Excitonic Insulator*

D. JÉROME,† T. M. RICE,‡ AND W. KOHN
University of California, San Diego, La Jolla, California
(Received 30 January 1967)

This paper presents theoretical considerations of a new kind of insulating phase which has recently been theoretically predicted but has as yet not been found experimentally. This phase is expected to occur when semiconductors with very small band gap or semimetals with very small band overlap are cooled to a sufficiently low temperature. The present paper first develops a BCS-like theory of the ground state and analyzes the nature of the response to a general perturbation, from which collective modes (of a sound-like nature), response to a static magnetic field, and conductivity are calculated. Finally, some discussion of the possible experimental realization of this new phase is presented.

I. INTRODUCTION

SEVERAL years ago Mott[1] made the observation that in a semimetal the electrons and holes will under certain circumstances form bound pairs, thereby leading to a nonconducting state. Shortly afterwards, Knox[2] made the remark that if, in a conventional insulator, the binding energy of an exciton $|E_B|$ exceeded the energy gap G, the normal insulating ground state (filled valence band) would be unstable against the formation of excitons.

From these observations, and subsequent developments,[3,4] there has now emerged the realization[5–9] that for solids with small energy gaps, there may exist, at sufficiently low temperatures, a new phase which we call the excitonic insulator and whose "phase diagram" is schematically shown[6] in Fig. 1.

Several studies of the properties of this phase have recently appeared in the literature. In the present paper we wish to report theoretical work which we have done over the last year and a half, which partly overlaps and extends those studies.

The paper consists of seven sections. Section II deals with the ground state in a manner formally analogous to the BCS theory of superconductivity. However, we show that the order is not of the off-diagonal kind,[10] characteristic of superfluid systems, but diagonal, as in an ordinary crystal. In Sec. III we derive a set of equations to describe the response of the system to electromagnetic perturbations used in later sections.

The translational symmetry of the ground state is broken by the introduction of a new characteristic wave vector $\mathbf{w}$ which represents the difference between the positions of one of the conduction-band minima and valence-band maxima in momentum space. Associated with this broken symmetry we find collective, longitudinal excited modes which have the dispersion relation of a longitudinal sound wave but are of entirely electronic (rather than ionic) character.[7] These are the subject of Sec. IV.

Section V deals with the response, in the ground state, to a static vector potential. In particular, we verify the absence of a Meissner effect in harmony with the previously established absence of off-diagonal long-range order.

In Sec. VI we calculate the frequency-dependent conductivity as a function of temperature, by utilizing ordinary and anomalous Green's functions. By means of the Kramers-Kronig relations, we then also predict the temperature dependence of the dc conductivity.

Finally, Sec. VII deals with the question of which materials have most promise for a realization of the excitonic insulating phase in the laboratory. It appears that a combination of high pressure (10–100 kbar) and low temperature (1–100°K) techniques is required and that Sr and Yb may be good prospects.

II. THE GROUND STATE OF THE EXCITONIC INSULATOR

In this section we develop the theory of the ground state of the excitonic insulator in formal analogy with the BCS theory of superconductivity.

For simplicity, we consider a system which, in the absence of interactions has a single valence-band maximum at $\mathbf{k}=0$ and a single conduction-band minimum

* Supported in part by the U.S. Office of Naval Research.
† Present address: Laboratoire de Physique du Solide, Faculté des Sciences, Orsay, Essonnes, France.
‡ Present address: Bell Telephone Laboratories, Murray Hill, New Jersey.
[1] N. F. Mott, Phil. Mag. **6**, 287 (1961).
[2] R. S. Knox, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Suppl. 5, p. 100.
[3] J. DesCloizeaux, J. Phys. Chem. Solids **26**, 259 (1965).
[4] L. V. Keldysh and Yu. V. Kopaev, Fiz. Tverd. Tela **6**, 2791 (1964) [English transl.: Soviet Phys.—Solid State **6**, 2219 (1965)].
[5] R. G. Arkhipov, Zh. Eksperim. i Teor. Fiz. **43**, 349 (1962) [English transl.: Soviet Phys.—JETP **16**, 251 (1962)].
[6] A. N. Kozlov and L. A. Maksimov, Zh. Eksperim. i Teor. Fiz. **48**, 1184 (1965) [English transl.: Soviet Phys.—JETP **21**, 790 (1965)].
[7] A. N. Kozlov and L. A. Maksimov, Zh. Eksperim. i Teor. Fiz. **49**, 1284 (1965) [English transl.: Soviet Phys.—JETP **22**, 889 (1966)].
[8] W. Kohn, in *Physics of Solids at High Pressures*, edited by C. T. Temiznka and R. M. Emrick (Academic Press Inc., New York, 1965).
[9] E. V. Baklanov and A. V. Chaplik, Fiz. Tverd. Tela **7**, 2768 (1965) [English transl.: Soviet Phys.—Solid State **7**, 2240 (1966)]. Yu. V. Kopaev, Fiz. Tverd. Tela **8**, 223 (1966) [English transl.: Soviet Phys.—Solid State **8**, 175 (1966)].
[10] V. L. Ginzburg and L. D. Landau, Zh. Eksperim. i Teor. Fiz. **20**, 1064 (1950); O. Penrose, Phil. Mag. **42**, 1373 (1951); O. Penrose and L. Onsager, Phys. Rev. **104**, 576 (1956); C. N. Yang, Rev. Mod. Phys. **34**, 694 (1962).

at $\mathbf{k}=\mathbf{w}$. We write the single-particle energies as

$$\epsilon_a(\mathbf{k}_a) = -\tfrac{1}{2}G-(2m_a)^{-1}k_a^2, \tag{2.1}$$

$$\epsilon_b(\mathbf{k}_b) = \tfrac{1}{2}G+(2m_b)^{-1}k_b^2, \tag{2.2}$$

where $\mathbf{k}_a$ and $\mathbf{k}_b$ refer to the respective band extrema. Energies are reckoned from the center of the gap and G may be positive or negative. For negative G we have, in the absence of interactions, a semimetal with Fermi wave vector given by

$$k_F^2=2\mu\,|\,G\,|, \tag{2.3}$$

where μ is the reduced mass

$$\mu^{-1}=m_a^{-1}+m_b^{-1}. \tag{2.4}$$

We introduce creation and destruction operators as follows: $a_\mathbf{k}^*$, $a_\mathbf{k}$ create and destroy electrons in band a with wave vector $\mathbf{k}$; $b_\mathbf{k}^*$, $b_\mathbf{k}$ create and destroy electrons in band b with wave vector $\mathbf{w}+\mathbf{k}$. We also introduce the (partial) charge-density operator[11]

$$\rho(\mathbf{q})=\sum_\mathbf{k} a_{\mathbf{k}+\mathbf{q}}^* a_\mathbf{k}+b_{\mathbf{k}+\mathbf{q}}^* b_\mathbf{k}. \tag{2.5}$$

Then, following Des Cloiseaux[3] we take for our model Hamiltonian

$$H=\sum_\mathbf{k}\epsilon_a(\mathbf{k})a_\mathbf{k}a_\mathbf{k}^*+\epsilon_b(\mathbf{k})b_\mathbf{k}^*b_\mathbf{k}+\tfrac{1}{2}\sum_\mathbf{q}V(\mathbf{q})\rho(\mathbf{q})\rho(-\mathbf{q}). \tag{2.6}$$

Here the sums go over the Brillouin zone and we take

$$V(\mathbf{q})=4\pi e^2/\epsilon(\mathbf{q})q^2, \tag{2.7}$$

where $\epsilon(\mathbf{q})$ is an effective dielectric constant.

The conventional, insulating ground state is given by

$$\Phi=\prod_\mathbf{k} a_\mathbf{k}^*\,|\,\mathrm{vac}\rangle, \tag{2.8}$$

where $|\,\mathrm{vac}\rangle$ is the state with no electrons and $\mathbf{k}$ runs over the Brillouin zone. The possible instability of this state against bound-pair formation suggests that we try a new Hartree-Fock ground state of the form

$$\Psi=\prod_\mathbf{k}\alpha_\mathbf{k}^*\,|\,\mathrm{vac}\rangle, \tag{2.9}$$

where $\alpha_\mathbf{k}^*$ creates an electron in a linear combination of band a and band b states,

$$\alpha_\mathbf{k}=u_\mathbf{k}a_\mathbf{k}-v_\mathbf{k}b_\mathbf{k}, \tag{2.10}$$

$$|\,u_\mathbf{k}\,|^2+|\,v_\mathbf{k}\,|^2=1. \tag{2.11}$$

A second linearly independent combination is

$$\beta_\mathbf{k}=+v_\mathbf{k}^*a_\mathbf{k}+u_\mathbf{k}^*b_\mathbf{k}. \tag{2.12}$$

This state can also be written in the equivalent form

$$\Psi=\prod_\mathbf{k}(u_\mathbf{k}^*-v_\mathbf{k}^*b_\mathbf{k}^*a_\mathbf{k})\Phi \tag{2.13}$$

which shows, in obvious parallelism to the BCS theory, that in this state a hole in $(a,\mathbf{k})$ and an electron in $(b,\ \mathbf{k}+\mathbf{w})$ are either both present or both absent.

Minimization of the total energy with respect to $u_\mathbf{k}$ and $v_\mathbf{k}$ gives the following results

$$u_\mathbf{k}=[\tfrac{1}{2}(1+(\xi_\mathbf{k}/E_\mathbf{k}))]^{1/2}, \tag{2.14}$$

$$v_\mathbf{k}=[\tfrac{1}{2}(1-(\xi_\mathbf{k}/E_\mathbf{k}))]^{1/2}(\Delta_\mathbf{k}/|\,\Delta_\mathbf{k}\,|), \tag{2.15}$$

where

$$\xi_\mathbf{k}=\tfrac{1}{2}[\epsilon_b(\mathbf{k})-\epsilon_a(\mathbf{k})], \tag{2.16}$$

$$E_\mathbf{k}^2=\xi_\mathbf{k}^2+|\,\Delta_\mathbf{k}\,|^2 \tag{2.17}$$

and the gap function $\Delta_\mathbf{k}$ is determined by the equation

$$\Delta_\mathbf{k}=\sum_\mathbf{p}V(\mathbf{k}-\mathbf{p})(\Delta_\mathbf{p}/2[\xi_\mathbf{p}^2+|\,\Delta_\mathbf{p}\,|^2]^{1/2}). \tag{2.18}$$

The phase of $\Delta_\mathbf{k}$ is arbitrary and the energy is independent of it. For most purposes we shall take $\Delta_\mathbf{k}$ and hence $u_\mathbf{k}$ and $v_\mathbf{k}$ as real and positive.

It is instructive to define a wave function

$$\varphi_\mathbf{p}\equiv\Delta_\mathbf{p}/2[\xi_\mathbf{p}^2+|\,\Delta_\mathbf{p}\,|^2]^{1/2}. \tag{2.19}$$

Then (2.18) may be written as

$$[(G+(k^2/2\mu))^2+4\,|\,\Delta_\mathbf{k}\,|^2]^{1/2}\varphi_\mathbf{k}=\sum_\mathbf{p}V(\mathbf{k}-\mathbf{p})\varphi_\mathbf{p}. \tag{2.20}$$

Comparison with the elementary equation for the exciton wave function,

$$[(k^2/2\mu)+|\,E_B\,|]\chi_\mathbf{k}=\sum_\mathbf{p}V(\mathbf{k}-\mathbf{p})\chi_\mathbf{p}, \tag{2.21}$$

confirms that $\Delta=0$ when $G\geq|\,E_B\,|$. One may also verify[6] that nontrivial solutions (i.e., $|\,\Delta\,|>0$) exist for all values of G, positive and negative, which are $\leq|\,E_B\,|$.

