Small-bipolaron formation

Morrel H. Cohen, E. N. Economou,* and Costas M. Soukoulis
Corporate Research Science Laboratories, Exxon Research and Engineering Company, Annandale, New Jersey 08801
(Received 4 October 1983)

Using the scaling arguments of Emin and Holstein, we examine the formation of small bipolarons by coupling of electrons to acoustic phonons. We find that for sufficiently small Coulomb interactions, the bipolaron forms at lower electron-phonon coupling constant λ than the polaron and is of lower energy than two polarons when these become stable. For larger Coulomb interactions, there can be a range of λ for which only the polaron is stable, and above that a range for which the bipolaron is the lower-energy form.

I. INTRODUCTION

Emin and Holstein[1] (EH) have introduced an elegantly simple method of studying polaron formation in the adiabatic limit. They have examined coupling of the electron, both to optical and acoustic modes, and have considered the effect of Coulombic impurity potential. The simplicity of their method derives from the simplicity of the scaling properties of the individual terms in the total energy of the electron-phonon system. We have extended their theory to the case of an impurity potential of finite range, showing that the existence of a characteristic length in the problem, the range, does not destroy the possibility of the scaling argument,[2] although it makes it more complicated. We have also found the EH method to provide a very powerful way to attack successfully, the far more difficult problem of the influence of the electron-phonon interaction on the mobility edge of a disordered semiconductor.[3] In the present note, we apply the EH method to the question of the bipolaron. Specifically, we ask whether two electrons will be bound together into a small bipolaron by coupling to acoustic phonons at some value of the electron-phonon coupling constant different, in general, from that required for small-polaron formation.

II. FORMALISM

The total energy of the two-electron—plus—phonon system can be written as

$$E_2=\langle H_{\rm el}\rangle+E_s\int d^3r_1\,d^3r_2\,|\psi(\vec{r}_1,\vec{r}_2)|^2 \times[\Delta(\vec{r}_1)+\Delta(\vec{r}_2)]+\tfrac{1}{2}K\int d^3r\,\Delta^2(r) \tag{1}$$

where $\psi(\vec{r}_1,\vec{r}_2)$ is the orbital part of the two-particle wave function and is symmetric with respect to the interchange $\vec{r}_1\leftrightarrow\vec{r}_2$. $H_{\rm el}$ is the sum of the band energies and the Coulomb repulsion between the electrons. $\Delta(\vec{r})$ is the dilation of the lattice at $\vec{r}$. The electron-phonon interaction has been assumed of the deformation-potential type: $H_{e\text{-ph}}(\vec{r})=E_s\Delta(\vec{r})$. K is a proper combination of elastic constants. Minimizing the total energy with respect to the dilation $\Delta(\vec{r})$, we find that $\Delta(\vec{r})=-2E_s\rho(\vec{r})/K$, where

$$\rho(\vec{r})=\int d^3r_1\,|\psi(\vec{r},\vec{r}_1)|^2\,. \tag{2}$$

Substituting back into Eq. (1), we obtain

$$E^{(2)}=\langle H_{\rm el}\rangle-4A\int d^3r\,\rho^2(\vec{r})\,, \tag{3}$$

where $A=E_s^2/2K$.

We now force the two electrons to be confined within a domain of linear dimension L. To find the L dependence of the various terms contributing to $E^{(2)}$, one must first examine the effects of the electron-electron correlation. For this purpose we have considered a trial wave function of the form

$$C\exp[-a(r_1^2+r_2^2)][1-\gamma\exp(-\nu a r_{12}^2)]\,,$$

where $a\sim 1/L^2$, C is the normalization factor, and the parameters γ,ν characterize the correlation. For each value of a we have found the values of γ and ν which minimize $\langle H_{\rm el}\rangle$. Our main results are the following. (i) The correlation is negligible for small L. (ii) The Coulomb interaction energy behaves as $\beta(L)/L$ with $\beta(L)$ a slowly varying function of L changing at most by a factor of $\frac{1}{2}$ as L changes from zero to infinity; $\beta(L)$ is about constant for $L\lesssim L_c$, where $L_c\approx 50$. Thus for the purposes of the present calculation, β can be taken as constant, equal to its value with no correlations. (iii) The kinetic energy behaves as $\alpha(L)/L^2$ with $\alpha(L)$ remaining about constant for $L\lesssim L_c$ and then steadily increasing, finally reaching (for $L\to\infty$) a value about 50 times larger than the $L=0$ (uncorrelated) value. However, at the values of L where $\alpha(L)$ is appreciably different from the $L=0$ value the kinetic energy is practically negligible relative to the Coulomb repulsion energy. Thus replacing $\alpha(L)$ by its uncorrelated value $\alpha(0)$ is a good approximation for the present purposes. In Fig. 1 we plot $\langle H_{\rm el}\rangle$ for the actual values of $\alpha(L)$ and $\beta(L)$ and for $\alpha(0)$ and $\beta(0)$.

