm{Tr}(\underline{V}\underline{C}) . \tag{B14}$$

Integrating this we get

$$F = F(0) - \frac{1}{2}(kT) \int_0^1 d\xi \, \mathrm{Tr}\{\underline{V}\underline{K}[\underline{I} + (1-\xi)\underline{V}\underline{K}]^{-1}\}. \tag{B15}$$

Let us define $V_i^{tot} = V_i^S + V_i^{ext}$. The "dielectric function" ϵ is defined through the variational derivative

$$\frac{\delta V_i^{tot}}{\delta V_j^{ext}} \equiv \epsilon_{ij}^{-1} = \delta_{ij} + \frac{\delta V_i^S}{\delta V_j^{ext}} . \tag{B16}$$

From (B6) and (B11) we get

$$\underline{\epsilon}^{-1} = \underline{I} + \underline{V}\underline{K} . \tag{B17}$$

Substituting (B17) into (B15) we obtain

$$F = F(0) + \frac{1}{2}(kT) \int_0^1 d\xi \, \mathrm{Tr}\{(\underline{\epsilon} - \underline{I})[\underline{I} + (1-\xi)(\underline{\epsilon} - \underline{I})]^{-1}\} . \tag{B18}$$

ϵ is, in general, a function of the parameter ξ. For $\xi = 0$, the dielectric function reduces to the

Hartree dielectric function ϵ_H. In terms of this the integral in (B18) can be performed and we get

$$F = F(0) + \tfrac{1}{2}(kT)\,\mathrm{Tr}\,\ln\epsilon_H .\qquad\qquad (B19)$$

$F(0)$ is just the self-consistent field result and the second term is the lowest-order correlation energy. But this is precisely what we obtained as the contribution to the free energy from the fluctuations of $O[\bar\phi^2]$. We have thus demonstrated that the fluctuations about the mean field do represent correlations in the system.

---

[1] N. F. Mott, Advan. Phys. 16, 49 (1967).

[2] M. H. Cohen, H. Fritzsche, and S. R. Ovshinsky, Phys. Rev. Letters 22, 1065 (1969).

[3] E. N. Economou, S. Kirkpatrick, M. H. Cohen, and T. P. Eggarter, Phys. Rev. Letters 25, 520 (1970); T. P. Eggarter and M. H. Cohen, ibid. 25, 807 (1970), and references contained therein.

[4] J. E. and M. G. Mayer, Statistical Mechanics (Wiley, New York, 1940).

[5] L. D. Landau and E. M. Lifshitz, Statistical Physics (Pergamon, New York, 1958), Chap. VII.

[6] J. E. Mayer, J. Chem. Phys. 18, 1426 (1950).

[7] H. Cramer, Mathematical Methods of Statistics (Princeton U.P., Princeton, N. J., 1946), p. 118.

[8] A. J. F. Siegert, in Brandeis University Summer Institute Lectures in Theoretical Physics, 1962 (Benjamin, New York, 1963), Vol. 3.

[9] H. L. Friedman, Ionic Solution Theory (Interscience, New York, 1962).

[10] In principle one should use the complete expression for F in (5.1). In this section, however, we shall be primarily interested in the mean occupation number at absolute zero. We shall see, in Sec. VI, that the free energy due to fluctuations vanishes at $T = 0$ °K. At finite temperatures the fluctuations will yield corrections to (5.2).

[11] We have evaluated the dominant contribution to $\langle\phi^2\rangle$ and it yields $\pi(e^4/\kappa)(N^L/\Omega)\langle f^2\rangle f''(\epsilon_l)$.

[12] M. H. Cohen, Phys. Rev. 130, 1301 (1963).

[13] I. G. Austin and N. F. Mott, Advan. Phys. 18, 14 (1969); M. Pollak and T. H. Geballe, Phys. Rev. 122, 1742 (1961).

[14] M. H. Cohen, in Proceedings of the Tenth International Conference on the Physics of Semiconductors, Cambridge, Mass., 1970 (AEC Division of Technical Information, Washington, D.C., 1970), p. 645.