In spite of the formal similarity with the BCS theory, the nature of the order in the excitonic insulator is entirely different. Yang[10] has shown that a superconductor is characterized by *off-diagonal* long-range order (ODLRO). This is manifested in the two-particle

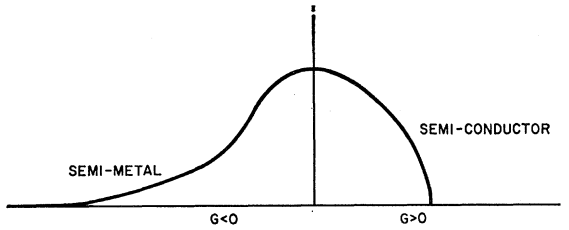

FIG. 1. The phase diagram of the excitonic insulator (Ref. 6).

[11] The total-charge-density operator is given by

$$\rho_T(\mathbf{q})=\sum_\mathbf{k}[f(\mathbf{k}+\mathbf{q},\mathbf{k})a_{\mathbf{k}+\mathbf{q}}^*a_\mathbf{k}+g(\mathbf{k}+\mathbf{w}+\mathbf{q},\mathbf{k}+\mathbf{w})b_{\mathbf{k}+\mathbf{q}}^*b_\mathbf{k}+h(\mathbf{k}+\mathbf{q},\mathbf{k})b_{\mathbf{k}-\mathbf{w}+\mathbf{q}}^*a_\mathbf{k}+h^*(\mathbf{k},\mathbf{k}+\mathbf{q})a_{\mathbf{k}+\mathbf{q}}^*b_{\mathbf{k}-\mathbf{w}}],$$

where f, g, and h are matrix elements and

$$\lim_{\mathbf{q}\to0}f(\mathbf{k}+\mathbf{q},\mathbf{k})=\lim_{\mathbf{q}\to0}g(\mathbf{k}+\mathbf{w}+\mathbf{q},\mathbf{k}+\mathbf{w})=1$$

and

$$\lim_{\mathbf{q}\to0}h(\mathbf{k}+\mathbf{q},\mathbf{k})=0.$$

density matrix,

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle \equiv \frac{\mathrm{Tr}\{e^{-\beta H}\psi^*(\mathbf{r}_2)\psi^*(\mathbf{r}_1)\psi(\mathbf{r}_1')\psi(\mathbf{r}_2')\}}{\mathrm{Tr}\{e^{-\beta H}\}}, \tag{2.22}$$

where we have suppressed spin indices. The criterion for ODLRO is that ρ_2 remains finite in the limit $|\mathbf{r}_1-\mathbf{r}_1'|\to\infty$, $\mathbf{r}_1\approx\mathbf{r}_2$ and $\mathbf{r}_1'\approx\mathbf{r}_2'$. In the BCS theory, one finds (for the well-known antiparallel spin orientation)[12]

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle = g(\mathbf{r}_1'-\mathbf{r}_1)g(\mathbf{r}_2'-\mathbf{r}_2) + f^*(\mathbf{r}_1-\mathbf{r}_2)f(\mathbf{r}_1'-\mathbf{r}_2'), \tag{2.23}$$

where

$$g(\mathbf{r}) = \Omega^{-1}\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 \exp(i\mathbf{k}\cdot\mathbf{r}), \tag{2.24}$$

$$f(\mathbf{r}) = \Omega^{-1}\sum_{\mathbf{k}} u_{\mathbf{k}}v_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{r}). \tag{2.25}$$

The second term in (2.23) leads to ODLRO.

In the excitonic insulator on the other hand we find, at $T=0$, in the Hartree-Fock approximation,

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle = \langle \mathbf{r}_2' \mid \rho_1 \mid \mathbf{r}_2\rangle\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle - \langle \mathbf{r}_2' \mid \rho_1 \mid \mathbf{r}_1\rangle\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_2\rangle, \tag{2.26}$$

where

$$\langle \mathbf{r}' \mid \rho_1 \mid \mathbf{r}\rangle = g(\mathbf{r},\mathbf{r}') - f(\mathbf{r},\mathbf{r}') \tag{2.27}$$

and, denoting the Bloch waves of bands a and b by $\varphi_{a\mathbf{k}}$ and $\varphi_{b\mathbf{k}+\mathbf{w}}$,

$$g(\mathbf{r},\mathbf{r}') = \Omega^{-1}\sum_{\mathbf{k}}\{|u_{\mathbf{k}}|^2\varphi^*_{a\mathbf{k}}(\mathbf{r})\varphi_{a\mathbf{k}}(\mathbf{r}') + |v_{\mathbf{k}}|^2\varphi_{b\mathbf{k}+\mathbf{w}}^*(\mathbf{r})\varphi_{b\mathbf{k}+\mathbf{w}}(\mathbf{r}')\}, \tag{2.28}$$

$$f(\mathbf{r},\mathbf{r}') = \Omega^{-1}\sum_{\mathbf{k}}\{u_{\mathbf{k}}^*v_{\mathbf{k}}\varphi_{b\mathbf{k}+\mathbf{w}}^*(\mathbf{r})\varphi_{a\mathbf{k}}(\mathbf{r}') + u_{\mathbf{k}}v_{\mathbf{k}}^*\varphi_{a\mathbf{k}}^*(\mathbf{r})\varphi_{b\mathbf{k}+\mathbf{w}}(\mathbf{r}')\}. \tag{2.29}$$

Both f and g are of short range in the sense that they vanish when $|\mathbf{r}-\mathbf{r}'|\to\infty$. Hence there is no ODLRO.

However, there is an additional diagonal long-range order, which we shall now discuss. Consider first, by way of background, an amorphous (disordered) solid. Then, for

$$|\mathbf{r}_1-\mathbf{r}_2|\to\infty, \quad \text{and} \quad \mathbf{r}_1'\approx\mathbf{r}_1, \quad \mathbf{r}_2'\approx\mathbf{r}_2, \tag{2.30}$$

one has, for the two-particle electronic density matrix,

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle \to \langle \mathbf{r}_2' \mid \rho_1 \mid \mathbf{r}_2\rangle\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle. \tag{2.31}$$

For a given configuration of the ions, $\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle$ is a function of both $(\mathbf{r}_1-\mathbf{r}_1')$ and $\frac{1}{2}(\mathbf{r}_1+\mathbf{r}_1')$, but if we average over the position of the center of mass of the entire system, $\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle_{\mathrm{Av}}$ is a function only of $(\mathbf{r}_1-\mathbf{r}_1')$,

$$\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle_{\mathrm{Av}} = l(\mathbf{r}_1'-\mathbf{r}_1). \tag{2.32}$$

The same average gives, under the conditions (2.30), for the two-particle density matrix

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle_{\mathrm{Av}} = \langle \mathbf{r}_2' \mid \rho_1 \mid \mathbf{r}_2\rangle_{\mathrm{Av}}\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle_{\mathrm{Av}} = l(\mathbf{r}_2'-\mathbf{r}_2)l(\mathbf{r}_1'-\mathbf{r}_1) \tag{2.33}$$

and in particular, for $\mathbf{r}_1'=\mathbf{r}_1$ and $\mathbf{r}_2'=\mathbf{r}_2$

$$\langle \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\rangle_{\mathrm{Av}} = \langle \rho_1(\mathbf{r}_1)\rangle_{\mathrm{Av}}\langle \rho_1(\mathbf{r}_2)\rangle_{\mathrm{Av}}. \tag{2.34}$$

There is no dependence on $(\mathbf{r}_2-\mathbf{r}_1)$, i.e., no long-range order.

Next consider an ordinary crystal, whose translational symmetry is characterized by the reciprocal lattice vectors $\mathbf{K}_\nu$. Here, if the ion positions are regarded as fixed, one has, under the conditions (2.30),

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle = \langle \mathbf{r}_2' \mid \rho_1 \mid \mathbf{r}_2\rangle\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle, \tag{2.35}$$

where $\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle$ has an expansion of the form

$$\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle = \sum_\nu z_\nu(\mathbf{r}_1-\mathbf{r}_1')\exp[i\mathbf{K}_\nu\cdot(\mathbf{r}_1+\mathbf{r}_1')/2]. \tag{2.36}$$

In particular, the density has the form

$$\rho(\mathbf{r}_1) = \langle \mathbf{r}_1 \mid \rho_1 \mid \mathbf{r}_1\rangle = \sum_\nu z_\nu(0)\exp(i\mathbf{K}_\nu\cdot\mathbf{r}_1). \tag{2.37}$$

If we now average over the positions of the center of mass of the crystal we find

$$\langle \mathbf{r}_1' \mid \rho_1 \mid \mathbf{r}_1\rangle_{\mathrm{Av}} = z_0(\mathbf{r}_1-\mathbf{r}_1') \tag{2.38}$$

as in the disordered case. The long-range order may, however, still be seen in the behavior of the two-particle density matrix, which now becomes

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle = \Omega^{-1}\int \sum_\nu z_\nu(\mathbf{r}_2-\mathbf{r}_2')\exp\{i\mathbf{K}_\nu\cdot[\tfrac{1}{2}(\mathbf{r}_2+\mathbf{r}_2')+\mathbf{R}]\}\times\sum_{\nu'} z_{\nu'}(\mathbf{r}_1-\mathbf{r}_1')\exp\{i\mathbf{K}_\nu\cdot[\tfrac{1}{2}(\mathbf{r}_1+\mathbf{r}_1')+\mathbf{R}]\}d\mathbf{R}$$
$$= \sum_\nu z_\nu(\mathbf{r}_2-\mathbf{r}_2')z_{-\nu}(\mathbf{r}_1-\mathbf{r}_1')\exp\{i\mathbf{K}_\nu\cdot[\tfrac{1}{2}(\mathbf{r}_2+\mathbf{r}_2')-\tfrac{1}{2}(\mathbf{r}_1+\mathbf{r}_1')]\}. \tag{2.39}$$

In particular, for $\mathbf{r}_1'=\mathbf{r}_1$ and $\mathbf{r}_2'=\mathbf{r}_2$

$$\langle \mathbf{r}_1\mathbf{r}_2 \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle_{\mathrm{Av}} = \langle \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\rangle_{\mathrm{Av}} = \sum_\nu z_\nu(0)z_{-\nu}(0)\exp[i\mathbf{K}_\nu\cdot(\mathbf{r}_2-\mathbf{r}_1)]. \tag{2.40}$$

We see that, even when $|\mathbf{r}_2-\mathbf{r}_1|\to\infty$, there remains a periodic dependence on $\mathbf{r}_2-\mathbf{r}_1$, which represents the diagonal long-range order of the crystal. We also note that, unlike the case of the disordered system,

$$\lim_{|\mathbf{r}_2-\mathbf{r}_1|\to\infty}\langle \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\rangle_{\mathrm{Av}} \neq \langle \rho_1(\mathbf{r}_1)\rangle_{\mathrm{Av}}\langle \rho_1(\mathbf{r}_2)\rangle_{\mathrm{Av}}. \tag{2.41}$$

In the excitonic insulator we consider throughout the positions of the ions as fixed. There is however an additional broken symmetry due to the arbitrariness in the phase of the gap function. Under the conditions

[12] M. Rensink (to be published).