We have also considered alternative ways of confining the electrons within a region of linear dimension L, e.g., by imposing rigid boundary conditions on the surface of a sphere or of a cube. In all cases we have defined L from the relation

$$\int d^3r\,\rho^2(L)=L^{-3}\,. \tag{4}$$

 ©1984 The American Physical Society

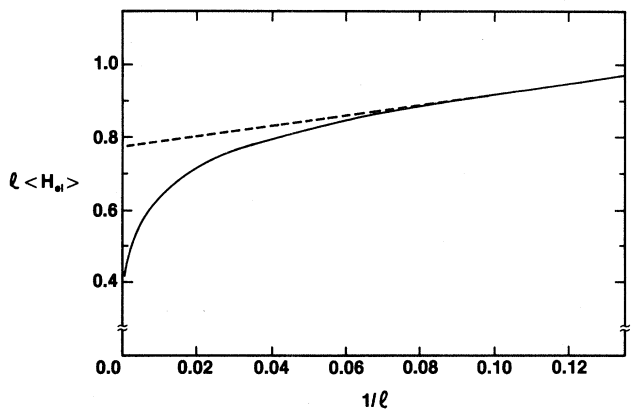

FIG. 1. Plot of the electronic energy of a correlated (solid line) and an uncorrelated (dashed line) Gaussian trial function (see text) vs $1/l$, where l is the dimensionless length given by Eq. (6).

This definition, which is similar to the way the participation ratio is defined in disordered systems,[4] is very convenient for our purposes in view of Eq. (3). We have found that the values of $\alpha(0)$ and $\beta(0)$ depend weakly on the choice of the particular way the confinement is implemented; the differences were no greater than 30%.

Defining

$$\epsilon^{(2)}=\frac{E^{(2)}}{E_H}\frac{\epsilon^2}{m^*}\ , \tag{5}$$

$$l=\frac{L}{a_B}\frac{m^*}{\epsilon}\ , \tag{6}$$

$$\lambda^*=\frac{E_s^2}{2KE_H a_B{}^3}\frac{m^{*2}}{\epsilon}\ , \tag{7}$$

$$\eta=\frac{\eta'}{a_B}\frac{m^*}{\epsilon}\ , \tag{8}$$

we can write the total energy as follows:

$$\epsilon^{(2)}=\frac{10}{l^2}+\frac{2}{l}-4\lambda^*\frac{1}{l^3+\eta^3}\ . \tag{9}$$

In the above formulas $a_B=\hbar^2/me^2=0.529$ Å is the Bohr radius, $E_H=e^4m/\hbar^2\simeq 27.2$ eV is the Hartree unit of energy, m^* is the ratio of the effective mass to the electronic mass, ϵ is the dielectric constant, and η' is a cutoff length roughly equal to the interatomic distance. In our numerical calculations we have taken $\eta'/a_B=5$. The presence of η in the last term of Eq. (9) ensures the obvious physical requirement that compressing the wave function below the interatomic distance does not increase further the lattice-mediated self-interaction, while it still increases the kinetic and Coulomb energies.

In our study of polaron formation in disordered media, we found it useful to introduce a phonon coupling λ defined by the relation

$$\lambda=\frac{E_s^2}{2K}\bar{\rho} \tag{10}$$

where $\bar{\rho}$ is the density of states per unit volume averaged over the band. Note that λ is very similar to the quantity denoted by the same symbol in the theory of superconductivity. The relation between λ^* and λ is

$$\lambda^*=\frac{1}{a_B^3E_H\bar{\rho}}\lambda\frac{m^{*2}}{\epsilon}\ . \tag{11}$$

In our subsequent calculations we have chosen $a_B^3E_H\bar{\rho}=0.1$, the typical value for Si.

The energy of two well-separated electrons interacting with the lattice is

$$2\epsilon^{(1)}=\frac{10}{l^2}-2\lambda^*\frac{1}{l^3+\eta^3}\ , \tag{12}$$

and

$$U=\epsilon^{(2)}_{\min}-2\epsilon^{(1)}_{\min}=\frac{2}{l}-2\lambda^*\frac{1}{l^3+\eta^3}\ . \tag{13}$$

III. RESULTS AND DISCUSSION

The important parameters in our expressions for $\epsilon^{(2)}(l)$ and $2\epsilon^{(1)}(l)$ are the dimensionless ratio m^*/ϵ, which determines the length scale, and the product λm^*, which to-