(2.30), ρ_2 of Eq. (2.26) reduces to

$$\langle \mathbf{r}_1'\mathbf{r}_2' \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle = \langle \mathbf{r}_2 \mid \rho_1 \mid \mathbf{r}_2'\rangle\langle \mathbf{r}_1 \mid \rho_1 \mid \mathbf{r}_1'\rangle \quad (2.42)$$

and we can see a new periodicity in ρ_1. From (2.28) and (2.29) we have for $\mathbf{r}_1=\mathbf{r}_1'$

$$\langle \rho_1(\mathbf{r})\rangle = \sum_\nu G_\nu \exp(i\mathbf{K}_\nu\cdot\mathbf{r}) + \sum_\nu (F_\nu \exp[i(\mathbf{K}_\nu+\mathbf{w})\cdot\mathbf{r}] + F_\nu^* \exp[-i(\mathbf{K}_\nu+\mathbf{w})\cdot\mathbf{r}]). \quad (2.43)$$

The first sum is of the same form as in an ordinary crystal [cf. Eq. (2.37)], but the second term exhibits a new periodicity characterized by the wave vector $\mathbf{w}$.

If one now averages over the phase of Δ (cf. the previous averaging over the center of mass $\mathbf{R}$), then, since v_k carries this phase we see from (2.29) that $F_\nu=0$. However, even with the phase averaging, a new diagonal long-range order remains in ρ_2. For example, setting $\mathbf{r}_2'=\mathbf{r}_2$ and $\mathbf{r}_1'=\mathbf{r}_1$, we find from (2.42), (2.28), and (2.29)

$$\langle \mathbf{r}_1\mathbf{r}_2 \mid \rho_2 \mid \mathbf{r}_1\mathbf{r}_2\rangle_{Av} = \langle \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\rangle_{Av} = [\sum_\nu G_\nu \exp(i\mathbf{K}_\nu\cdot\mathbf{r}_1)][\sum_{\nu'} G_{\nu'} \exp(i\mathbf{K}_{\nu'}\cdot\mathbf{r}_2)] + \sum_{\nu,\nu'}\{F_\nu F_{\nu'}^* \exp[i(\mathbf{K}_\nu\cdot\mathbf{r}_1-\mathbf{K}_{\nu'}\cdot\mathbf{r}_2)]\exp[i\mathbf{w}\cdot(\mathbf{r}_1-\mathbf{r}_2)]+\text{c.c.}\}. \quad (2.44)$$

In summary, then, we have a new periodicity characterized by $\mathbf{w}$, whose phase with respect to the lattice is arbitrary.

Throughout this paper we neglect the effects of electron spin. If we include the electron spin, then the pairing can take place in either a singlet or a triplet state. If the pairing is in a triplet state, there will be additional magnetic effects, which are currently under study.

III. GENERAL RESPONSE THEORY

We turn now to the general response of the excitonic insulator to a perturbation. We will employ a formalism similar to that used by Ambegaokar and Kadanoff[13] for superconductors. We introduce an anomalous Green's function to describe the excitonic insulator in direct analogy to the Gorkov F function used in the theory of superconductivity. It is more convenient to work in position space, and introducing the operators

$$\psi_a(\mathbf{x}) = \sum_k a_k \exp(i\mathbf{k}\cdot\mathbf{x}),$$
$$\psi_b(\mathbf{x}) = \sum_k b_k \exp(i\mathbf{k}\cdot\mathbf{x}), \quad (3.1)$$

we can write the model Hamiltonian [Eq. (2.6)] in position space as

$$H = \sum_{i=a,b}\int \psi_i^\dagger(\mathbf{x})\epsilon_i(\mathbf{x})\psi_i(\mathbf{x})\,d\mathbf{x} + \frac{1}{2}\int \rho(\mathbf{x})\rho(\mathbf{x}')V(\mathbf{x}-\mathbf{x}')\,d\mathbf{x}\,d\mathbf{x}'. \quad (3.2)$$

$\epsilon_i(\mathbf{x})$ and $\rho(\mathbf{x})$ are defined as follows:

$$\epsilon_a(\mathbf{x}) = -\tfrac{1}{2}G + (2m_a)^{-1}\nabla^2; \qquad \epsilon_b(\mathbf{x}) = \tfrac{1}{2}G - (2m_b)^{-1}\nabla^2;$$
$$\rho(\mathbf{x}) = \psi_a^\dagger(\mathbf{x})\psi_a(\mathbf{x}) + \psi_b^\dagger(\mathbf{x})\psi_b(\mathbf{x}). \quad (3.3)$$

[13] V. Ambegaokar and L. P. Kadanoff, Nuovo Cimento **22**, 914 (1961).

The equation of motion for the Heisenberg operator $\tilde{\psi}_b(1)$ is

$$[i(\partial/\partial t_1) - \epsilon_b(1)]\tilde{\psi}_b(1) = V(1-\bar{1})\rho(\bar{1})\tilde{\psi}_b(1), \quad (3.4)$$

where $1\equiv(\mathbf{x}_1, t_1)$, and the bar denotes integration over that variable. The two-body interaction is written as

$$V(1-1') = V(\mathbf{x}_1-\mathbf{x}_1')\delta(t_1-t_1'). \quad (3.5)$$

The zero-temperature Green's functions for the a and b bands are defined as

$$G_a(1,1') = -i\langle T\tilde{\psi}_a(1)\tilde{\psi}_a^\dagger(1')\rangle,$$
$$G_b(1,1') = -i\langle T\tilde{\psi}_b(1)\tilde{\psi}_b^\dagger(1')\rangle. \quad (3.6)$$

Multiplying Eq. (3.4) across by $\tilde{\psi}_b^\dagger(1')$, we get the equation of motion for the Green's function

$$[i(\partial/\partial t_1) - \epsilon_b(1)]G_b(1,1') = \delta(1,1') - iV(1-\bar{1})\langle T\rho(\bar{1})\tilde{\psi}_b(1)\tilde{\psi}_b^\dagger(1')\rangle. \quad (3.7)$$

Introducing the anomalous Green's functions

$$F^\dagger(1,1') = -i\langle T\tilde{\psi}_a(1)\tilde{\psi}_b^\dagger(1')\rangle,$$
$$F(1,1') = -i\langle T\tilde{\psi}_b(1)\tilde{\psi}_a^\dagger(1')\rangle, \quad (3.8)$$

we factorize the product on the right-hand side of (3.7) and we obtain

$$[i(\partial/\partial t_1) - \epsilon_b(1)]G_b(1,1') - iV(1-\bar{1})F^\dagger(\bar{1},1')F(1,\bar{1}) = \delta(1,1'). \quad (3.9)$$

We omit the Hartree and Hartree-Fock terms and use the real (i.e., observed) masses in $\epsilon_i(1)$. In a similar way we obtain an equation for $F^\dagger$

$$[i(\partial/\partial t_1) - \epsilon_a(1)]F^\dagger(1,1') - iV(1-\bar{1})F^\dagger(1,\bar{1})G_b(\bar{1},1') = 0. \quad (3.10)$$

If we take the Fourier transforms of (3.9) and (3.10)

then we find

$$[\epsilon-\epsilon_b(\mathbf{p})]G_b(\mathbf{p},\epsilon)-\Delta(\mathbf{p})F^\dagger(\mathbf{p},\epsilon)=1,$$
$$[\epsilon-\epsilon_a(\mathbf{p})]F^\dagger(\mathbf{p},\epsilon)-\Delta^\dagger(\mathbf{p})G_b(\mathbf{p},\epsilon)=0, \quad (3.11)$$

where

$$G(1,1')=\sum_{\mathbf{p}}\int\frac{d\epsilon}{2\pi}G(\mathbf{p},\epsilon) \times \exp i\cdot\{\mathbf{p}(\mathbf{x}_1-\mathbf{x}_1')-\epsilon(t_1-t_1')\} \quad (3.12)$$

and we have introduced the gap function $\Delta(\mathbf{p})$ defined by

$$\Delta(\mathbf{p})\equiv i\sum_{\mathbf{k}}\int\frac{d\nu}{2\pi}V(\mathbf{p}-\mathbf{k})F(\mathbf{k},\nu),$$
$$\Delta^\dagger(\mathbf{p})\equiv i\sum_{\mathbf{k}}\int\frac{d\nu}{2\pi}V(\mathbf{p}-\mathbf{k})F^\dagger(\mathbf{k},\nu). \quad (3.13)$$

Solving Eq. (3.11) for G_b and $F^\dagger$ in terms of Δ, we obtain

$$G_b(\mathbf{p},\epsilon)=\frac{\epsilon-\epsilon_a(\mathbf{p})}{[\epsilon-\epsilon_a(\mathbf{p})][\epsilon-\epsilon_b(\mathbf{p})]-|\Delta(\mathbf{p})|^2},$$
$$F^\dagger(\mathbf{p},\epsilon)=\frac{\Delta^\dagger(\mathbf{p})}{[\epsilon-\epsilon_a(\mathbf{p})][\epsilon-\epsilon_b(\mathbf{p})]-|\Delta(\mathbf{p})|^2}. \quad (3.14)$$

Substituting into Eq. (3.13) and integrating we get the complex conjugate of the gap equation, Eq. (2.18). In a similar way we find for G_a and F

$$G_a(\mathbf{p},\epsilon)=\frac{\epsilon-\epsilon_b(\mathbf{p})}{[\epsilon-\epsilon_a(\mathbf{p})][\epsilon-\epsilon_b(\mathbf{p})]-|\Delta(\mathbf{p})|^2},$$
$$F(\mathbf{p},\epsilon)=\frac{\Delta(\mathbf{p})}{[\epsilon-\epsilon_a(\mathbf{p})][\epsilon-\epsilon_b(\mathbf{p})]-|\Delta(\mathbf{p})|^2}. \quad (3.15)$$

Consider now the effect of an external perturbation U. The set of Eqs. (3.9) and (3.10) are still valid in the presence of the perturbation and the formally self-consistent field. Thus Eqs. (3.9) and (3.10) become

$$[i(\partial/\partial t_1)\delta(1,\bar1)-h_b(1,\bar1;U)]G_b(\bar1,1';U)-iV(1-\bar1)F^\dagger(\bar1,1';U)F(1,\bar1;U)=\delta(1,1'), \quad (3.16)$$
$$[i(\partial/\partial t_1)\delta(1,\bar1)-h_a(1,\bar1;U)]F^\dagger(\bar1,1';U)-iV(1-\bar1)F^\dagger(1,\bar1;U)G_b(\bar1,1';U)=0, \quad (3.17)$$

with

$$h_i(1,1';U)=\epsilon_i(1)\delta(1,1')-U(1,1')-v(1-\bar1)\langle\rho(\bar1;U)\rangle\delta(1,1'). \quad (3.18)$$

The last term in Eq. (3.18) represents the Hartree field due to the perturbation. Note that $v(1-\bar1)$ is the unscreened potential. Let us now expand Eqs. (3.16) and (3.17) and keep the linear term in U. We could do this by taking the functional derivative with respect to U evaluated at $U=0$. However, it is more convenient to rewrite Eqs. (3.16) and (3.17) in a slightly different way before we take the derivative.

$$G_{0b}^{-1}(1,\bar1;U)G_b(\bar1,1';U)-iV(1-\bar1)F(1,\bar1;U)F^\dagger(\bar1,1';U)=\delta(1,1'), \quad (3.19)$$
$$G_{0a}^{-1}(1,\bar1;U)F^\dagger(\bar1,1';U)-iV(1-\bar1)F^\dagger(1,\bar1;U)G_b(\bar1,1';U)=0, \quad (3.20)$$

with the following definition for the operator $G_{0i}^{-1}(1,1';U)$:

$$G_{0i}^{-1}(1,1';U)=[i(\partial/\partial t_1)-\epsilon_i(1)]\delta(1,1')-U(1,1')-v(1-\bar1)\langle\rho(\bar1;U)\rangle\delta(1,1'), \quad (3.21)$$

where $i=a, b$. Now, we multiply Eq. (3.20) across by $G_{0a}(2,1)$ and we integrate over the variable 1. We get, after an obvious change of variable,

$$F^\dagger(1,1';U)=iV(\bar1-\bar2)F^\dagger(\bar2,\bar1;U)G_b(\bar1,1';U)G_{0a}(1,\bar2;U). \quad (3.22)$$

Using the symmetry between a and b bands we get a similar equation for F;

$$F(1,1';U)=iV(\bar1-\bar2)F(\bar2,\bar1;U)G_a(\bar1,1';U)G_{0b}(1,\bar2;U). \quad (3.23)$$

After substitution for $F^\dagger$ from Eq. (3.22), (3.19) becomes

$$G_{0b}^{-1}(1,\bar1;U)G_b(\bar1,1';U)+V(1-\bar1)V(\bar3-\bar2)F(1,\bar1;U)F^\dagger(\bar3,\bar2;U)G_b(\bar2,1';U)G_{0a}(\bar1,\bar3;U)=\delta(1,1'). \quad (3.24)$$

We multiply Eq. (3.24) across by $G_b^{-1}(1',4;U)$ and integrate over the variable $1'$ and obtain after a change of variable

$$G_b^{-1}(1,1',U)=G_{0b}^{-1}(1,1',U)+V(1-\bar1)V(\bar2-1')F(1,\bar1,U)F^\dagger(\bar2,1',U)G_{0a}(\bar1,\bar2;U). \quad (3.25)$$

Differentiating with respect to U we get

$$-G_b^{-1}(3,\bar1)\frac{\partial G_b(\bar1,\bar2)}{\partial U(2,2')}G_b^{-1}(\bar2,3')=\frac{\partial h_b(3,3')}{\partial U(2,2')}-V(3-\bar1)V(\bar2-3')F(3,\bar1)F^\dagger(\bar2,3')G_{0a}(\bar1,\bar3)\frac{\partial h_a(\bar3,\bar4)}{\partial U(2,2')}G_{0a}(\bar4,\bar2)$$
$$+V(3-\bar1)V(\bar2-3')\frac{\partial F(3,\bar1)}{\partial U(2,2')}F^\dagger(\bar2,3')G_{0a}(\bar1,\bar2)+V(3-\bar1)V(\bar2-3')F(3,\bar1)\frac{\partial F^\dagger(\bar2,3')}{\partial U(2,2')}G_{0a}(\bar1,\bar2). \quad (3.26)$$

We multiply across by the product $G_b(13)G_b(3'1')$ and integrate over the variables 3 and 3'. Using Eqs. (3.22)

and (3.23), we find after some algebra

$$\frac{\partial G_b(1,1')}{\partial U(2,2')} = -G_b(1,\bar{1})G_b(\bar{2},1')\frac{\partial h_b(\bar{1},\bar{2})}{\partial U(2,2')} - F(1,\bar{1})F^\dagger(\bar{2},1')\frac{\partial h_a(\bar{1},\bar{2})}{\partial U(2,2')}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F(\bar{1},\bar{2})}{\partial U(2,2')}G_b(1,\bar{1})F^\dagger(\bar{2},1') + iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{1},\bar{2})}{\partial U(2,2')}G_b(\bar{2},1')F(1,\bar{1}). \quad (3.27)$$

If we differentiate (3.22) with respect to U we find

$$\frac{\partial F^\dagger(1,1')}{\partial U(2,2')} = -G_{0a}(1,\bar{1})F^\dagger(\bar{2},1')\frac{\partial h_a(\bar{1},\bar{2})}{\partial U(2,2')}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{2},\bar{1})}{\partial U(2,2')}G_b(\bar{1},1')G_{0a}(1,\bar{2}) + iV(\bar{1}-\bar{2})F^\dagger(\bar{2},\bar{1})\frac{\partial G_b(\bar{1},1')}{\partial U(2,2')}G_{0a}(1,\bar{2}), \quad (3.28)$$

and after substituting for $\partial G_b/\partial U$ from (3.27) we get

$$\frac{\partial F^\dagger(1,1')}{\partial U(2,2')} = -G_a(1,\bar{1})F^\dagger(\bar{2},1')\frac{\partial h_a(\bar{1},\bar{2})}{\partial U(2,2')} - G_b(\bar{2},1')F(1,\bar{1})\frac{\partial h_b(\bar{1},\bar{2})}{\partial U(2,2')}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{1},\bar{2})}{\partial U(2,2')}G_b(\bar{2},1')G_a(1,\bar{1}) + iV(\bar{1}-\bar{2})\frac{\partial F(\bar{1},\bar{2})}{\partial U(2,2')}F^\dagger(1,\bar{1})F^\dagger(\bar{2},1'). \quad (3.29)$$

We can use the symmetry between the a and b bands to write the functional derivatives

$$\partial G_a(1,1')/\partial U(2,2') \quad \text{and} \quad \partial F(1,1')/\partial U(2,2'):$$

$$\frac{\partial G_a(1,1')}{\partial U(2,2')} = -G_a(1,\bar{1})G_a(\bar{2},1')\frac{\partial h_a(\bar{1},\bar{2})}{\partial U(2,2')} - F^\dagger(1,\bar{1})F^\dagger(\bar{2},1')\frac{\partial h_b(\bar{1},\bar{2})}{\partial U(2,2')}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{1},\bar{2})}{\partial U(2,2')}G_a(1,\bar{1})F(\bar{2},1') + iV(\bar{1}-\bar{2})\frac{\partial F(\bar{1},\bar{2})}{\partial U(2,2')}G_a(\bar{2},1')F^\dagger(1,\bar{1}), \quad (3.30)$$

$$\frac{\partial F(1,1')}{\partial U(2,2')} = -G_b(1,\bar{1})F(\bar{2},1')\frac{\partial h_b(\bar{1},\bar{2})}{\partial U(2,2')} - G_a(\bar{2},1')F(1,\bar{1})\frac{\partial h_a(\bar{1},\bar{2})}{\partial U(2,2')}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F(\bar{1},\bar{2})}{\partial U(2,2')}G_a(\bar{2},1')G_b(1,\bar{1}) + iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{1},\bar{2})}{\partial U(2,2')}F(\bar{2},1')F(1,\bar{1}). \quad (3.31)$$

The set of Eqs. (3.27), (3.29)–(3.31), determine the response of a system to a general perturbation and we shall take these as our starting point, in the following sections, to calculate the collective modes and the electromagnetic response function.

IV. THE COLLECTIVE MODES

We can use the equations derived in the last section to calculate the energy spectrum of the collective modes. If the external perturbation U has a wave vector and frequency equal to that of a collective mode to which it is coupled, the response will be infinite, like that of a forced oscillator at its resonant frequency. The dispersion relation of the collective modes can therefore be obtained from the condition that the response equations (3.28)–(3.31) have a solution even without an external driving term.

Let us take U as a scalar potential, i.e., $U(2,2') = U(2)\delta(2,2')$. Then we have from the definition (3.18)

$$\partial h(1,1')/\partial U(2) = -\delta(1,2)\delta(2,1') - v(1-\bar{1})[\partial\langle\rho(\bar{1})\rangle/\partial U(2)]\delta(1,1'). \quad (4.1)$$

Substituting this into (3.29) and (3.31), but omitting the first (driving) term on the right-hand side of (4.1) we obtain a set of integral for $\partial F^\dagger/\partial U$ and $\partial F/\partial U$ involving also $\partial\langle\rho\rangle/\partial U$. But

$$\langle\rho(1)\rangle = -i[G_a(1,1^+) + G_b(1,1^+)] \qquad [1^+ = (\mathbf{x}_1, t_1+0)], \quad (4.2)$$

and using (3.27) and (3.30) we obtain

$$\frac{\partial[G_a(1,1^+)+G_b(1,1^+)]}{\partial U(2)}=iV(\bar{1}-\bar{2})\frac{\partial F(\bar{1},\bar{2})}{\partial U(2)}\{F^\dagger(1,\bar{1})G_a(\bar{2},1^+)+G_b(1,\bar{1})F^\dagger(\bar{2},1^+)\}$$

$$+iV(\bar{1}-\bar{2})\frac{\partial F^\dagger(\bar{1},\bar{2})}{\partial U(2)}\{G_a(1,\bar{1})F(\bar{2},1^+)+F(1,\bar{1})G_b(\bar{2},1^+)\}-iv(\bar{1}-\bar{2})\frac{\partial[G_a(\bar{2},\bar{2}^+)+G_b(\bar{2},\bar{2}^+)]}{\partial U(2)}$$

$$\times\{\sum_{i=a,b}G_i(1,\bar{1})G_i(\bar{1},1^+)+F(1,\bar{1})F^\dagger(\bar{1},1^+)+F^\dagger(1,\bar{1})F(\bar{1},1^+)\}. \tag{4.3}$$