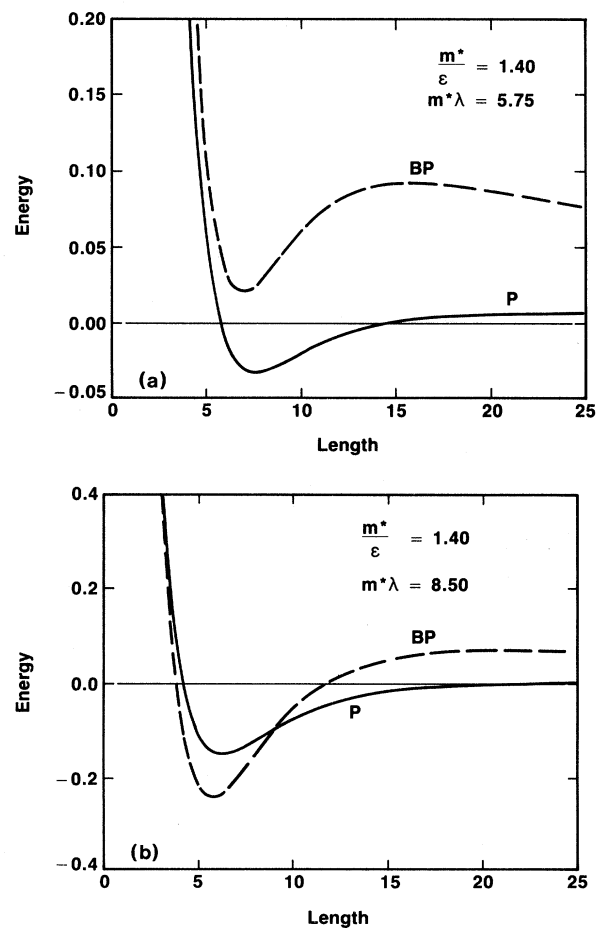

FIG. 2. Plot of the energy $\epsilon^{(2)}$ vs length l (dashed line) and energy $2\epsilon^{(1)}$ vs l (solid line) for two characteristic values of the parameters m^*/ϵ and $m^*\lambda$. The units of $\epsilon^{(2)}$ and l are given in Eqs. (5) and (6), respectively.

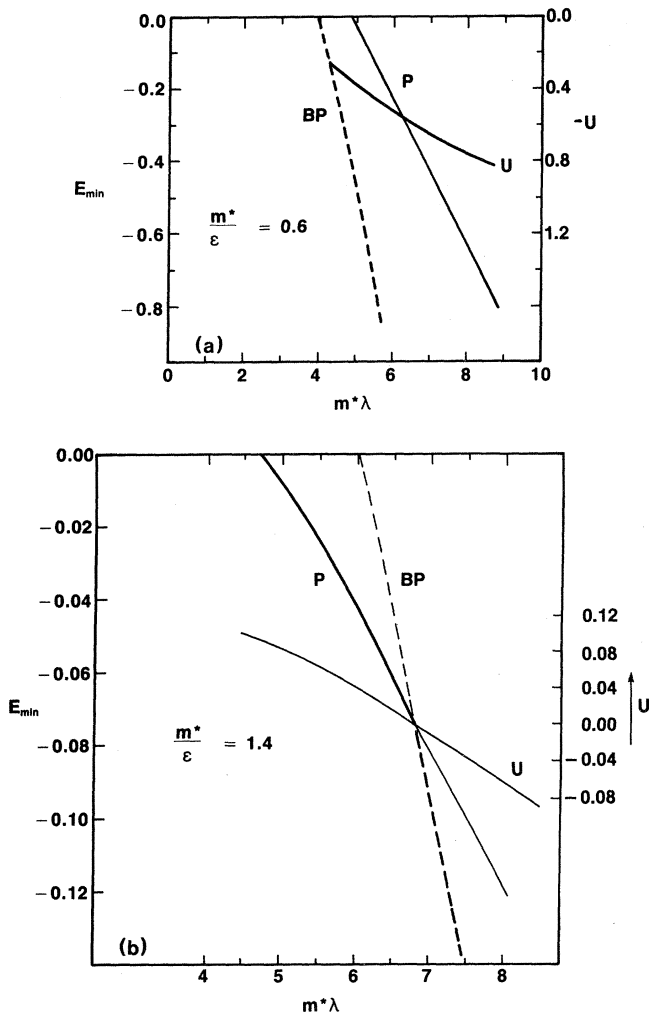

FIG. 3. Plot of the energy $\epsilon^{(2)}_{\min}$ (dashed line), $2\epsilon^{(1)}_{\min}$ (solid line), and $U=\epsilon^{(2)}_{\min}-2\epsilon^{(1)}_{\min}$ (solid line) vs λm^* for two different values of m^*/ϵ. m^* is the ratio of the effective mass to the electronic mass, and λ is the electron-phonon coupling constant.

gether with m^*/ϵ determines the relative strength of the lattice-mediated self- and mutual interactions.