We now have a set of coupled integral equations linking the quantities $\partial F^\dagger/\partial U$, $\partial F/\partial U$, and $\partial(G_a+G_b)/\partial U$. Let us transform to momentum space. If we define

$$\frac{\partial F(1,1')}{\partial U(2)}=\sum_{k,q}\int\frac{d\nu d\omega}{4\pi^2}t_k(q)\ \exp i\{\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-\omega(t_1-t_2)+\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_1')-\nu(t_1-t_1')\},$$

$$\frac{\partial F^\dagger(1,1')}{\partial U(2)}=\sum_{k,q}\int\frac{d\nu d\omega}{4\pi^2}t_k^*(q)\ \exp i\{\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-\omega(t_1-t_2)+\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_1')-\nu(t_1-t_1')\},$$

$$\frac{\partial[G_a(1,1^+)+G_b(1,1^+)]}{\partial U(2)}=\sum_q\int\frac{d\omega}{2\pi}r(q)\ \exp i\{\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-\omega(t_1-t_2)\}, \tag{4.4}$$

using the four-vectors $k=(\mathbf{k},\nu)$, $q=(\mathbf{q},\omega)$, $p=(\mathbf{p},\epsilon)$, then

$$t_k(q)=i\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})F(k)F(k+q)t_p^*(q)$$

$$+i\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})G_a(k)G_b(k+q)t_p(q)-iv(\mathbf{q})r(q)\{F(k+q)G_a(k)+G_b(k+q)F(k)\}, \tag{4.5}$$

$$t_k^*(q)=i\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})F^\dagger(k)F^\dagger(k+q)t_p(q)$$

$$+i\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})G_b(k)G_a(k+q)t_p^*(q)-iv(\mathbf{q})r(q)\{F^\dagger(k+q)G_b(k)+G_a(k+q)F^\dagger(k)\}, \tag{4.6}$$

$$r(q)=i\sum_{k,p}\int\frac{d\epsilon d\nu}{4\pi^2}V(\mathbf{p}-\mathbf{k})t_p(q)\{F^\dagger(k+q)G_a(k)+G_b(k+q)F^\dagger(k)\}$$

$$+i\sum_{k,p}\int\frac{d\epsilon d\nu}{4\pi^2}V(\mathbf{p}-\mathbf{k})t_p^*(q)\{G_a(k+q)F(k)+F(k+q)G_b(k)\}$$

$$-i\sum_p\int\frac{d\epsilon}{2\pi}v(\mathbf{q})r(q)\{\sum_{i=a,b}G_i(p+q)G_i(p)+F(p+q)F^\dagger(p)+F^\dagger(p+q)F(p)\}. \tag{4.7}$$

We can express these equations in terms of the variables

$$A_k(q)=\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})[t_p(q)-t_p^*(q)],$$

$$B_k(q)=\sum_p\int\frac{d\epsilon}{2\pi}V(\mathbf{p}-\mathbf{k})[t_p(q)+t_p^*(q)],$$

and

$$L(q)=[v(\mathbf{q})]^{1/2}r(q).$$

Then using the following form for G and F

$$G_b(p)=u_p^2/(\epsilon-\zeta_p-E_p+i\delta)+v_p^2/(\epsilon-\zeta_p+E_p-i\delta),$$
$$F(p)=F^\dagger(p)=u_pv_p\{(\epsilon-\zeta_p-E_p+i\delta)^{-1}-(\epsilon-\zeta_p+E_p-i\delta)^{-1}\},$$
$$G_a(p)=v_p^2/(\epsilon-\zeta_p-E_p+i\delta)+u_p^2/(\epsilon-\zeta_p+E_p-i\delta), \tag{4.8}$$

where $\zeta_p=\frac{1}{2}[\epsilon_a(\mathbf{p})+\epsilon_b(\mathbf{p})]$, we arrive at a set of coupled integral equations

$$A_k(q)=-\sum_p \frac{V(\mathbf{k}-\mathbf{p})l_p^2(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}A_p(q)-\sum_p \frac{V(\mathbf{k}-\mathbf{p})l_p(\mathbf{q})n_p(\mathbf{q})(\omega+\zeta-\zeta')}{(\omega+\zeta-\zeta')^2-(E+E')^2}B_p(q)$$
$$+\sum_p \frac{2V(\mathbf{k}-\mathbf{p})[v(\mathbf{q})]^{1/2}l_p(\mathbf{q})p_p(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}L(q), \quad (4.9)$$

$$B_k(q)=-\sum_p \frac{V(\mathbf{k}-\mathbf{p})l_p(\mathbf{q})n_p(\mathbf{q})(\omega+\zeta-\zeta')}{(\omega+\zeta-\zeta')^2-(E+E')^2}A_p(q)-\sum_p \frac{V(\mathbf{k}-\mathbf{p})n_p^2(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}B_p(q)$$
$$+\sum_p \frac{2V(\mathbf{k}-\mathbf{p})[v(\mathbf{q})]^{1/2}p_p(\mathbf{q})n_p(\mathbf{q})(\omega+\zeta-\zeta')}{(\omega+\zeta-\zeta')^2-(E+E')^2}L(q), \quad (4.10)$$

$$L(q)=\sum_p \frac{2v(\mathbf{q})p_p^2(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}L(q)-\sum_p \frac{[v(\mathbf{q})]^{1/2}n_p(\mathbf{q})p_p(\mathbf{q})(\omega+\zeta-\zeta')}{(\omega+\zeta-\zeta')^2-(E+E')^2}B_p(q)$$
$$-\sum_p \frac{[v(\mathbf{q})]^{1/2}l_p(\mathbf{q})p_p(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}A_p(q). \quad (4.11)$$

l, m, n, and p are the coherence factors given by

$$l_p(\mathbf{q})=u_pu_{p+q}+v_pv_{p+q}, \qquad m_p(\mathbf{q})=u_pv_{p+q}+v_pu_{p+q},$$
$$n_p(\mathbf{q})=u_pu_{p+q}-v_pv_{p+q}, \qquad p_p(\mathbf{q})=u_pv_{p+q}-v_pu_{p+q}, \quad (4.12)$$

and we have used the notation that $E=E_p$, $E'=E_{p+q}$, $\zeta=\zeta_p$, and $\zeta'=\zeta_{p+q}$.

We now assign $\mathbf{q}$ a definite value and look for the values of ω which are eigenvalues of these equations. At $\mathbf{q}=0$; $\omega=0$ is an eigenvalue and the corresponding eigenfunctions are $A_k(0)=\Delta_{|k|}$; $B_k(0)=L(0)=0$. A careful examination at finite but small $\mathbf{q}$ shows that there is a solution with a phonon-like dispersion relation, i.e., $\omega\sim|\mathbf{q}|$. However it is very difficult to solve exactly for the velocity except in one limiting case. In the extreme semimetallic limit the interaction potential V becomes more and more short range so that one may eventually replace $V(\mathbf{k})$ by a constant, V. Note that the direct interaction $v(\mathbf{q})$ is not screened. With this simplification the coupled set of integral equations reduces to a set of linear equations for the quantities $L(q)$, $A(q)$, and $B(q)$. In this region we may also make use of particle-hole symmetry, i.e., any quantity which is summed over $\mathbf{p}$ and which is odd in $\epsilon_b(\mathbf{p})$ or $\epsilon_a(\mathbf{p})$ can be neglected. The products

$$p_p(\mathbf{q})l_p(\mathbf{q})=\Delta(\xi-\xi')/2EE' \quad (4.13)$$

and

$$l_p(\mathbf{q})n_p(\mathbf{q})=\xi/2E+\xi'/2E' \quad (4.14)$$

are odd in ξ and thus the terms which couple $L(q)$ and $A(q)$ are zero. The terms which couple $A(q)$ and $B(q)$ are each proportional to q^2 as ω, $\mathbf{q}\to 0$. Thus the $A(q)$ mode is essentially uncoupled in this limit, and its dispersion relation is given by

$$1=-\sum_p \frac{Vl_p^2(\mathbf{q})(E+E')}{(\omega+\zeta-\zeta')^2-(E+E')^2}. \quad (4.15)$$

Expanding in powers of ω, $|\mathbf{q}|$ and using the gap equation, we get

$$0=V\sum_p\left\{\frac{(\omega+\zeta-\zeta')^2}{(2E)^3}-\frac{(\xi-\xi')^2}{(2E)^3}\right\}, \quad (4.16)$$

or

$$\omega^2=k_F^2q^2/3m_am_b. \quad (4.17)$$

We note that this mode has the same dispersion relation as the well-known "ionic sound" mode in a two-component plasma. This acoustic mode is associated with the broken symmetry in the new phase and thus will not disappear when interband scattering terms are included, contrary to the remark in Kozlov and Maksimov.[7]

V. THE ELECTROMAGNETIC RESPONSE AT $T=0$

In Sec. III we discussed the general response of the excitonic insulator to a perturbation. In this section we will use those results to derive the linear electromagnetic response function at zero temperature. We will show that the excitonic insulator is in fact an insulator and that there is no Meissner effect. Thus despite the close formal similarity of the excitonic insulating state and the superconducting state, the physical properties of the two states are very different. This is what we expect in view of the very different nature of the electronic order in the two states.

We study the effect of an external electromagnetic

field by introducing a vector potential $\mathfrak{A}(\mathbf{x},t)$ into the Hamiltonian. In the presence of the field the single-particle terms in the Hamiltonian are

$$h_i(1,1')=\epsilon_i[\hat{\mathbf{p}}_1-e\mathfrak{A}(1)]\delta(1,1') -v(1-\bar{1})\langle\rho(\bar{1};\mathfrak{A})\rangle\delta(1,1'), \quad i=a,b. \tag{5.1}$$

The current generated by the external field can be split into two terms, a diamagnetic term which is the usual gauge term and a paramagnetic term:

$$\mathbf{J}(1)=\mathbf{J}^d(1)+\mathbf{J}^p(1), \tag{5.2}$$

where

$$\mathbf{J}^d(1)=-(n_c e^2/\mu)\mathfrak{A}(1)$$

and n_c is the number of carriers in either band,

$$n_c=\sum_{\mathbf{p}}\langle a_{\mathbf{p}}a_{\mathbf{p}}^\dagger\rangle=\sum_{\mathbf{p}}\langle b_{\mathbf{p}}^\dagger b_{\mathbf{p}}\rangle=\sum_{\mathbf{p}}v_{\mathbf{p}}^2. \tag{5.3}$$

The paramagnetic term is given by

$$\mathbf{J}^p(1)=-\tfrac{1}{2}e(\nabla_1-\nabla_{1'})\left\{m_b^{-1}\frac{\delta G_b(1,1')}{\delta\mathfrak{A}(\bar{2})}-m_a^{-1}\frac{\delta G_a(1,1')}{\delta\mathfrak{A}(\bar{2})}\right\}_{1'=1}\cdot\mathfrak{A}(\bar{2}). \tag{5.4}$$

Taking the functional derivative of Eq. (5.1) we find

$$\delta h_b(1,1')/\delta\mathfrak{A}(2)=-(e/2m_b i)\{\nabla_2-\nabla_{2'}\}\delta(1-2)\delta(1'-2')\,|_{2'=2}-v(1-\bar{1})[\delta\langle\rho(\bar{1})\rangle/\delta\mathfrak{A}(2)]\delta(1,1'). \tag{5.5}$$

Substituting in (3.27) and (3.29) and taking the Fourier transform we obtain

$$\begin{aligned}\Lambda_p{}^b(q)=&-(e/2m_b)(2\mathbf{p}+\mathbf{q})G_b(p+q)G_b(p)+(e/2m_a)(2\mathbf{p}+\mathbf{q})F(p+q)F^\dagger(p)\\&+i\sum_{\mathbf{k}}\int\frac{dv}{2\pi}V(\mathbf{k})[\lambda_{p-k}(q)F^\dagger(p)G_b(p+q)+\lambda_{p-k}{}^*(q)F(p+q)G_b(p)]\\&-i\sum_{\mathbf{k}}\int\frac{dv}{2\pi}v(\mathbf{q})[G_b(p+q)G_b(p)+F(p+q)F^\dagger(p)]\{\Lambda_k{}^a(q)+\Lambda_k{}^b(q)\},\end{aligned} \tag{5.6}$$

and

$$\begin{aligned}\lambda_p{}^*(q)=&+(e/2m_a)(2\mathbf{p}+\mathbf{q})G_a(p+q)F^\dagger(p)-(e/2m_b)(2\mathbf{p}+\mathbf{q})G_b(p)F(p+q)\\&+i\sum_{\mathbf{k}}\int\frac{dv}{2\pi}V(\mathbf{k})[\lambda_{p-k}(q)F^\dagger(p+q)F^\dagger(p)+\lambda_{p-k}{}^*(q)G_a(p+q)G_b(p)]\\&-i\sum_{\mathbf{k}}\int\frac{dv}{2\pi}v(\mathbf{q})[F^\dagger(p+q)G_b(p)+F^\dagger(p)G_a(p+q)][\Lambda_k{}^a(q)+\Lambda_k{}^b(q)],\end{aligned} \tag{5.7}$$

where $\Lambda_p{}^i$ and $\lambda_p{}^*$ are defined by

$$\frac{\partial G_i(1,1')}{\delta\mathfrak{A}(2)}=\sum_{\mathbf{p},\mathbf{q}}\int\frac{d\epsilon d\omega}{4\pi^2}\Lambda_p{}^i(q)\exp\{i\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-i\omega(t_1-t_2)+i\mathbf{p}\cdot(\mathbf{x}_1-\mathbf{x}_1')-i\epsilon(t_1-t_1')\} \tag{5.8}$$

and

$$\frac{\partial F^\dagger(1,1')}{\delta\mathfrak{A}(2)}=\sum_{\mathbf{p},\mathbf{q}}\int\frac{d\epsilon d\omega}{4\pi^2}\lambda_p{}^*(q)\exp\{i\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-i\omega(t_1-t_2)+i\mathbf{p}\cdot(\mathbf{x}_1-\mathbf{x}_1')-i\epsilon(t_1-t_1')\}. \tag{5.9}$$

The Fourier transform of the current can be expressed in terms of the Λ and we find using (5.4)

$$\mathbf{J}^p(q)=-ie\sum_{\mathbf{p}}\int\frac{d\epsilon}{4\pi}(2\mathbf{p}+\mathbf{q})\left\{\frac{\Lambda_p{}^b(q)}{m_b}-\frac{\Lambda_p{}^a(q)}{m_a}\right\}\cdot\mathfrak{A}(q). \tag{5.10}$$

The calculation of the current requires the solution of the integral Eq. (5.7) for λ^* and the companion equation for λ. Let us look first at a static external field and for simplicity restrict ourselves to the equal mass case. With these simplifications the equations can be written as

$$\mathbf{J}^p(\mathbf{q})=(e/2m)\sum_{\mathbf{p}}(2\mathbf{p}+\mathbf{q})[m_{\mathbf{p}}(\mathbf{q})/(E_{\mathbf{p}}+E_{\mathbf{p}+\mathbf{q}})]\{n_{\mathbf{p}}(\mathbf{q})\beta_{\mathbf{p}}(\mathbf{q})+(e/m)(2\mathbf{p}+\mathbf{q})\cdot\mathfrak{A}(\mathbf{q})m_{\mathbf{p}}(\mathbf{q})\}, \tag{5.11}$$

where

$$\beta_{\mathbf{p}}(\mathbf{q})=i\sum_{\mathbf{k}}\int\frac{dv}{2\pi}V(\mathbf{p}-\mathbf{k})[\lambda_{\mathbf{k}}{}^*(\mathbf{q})+\lambda_{\mathbf{k}}(\mathbf{q})]\cdot\mathfrak{A}(\mathbf{q}), \tag{5.12}$$

and $\beta_p(\mathbf{q})$ is determined by an inhomogeneous linear integral equation

$$\beta_{p'}(\mathbf{q})=\sum_p V(\mathbf{p}-\mathbf{p}')[n_p(\mathbf{q})/(E_p+E_{p+q})]\{n_p(\mathbf{q})\beta_p(\mathbf{q})+(e/m)(2\mathbf{p}+\mathbf{q})\cdot\mathfrak{A}(\mathbf{q})m_p(\mathbf{q})\}. \tag{5.13}$$

The functions $m_p(\mathbf{q})$, $n_p(\mathbf{q})$ are the coherence factors introduced earlier in Eqs. (4.12).

Let us first verify that our equations are gauge-invariant. This we do by taking $\mathfrak{A}(\mathbf{q})$ as a longitudinal vector potential and demonstrating that the system does not respond. Then we shall investigate the Meissner effect by examining the response to a general potential in the limit $\mathbf{q}\to 0$. If we write $\mathfrak{A}(\mathbf{q})=\mathbf{q}\mathfrak{A}_0(\mathbf{q})$, the Eq. (5.12) for $\beta_p(\mathbf{q})$ may be rewritten as

$$\beta_{p'}(\mathbf{q})=\sum_p V(\mathbf{p}-\mathbf{p}')[n_p(\mathbf{q})/(E_p+E_{p+q})]\{n_p(\mathbf{q})\beta_p(\mathbf{q})+2em_p(\mathbf{q})\mathfrak{A}_0(\mathbf{q})[\xi(\mathbf{p}+\mathbf{q})-\xi(\mathbf{p})]\}. \tag{5.14}$$

The solution of this equation is

$$\beta_p(\mathbf{q})=2e(\Delta_p-\Delta_{p+q})\mathfrak{A}_0(\mathbf{q}) \tag{5.15}$$

as we shall verify by direct substitution. The right-hand side becomes

$$\beta_{p'}(\mathbf{q})=2e\mathfrak{A}_0\sum_p V(\mathbf{p}-\mathbf{p}')[n_p(\mathbf{q})/(E_p+E_{p+q})]\{n_p(\mathbf{q})(\Delta_p-\Delta_{p+q})+m_p(\mathbf{q})[\xi(\mathbf{p}+\mathbf{q})-\xi(\mathbf{p})]\} \tag{5.16}$$

and, after some algebra, we find that this can be rewritten as

$$\beta_{p'}(\mathbf{q})=-2e\mathfrak{A}_0\sum_p V(\mathbf{p}-\mathbf{p}')n_p(\mathbf{q})p_p(\mathbf{q}) \tag{5.17}$$

$$=-2e\mathfrak{A}_0\sum_p V(\mathbf{p}-\mathbf{p}')\frac{1}{2}\left(\frac{\Delta_{p+q}}{E_{p+q}}-\frac{\Delta_p}{E_p}\right) \tag{5.18}$$

$$=2e\mathfrak{A}_0(\Delta_{p'}-\Delta_{p'+q}). \tag{5.19}$$

This verifies that Eq. (5.15) is a solution of Eq. (5.14) for a longitudinal vector potential. It remains to substitute this solution into Eq. (5.11) and then to evaluate the current

$$\mathbf{J}^p(\mathbf{q})=\frac{e^2\mathfrak{A}_0}{m}\sum_p(2\mathbf{p}+\mathbf{q})\frac{m_p(\mathbf{q})}{E_p+E_{p+q}}\times\{n_p(\mathbf{q})(\Delta_p-\Delta_{p+q})+[\xi(\mathbf{p}+\mathbf{q})-\xi(\mathbf{p})]m_p(\mathbf{q})\} \tag{5.20}$$

$$=(-2e^2/m)\mathfrak{A}_0\sum_p(\mathbf{p}+\tfrac{1}{2}\mathbf{q})m_p(\mathbf{q})p_p(\mathbf{q}) \tag{5.21}$$

$$=(-2e^2/m)\mathfrak{A}_0\sum_p(\mathbf{p}+\tfrac{1}{2}\mathbf{q})(u_p^2v_{p+q}^2-v_p^2u_{p+q}^2) \tag{5.22}$$

$$=(-2e^2/m)\mathfrak{A}_0\sum_p(\mathbf{p}+\tfrac{1}{2}\mathbf{q})(v_{p+q}^2-v_p^2) \tag{5.23}$$

$$=(2e^2/m)\mathfrak{A}_0\mathbf{q}\sum_p v_p^2 \tag{5.24}$$

$$=(n_ce^2/\mu)\mathfrak{A}(\mathbf{q}). \tag{5.25}$$

Combining the paramagnetic and gauge contributions to the current, we see at once that they cancel exactly and the total current is zero. This demonstrates that the theory is gauge-invariant.

Now let us examine the integral equation, Eq. (5.13) in the limit as $\mathbf{q}\to 0$ for a general vector potential. The integral equation is perfectly regular in this limit. By contrast, in a superconductor the corresponding integral equation [see Rickayzen,[14] Eq. (6.3)] is singular in this limit. It is this singularity, which leads to the very different behavior of the response function of a superconductor to a transverse or longitudinal vector potential. In the excitonic insulator the variable $\beta_p(q)$ is independent of the polarization as $\mathbf{q}\to 0$. Thus the response function is independent of the polarization in the limit as $\mathbf{q}\to 0$ and there is no Meissner effect.

Lastly, let us examine the response in the low-frequency limit. If we apply a low-frequency electric field to the system, we represent the perturbation by a vector potential

$$\mathfrak{A}(\mathbf{q},\omega)=-(i/\omega)\mathbf{E}(\mathbf{q},\omega). \tag{5.26}$$

Then the electromagnetic kernel K_{ij} defined by

$$J_i(\mathbf{q},\omega)=K_{ij}(\mathbf{q},\omega)\mathfrak{A}_j(\mathbf{q},\omega), \tag{5.27}$$

and the conductivity σ given by

$$\mathbf{J}(\mathbf{q},\omega)=\sigma(\mathbf{q},\omega)\mathbf{E}(\mathbf{q},\omega), \tag{5.28}$$

are related as follows:

$$\sigma(\mathbf{q},\omega)=-i\,\mathrm{Re}\{(1/\omega)K_{ii}(\mathbf{q},\omega)\}. \tag{5.29}$$

The dc conductivity is determined by the response to a very low-frequency uniform electric field, i.e., $\lim_{\omega\to 0}(\lim_{q\to 0}K_{ii}(\mathbf{q},\omega))$. Now the electromagnetic kernel is determined by (5.6), (5.7), and (5.10). At $T=0$ an examination of the integrals over the products of Green's function entering (5.6) and (5.7) shows that they are regular in the limit $\mathbf{q},\omega\to 0$. Thus $\lim_{q,\omega\to 0}K(\mathbf{q},\omega;T=0)$ is well defined. (For nonzero temperatures this is not so and the limit depends critically on the ratio $|\mathbf{q}|/\omega$ as $\mathbf{q},\omega\to 0$.) We have shown above that

$$K_{ij}(\mathbf{q}=0;\omega=0;T=0)=0,$$

[14] G. Rickayzen, Phys. Rev. **115**, 795 (1959).

thus the dc conductivity at zero temperature is zero and the new phase is an insulating phase.

VI. THE CONDUCTIVITY OF THE EXCITONIC INSULATOR

We turn now to a discussion of the finite-temperature resistivity of the excitonic insulator. We are interested in the resistivity, since, in general, a transition between metallic and insulating behavior is accompanied by a large change in the resistivity. Thus we expect that the transition will be most easily seen experimentally in the resistivity.[15] In the last section the electromagnetic kernel K was calculated at zero temperature. We can generalize the results to finite temperature by using the finite-temperature time-dependent Green's functions. These are defined as

$$\mathcal{G}_i(1,1') = -i\langle\langle T\tilde{\psi}_i(1)\tilde{\psi}_i^\dagger(1')\rangle\rangle, \qquad i=a,b$$
$$\mathcal{F}^\dagger(1,1') = -i\langle\langle T\tilde{\psi}_a(1)\tilde{\psi}_b^\dagger(1')\rangle\rangle, \tag{6.1}$$

where $\langle\langle\cdots\rangle\rangle$ denotes the thermal average of the time-ordered product. The form of the Green's-function equations is unchanged at finite temperature. However, these determine only the real part of $\mathcal{G}$. The imaginary part can be obtained by using the relation due to Landau[16]

$$\mathrm{Re}\mathcal{G}(\epsilon) = -\pi^{-1}\int_{-\infty}^{+\infty} \coth\tfrac{1}{2}(\beta x)\,\frac{\mathrm{Im}\mathcal{G}(x)}{\epsilon - x}\,dx, \tag{6.2}$$

and we find, in the equal-mass case,

$$\mathcal{G}_b(\mathbf{p},\epsilon) = \frac{u_p^2(1-n_p)}{\epsilon-E_p+i\delta} + \frac{v_p^2 n_p}{\epsilon+E_p+i\delta} + \frac{v_p^2(1-n_p)}{\epsilon+E_p-i\delta} + \frac{u_p^2 n_p}{\epsilon-E_p-i\delta},$$

$$\mathcal{G}_a(\mathbf{p},\epsilon) = \frac{v_p^2(1-n_p)}{\epsilon-E_p+i\delta} + \frac{u_p^2 n_p}{\epsilon+E_p+i\delta} + \frac{u_p^2(1-n_p)}{\epsilon+E_p-i\delta} + \frac{v_p^2 n_p}{\epsilon-E_p-i\delta},$$

$$\mathcal{F}(\mathbf{p},\epsilon) = \mathcal{F}^\dagger(\mathbf{p},\epsilon) = u_p v_p\left\{\frac{1-n_p}{\epsilon-E_p+i\delta} - \frac{n_p}{\epsilon+E_p+i\delta} - \frac{1-n_p}{\epsilon+E_p-i\delta} + \frac{n_p}{\epsilon-E_p-i\delta}\right\}, \tag{6.3}$$

where $n_p=[\exp(\beta E_p)+1]^{-1}$ and we have chosen the phase of the gap function to be real. The gap function Δ_p is determined by the usual temperature-dependent equation

$$\Delta_p = \sum_k V(\mathbf{p}-\mathbf{k})\,(\Delta_k/2E_k)\,\tanh(\tfrac{1}{2}\beta E_k). \tag{6.4}$$

The calculation of the finite-temperature electromagnetic kernel K proceeds in an analogous fashion to the zero-temperature calculation carried out in Sec. V. The finite-temperature generalization of Eqs. (5.6), (5.7), and (5.10) are found by replacing G and F by $\mathcal{G}$ and $\mathcal{F}$. Thus we find after some algebra

$$\lim_{\omega\to0}[\lim_{q\to0} K_{ij}{}^p(\mathbf{q},\omega)] = \frac{e^2}{m^2}\sum_p p_i p_j \frac{4u_p^2 v_p^2}{E_p}(1-2n_p) + \frac{e}{m}\sum_p p_i\gamma_j(\mathbf{p})\,\frac{u_p v_p}{E_p}(u_p^2-v_p^2)(1-2n_p), \tag{6.5}$$

where

$$\gamma_j(\mathbf{p}) = +i\sum_k \int \frac{d\nu}{2\pi} V(\mathbf{p}-\mathbf{k})[\lambda_k^*(0)+\lambda_k(0)]_j \tag{6.6}$$

and satisfies the equation

$$\gamma_j(\mathbf{p}) = \sum_k V(\mathbf{p}-\mathbf{k})\,\frac{(u_k^2-v_k^2)^2}{2E_k}(1-2n_k)\gamma_j(\mathbf{k}) + \frac{e}{m}\sum_k V(\mathbf{p}-\mathbf{k})\,2k_j\,\frac{u_k v_k}{E_k}(u_k^2-v_k^2)(1-2n_k). \tag{6.7}$$

Note the order of the limits $\mathbf{q},\omega\to0$ is important at finite temperatures. The dc conductivity is determined by the response to a very low-frequency uniform electric field. Thus $\lim\mathbf{q}\to0$ must be taken before $\lim\omega\to0$.

The calculation of the conductivity in general involves the solution of the inhomogeneous integral Eq. (6.7). However in the most interesting limiting case the solution is trivial. In the semimetal limit the inhomogeneous term in (6.7) is zero by particle-hole symmetry. Thus $\gamma(p)=0$ is the solution and one finds in this limit

$$\lim_{\omega\to0}[\lim_{q\to0} K_{ij}{}^p(\mathbf{q},\omega)] = \frac{e^2}{m^2}\sum_p p_i p_j \frac{\Delta^2}{E_p{}^3}(1-2n_p) \tag{6.8}$$

$$= \frac{e^2}{3m^2}\sum_p \frac{p^2\Delta^2}{E_p{}^3}(1-2n_p)\delta_{i,j}. \tag{6.9}$$

[15] It has recently been suggested by Kozlov and Maksimov [A. N. Kozlov and L. A. Maksimov, Zh. Eksperim. i Teor. Fiz. **50**, 131 (1966) [English transl.: Soviet Phys.—JETP **23**, 88 (1966)]], that an excitonic insulator would be a "superthermal conductor." They consider the properties of the states of the excitonic insulator which are the analogues of the current-carrying superconducting BCS states. In the excitonic insulator such states do not involve mass or charge transport. Moreover in common with supercurrents in metals and He^4, a "supercurrent" in an excitonic insulator will not carry entropy. Thus it will not contribute to heat transport or lead to "superthermal conductivity."

[16] L. D. Landau, Zh. Eksperim. i Teor. Fiz. **34**, 262 (1958) [English transl.: Soviet Phys.—JETP **7**, 182 (1958)].

The diamagnetic contribution to the kernel is

$$K_{ij}^{d}(0,0) = -\frac{e^2}{m}\sum_{\mathrm{p}}\langle\langle a_{\mathrm{p}}a_{\mathrm{p}}^{\dagger}+b_{\mathrm{p}}^{\dagger}b_{\mathrm{p}}\rangle\rangle\delta_{i,j} \tag{6.10}$$

$$= -\frac{e^2}{m}\sum_{\mathrm{p}}[2v_{\mathrm{p}}^2+(u_{\mathrm{p}}^2-v_{\mathrm{p}}^2)n_{\mathrm{p}}]\delta_{i,j}. \tag{6.11}$$

In the semimetal limit, the second term on the right vanishes by particle-hole symmetry and combining (6.9) and (6.11) we find for low frequencies

$$\sigma(0,\omega) = -(i/\omega)\ \mathrm{Re}\{\lim_{\omega\to 0}(\lim_{q\to 0}[K_{ii}^{p}(\mathbf{q},\omega)+K_{ii}^{d}(\mathbf{q},\omega)])\} \tag{6.12}$$

$$= -\frac{ie^2}{m\omega}\sum_{\mathrm{p}}\left\{\frac{p^2\Delta^2}{3mE_{\mathrm{p}}^3}(1-2n_{\mathrm{p}}) - \left(1-\frac{\xi_{\mathrm{p}}}{E_{\mathrm{p}}}\right)\right\}. \tag{6.13}$$

The temperature-independent term is identically zero, as can be seen by integrating by parts, and we are left with

$$\sigma(0,\omega) = +i(n^{\mathrm{eff}}e^2/m\omega), \tag{6.14}$$

where

$$n^{\mathrm{eff}} = \tfrac{2}{3}\sum_{\mathrm{p}}(p^2\Delta^2/mE_{\mathrm{p}}^3)[\exp(\beta E_{\mathrm{p}})+1]^{-1} \tag{6.15}$$

$$= 2n_c\int_0^{\infty} d\xi\,\frac{\Delta^2}{(\xi^2+\Delta^2)^{3/2}}\{\exp[\beta(\xi^2+\Delta^2)^{1/2}]+1\}^{-1} \tag{6.16}$$

$$= 2n_c\int_0^{\infty}\frac{dt}{(t^2+1)^{3/2}}\{\exp[\beta\Delta(t^2+1)^{1/2}]+1\}^{-1}; \tag{6.17}$$

n_c is the number of carrier in a band defined in (5.3). In the semimetallic limit $n_c = k_F^3/3\pi^2$.

The foregoing discussion has referred to a "perfect" system without scattering, for which we found that for $\omega\to 0$, the conductivity has the form

$$\sigma(\omega) = \sigma_1(\omega)+i\sigma_2(\omega)$$
$$\sigma_2(\omega) = c/\omega \qquad \omega>0. \tag{6.18}$$

Such a form, implies by the Kramers-Kronig relationship

$$\sigma_2(\omega) = -\frac{2\omega}{\pi}\int_0^{\infty}\frac{\sigma_1(\omega')}{\omega'^2-\omega^2}\,d\omega', \tag{6.19}$$

that

$$\sigma_1(\omega) = 0, \qquad \omega>0 \tag{6.20a}$$

$$\int_0^{\infty}\sigma_1(\omega)\,d\omega = \tfrac{1}{2}\pi c. \tag{6.20b}$$

Thus $\sigma_1(\omega)$ has a δ-function-like character.

When scattering is present but weak, $\sigma_1(\omega)$ is somewhat spread out (up to $\omega\sim 1/\tau$, where τ is the effective relaxation time) but (6.20b) continues to hold. If the functional form of $\sigma_1(\omega)$ is adequately given by a simple relaxation time expression

$$\sigma_1(\omega) = \sigma_1(0)\,1/[1+(\omega\tau)^2]. \tag{6.21}$$

Then, by (6.22b) the dc conductivity $\sigma_1(0)$ is given in terms of our calculated constant c and the relaxation time τ by

$$\sigma_1(0) = c\tau. \tag{6.22}$$

Let us examine the qualitative behavior of the dc conductivity as a function of the temperature T and the bandgap G. Consider first $\sigma(T)$ for a fixed-value (negative) of the bandgap in the semimetal region. At the transition temperature the conductivity is a continuous function of temperature but there is a discontinuity in slope.

$$\frac{d\sigma}{dT} = \frac{\tau e^2}{m}\left\{\frac{\partial n^{\mathrm{eff}}}{\partial T}+\frac{\partial n^{\mathrm{eff}}}{\partial\Delta}\frac{d\Delta}{dT}\right\}. \tag{6.23}$$

As $T\to T_c$ the second term on the right diverges and $d\sigma/dT\to+\infty$. As $T\to 0$ then $d\sigma/dT\to 0$. The qualitative behavior of $\sigma(T)$ is sketched in Fig. 2 in the semimetal limit.

Also of physical interest is the behavior of σ at a fixed temperature as the gap G is changed, for example by varying the external pressure. In the normal (semimetal) state the dc conductivity is proportional to the number of carriers and $\sigma\propto(-G)^{3/2}$. At the transition $\sigma(G)$ is continuous, but there is a discontinuity in slope.

$$\frac{d\sigma}{dG} = \frac{\tau e^2}{m}\left\{\frac{\partial n^{\mathrm{eff}}}{\partial G}+\frac{\partial n^{\mathrm{eff}}}{\partial\Delta}\frac{d\Delta}{dG}\right\}. \tag{6.24}$$

The second term on the right-hand side diverges as $G\to G_c$ and $d\sigma/dG = -\infty$ at $G=G_c$. In Fig. 3 the rough behavior of $\sigma(G)$ is shown.

VII. CHOICE OF MATERIALS

In the preceding sections we have discussed some of the main physical features associated with the excitonic

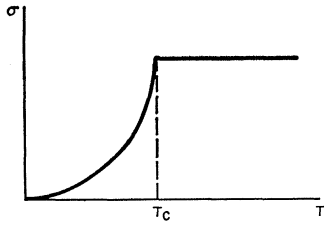

FIG. 2. The dc conductivity as a function of temperature in the semimetallic limit.

insulator phase. No experimental evidence for the existence of this phase has yet been reported, so we now turn to questions connected with its realizability in the laboratory.

Kozlov and Maksimov[6] have calculated the main features of the phase diagram which is shown in Fig. 1. The scale of this diagram is determined by the exciton-binding energy in the normal insulating phase,

$$|E_B| = \tfrac{1}{2}(\mu^*/m)(1/\epsilon_0^2)\ \text{Ry}, \tag{7.1}$$

where μ^* and m are, respectively, the reduced effective mass and free-electron mass, and ϵ_0 is the static dielectric constant. The cutoff for positive G is given by

$$G_1 = |E_B|; \tag{7.2}$$

for negative G there is no sharp cutoff, but the transition temperature falls exponentially on a scale given by $|E_B|$, so that we may take the effective cutoff at

$$G_2 = -|E_B|; \tag{7.3}$$

the maximum transition temperature, according to Ref. 6, is given approximately by

$$k\bar{T}_c \approx |E_B|. \tag{7.4}$$

We now turn to the question of the experimental realizability of the excitonic phase. Since this phase occurs when the bandgap is approximately zero, one is led to consider materials which under normal conditions have a small (positive or negative) gap. Further, one requires that the gap can be varied continuously through zero by some external means without the occurrence, in the "normal" phase, of a polymorphic transition. These considerations suggest two main groups of materials: Divalent metals and Group V semimetals and related alloys.

Among possible means of varying the gap, the most useful appeared to be the application of hydrostatic pressure. Several measurements of resistance and crystal structure as function of pressure and temperature do in fact indicate the existence of some materials which may satisfy the necessary requirements. They are discussed in more detail, in the subsequent two sections.

1. Divalent Metals

Pressure measurements on these metals have recently been summarized in a review article by Drickamer.[17] Calcium, strontium, and ytterbium, under application of pressure, all appear to change from metallic to insulating behavior as shown by the temperature coefficient of their resistance.

Strontium

This material appears to us the most promising of all. Being a divalent element, strontium owes its normal metallic character to band overlap. Recent high-pressure measurements[18] have shown that at both 77 and 300°K this material maintains its fcc structure up to a pressure of at least 37 kbar.[19] The following remarks are restricted to pressures below this value. Resistance measurements at both temperatures indicate that as pressure is increased the electronic properties tend towards those of an insulator: (1) At both 77 and 300°K, the resistance increases strongly as the pressure is increased, and (2) at pressures around 37 kbar, the resistance at 77°K is higher than at 300°K, and at 300°K a slightly negative temperature coefficient has been reported.

By themselves these experimental results would strongly suggest that under pressure, Sr develops a positive gap G. However, band-theoretic considerations lead to the conclusion that the two bands in question cannot actually separate.[20] If one accepts these conclusions one is led to the belief that, under high pressure, Sr becomes a semimetal with a degeneracy temperature, T_F, of less than about 300°K, perhaps much lower. Resistance measurements under high pressure and in the temperature range from 4.2 to 77°K would be most helpful in clarifying the situation.

We can make only the roughest estimate of $|E_B|$ since neither μ^* nor ϵ_0 is known for Sr. By analogy with Ca (see below), we may guess $\mu^* \approx 0.25\ m$ and $\epsilon_0 \approx 30$, which gives

$$|E_B| \approx 4\times10^{-3}\ \text{eV}, \tag{7.5}$$

and hence

$$\bar{T}_c \approx 0.45\,|E_B|/k \approx 20°\text{K}. \tag{7.6}$$

This is only an order-of-magnitude estimate. If Sr

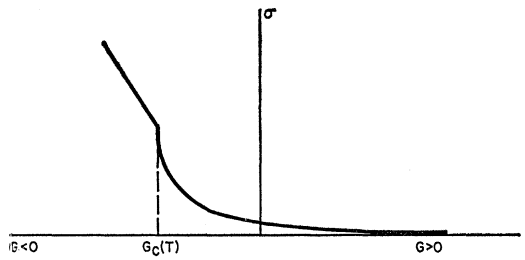

FIG. 3. The dc conductivity as a function of bandgap G at low temperatures.

[17] H. G. Drickhamer, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1965), Vol. 17.

[18] R. A. Stager and H. G. Drickhamer, Phys. Rev. **131**, 2524 (1963).

[19] A. Jayaraman, W. Klement, and G. C. Kennedy, Phys. Rev. **132**, 1620 (1963).

[20] B. Vasvari, A. O. E. Animalu, and V. Heine, Phys. Rev. **154**, 535 (1967); B. Vasvari and V. Heine (to be published).

remains a semimetal, a condition for the new phase to be observable would be that T_F is not much larger than $\bar{T}_c$ (see Fig. 1). There appears to be some possibility that this is in fact the case.

The pressure range ΔP over which the new phase might be observed is given by

$$(\partial G/\partial P)\,\Delta P \approx |\,E_B\,|. \tag{7.7}$$

If we estimate, very roughly, $(\partial G/\partial P)\approx(0.5\text{ eV})/25$ kbar we obtain

$$\Delta P \approx 0.2\text{ kbar}. \tag{7.8}$$

Of course this is again only a very rough estimate but it points to the need for accurate pressure control, quite likely not yet within the reach of present low-temperature–high-pressure techniques.

Calcium

This is qualitatively similar to Sr, except that the interesting region occurs at much higher pressure.[18] There is a first-order phase transition at about 350 kbar, below which Ca has the fcc structure. Near 300 kbar, Ca has a negative temperature coefficient of the resistance, suggesting a positive or at least-small negative gap.

Cyclotron resonance gives values of about $m^*=0.5\,m$ for electrons and holes, leading to $\mu^*\approx 0.25\,m$.[21]

The static dielectric constant we have estimated as best we could by means of the Kramers-Kronig relations, using the available optical data for the energy range $2.16\text{ eV}\leq\hbar\omega\leq 3.08$ eV and theoretical considerations for both Al and Ca.[22–24] We conclude that

$$\epsilon_0\approx 30, \tag{7.9}$$

with an uncertainty of about a factor of 2.

Finally, we use the theoretical band overlap of 0.5 eV to estimate

$$\partial G/\partial P = (0.5\text{ eV}/300\text{ kbar}). \tag{7.10}$$

Combining these estimates we find

$$\bar{T}_c\approx 20°\text{K}, \qquad \Delta P\approx 2\text{ kbar}. \tag{7.11}$$

Ytterbium

This element has been extensively studied under pressure. It keeps its fcc structure up to 40 kbar. At about 18 kbar the temperature coefficient of the resistance changes sign indicating a metal to insulator transition.[25] Hall and Merrill[26] report a kink in the compressibility at this pressure, but no such effect appears in the data of Stephens.[27] A positive gap above 15 kbar has been quite conclusively established by the resistivity measurements of Souers and Jura.[25] We have no independent information about m^* and ϵ_0.

Thus the metal-insulator transition seems more firmly established for this material than for Sr and Ca. However, there are elements of uncertainty due to conflicting experimental results on the compressibility and to the possible role of the $4f$ electrons.

Note added in proof. The electrical resistivity of Sr and Yb under pressure has recently been studied by D. B. McWhan [Bull. Am. Phys. Soc. **12,** 356 (1967), and to be published] in the region 4.2 to 300°K. He finds that Yb in the fcc phase undergoes a change from a metallic to a semiconducting state with increasing pressure. In Sr there is evidence for a small energy gap in the fcc phase just before the crystal phase change to bcc. The samples, however, are not pure enough to be intrinsic semiconductors and the excitonic insulator is masked by impurity effects.

The optical properties of Yb and Sr have been studied by W. E. Müller [Phys. Letters **17,** 82 (1965), and to be published]. Due to the presence of low-energy interband transitions it is not possible to separate clearly the interband and free-carrier contributions, so that one cannot make a reliable estimate of the dielectric constant. We estimate $\epsilon_0\approx 50$ at 1 atm in Yb, though it could be much less, which would give $T_c\approx$ 13°K for Yb. Sr appears to be very similar to Yb in its optical properties.

2. Semimetals

The elements As, Sb, and Bi at first sight appear very promising. A continuous transition from metallic to insulating behavior under pressure has been clearly established for Bi by Jaggi.[28] At 4.2°K, the gap changes sign at about 6 kbar. A polymorphic transition does not occur until about 25 kbar. Similar behavior may be expected for Sb and As.[29]

However these materials do, nevertheless, not appear to be good candidates for the observation of the new phase, because the small effective mass and high-dielectric constant lead to a very small exciton-binding energy. A rough estimate using $\mu^*=0.01\,m$ and $\epsilon_0=100$,[30] gives

$$|\,E_B\,| = 1.3\times 10^{-5}\text{ eV}, \tag{7.12}$$

$$\bar{T}_c = 0.05°\text{K}. \tag{7.13}$$

Also the total number of electrons and holes in the new phase would be $\approx 10^{12}$ cm^{-3}, so that impurities would entirely mask any intrinsic effects. Finally, the pressure range ΔP over which the new phase would exist is only about

$$\Delta P\approx 2\times 10^{-3}\text{ kbar}, \tag{7.14}$$

far less than present pressure control permits.

[21] J. H. Condon and J. A. Marcus, Phys. Rev. **134,** A446 (1964).
[22] H. M. O'Bryan, J. Opt. Soc. Am. **26,** 122 (1963).
[23] W. A. Harrison, Phys. Rev. **147,** 467 (1966).
[24] H. Ehrenreich, H. R. Philipp, and B. Segall, Phys. Rev. **132,** 1918 (1963).
[25] P. C. Souers and G. Jura, Science **140,** 481 (1963).
[26] H. T. Hall and L. Merrill, Inorg. Chem. **2,** 618 (1963).
[27] D. R. Stephens, J. Phys. Chem. Solids **25,** 423 (1964).
[28] R. Jaggi, in *Proceedings of the Conference on Semiconductors, Paris, 1964* (Dunod Cie, Paris, 1964).
[29] L. M. Falicov, in *Physics of Solids at High Pressures,* edited by C. T. Tomizuka and R. M. Emrick (Academic Press Inc., New York, 1965).
[30] W. S. Boyle and A. D. Brailsford, Phys. Rev. **120,** 1943 (1960).