In Figs. 2(a) and 2(b) we plot $\epsilon^{(2)}$ and $2\epsilon^{(1)}$ vs l. In Fig. 2(a), corresponding to $m^*/\epsilon=1.4$ and $\lambda m^*=5.75$, the $2\epsilon^{(1)}$ curve has a minimum at an l about equal to the interatomic distance. The minimum corresponds to an energy lower than zero (the value at $l=\infty$). Thus, two independent small polarons are formed. On the other hand, the $\epsilon^{(2)}$-vs-l curve has a minimum which does not cross the horizontal axis showing that no bipolaron is formed. The difference between the two minima can be interpreted as a renormalized U in the Hubbard sense: $U=\epsilon^{(2)}_{\min}-2\epsilon^{(1)}_{\min}$. In the case of Fig. 2(a), U is positive (repulsive) and about 2 eV (for $m^*\approx 1$). In Fig. 2(b) we see that both minima are lower than the energy of the extended ($l=\infty$) state. However, now the bipolaron is more stable than the polaron, or, equivalently, U is negative (attractive).

In Fig. 3(b) we plot the bipolaron and polaron minima as well as their difference U. As λm^* increases the ground state starts from extended ($l=\infty$), and then

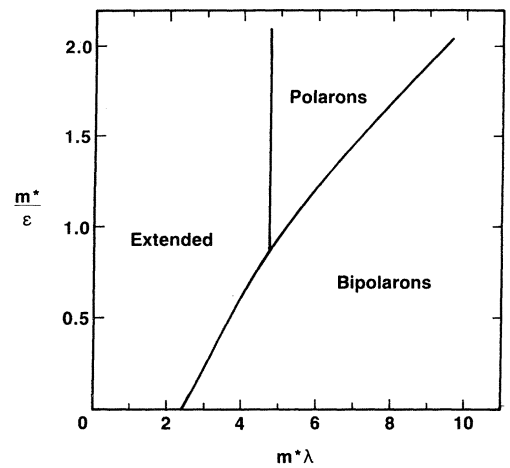

FIG. 4. Phase diagram in the plane $m^*/\epsilon,\lambda m^*$ indicating the regions of extended eigenstates ($l=\infty$), small polarons (positive U), and small bipolarons (negative U). m^* is the ratio of the effective mass to the electronic mass, ϵ is the dielectric constant, and λ is the electron-phonon coupling constant.

switches to small polarons with a positive (repulsive) U. Finally, at a still higher value of λm^*, U becomes negative (attractive) and, consequently, the ground state is a small bipolaron. This sequence of events is expected on physical grounds, since increasing λm^* implies an increasing self-attraction which first overcomes the kinetic energy [when m^* is large (i.e., small kinetic energy) and ϵ is small (i.e., large Coulomb repulsion)] to form polarons, and then the Coulomb repulsion as well, to form bipolarons. On the other hand, for small values of m^*/ϵ, i.e., large kinetic energy and/or small Coulomb repulsion, the increasing self-attraction is expected to first overcome the Coulomb repulsion and then the kinetic energy; this is what happens [see Fig. 3(a)]. The ground state switches directly from extended states to small bipolarons. U is negative for all values of λm^*.

In Fig. 4 we summarize our findings in a phase diagram. For small λ and/or small m^* (i.e., wide-band materials), no polarons or bipolarons are formed. For narrow-band materials (large m^*) with small ϵ, polaron formation is favored with U being repulsive. On the other hand, for large ϵ and/or large λ, U becomes negative (attractive), and the ground state switches to bipolaron. Within the framework of our approach the transition from extended to polarons or bipolarons is first order, while the transition from polaron to bipolaron is second order. Some of the these findings were already known.[5] However, it is worth pointing out how effortlessly they can be obtained by the scaling procedure.

Finally, we mention that the critical value of λ for small-polaron or -bipolaron formation is about 5 for $m^*\approx 1$. This estimate is no more accurate than a factor of 2. Because λ is almost identical to the λ entering the theory of superconductivity, it follows that materials of potentially high T_c ($\lambda \gtrsim 3$) are unstable towards bipolaron formation, thus becoming insulators instead of high-T_c superconductors. The best-known example where this instability actually occurs is the case of solid hydrogen, where the metallic phase transforms to the ultimate bipo-

laron: two electrons pulling together two hydrogen nuclei to form what is more commonly described as molecular hydrogen. However, there may be an intermediate regime in which bipolarons form but are not localized so that bipolaronic superconductivity occurs, as proposed by Alexandrov and Ranninger.[5]

---

*Permanent address: Department of Physics, University of Crete and Research Center of Crete, Heraklio, Crete, Greece.

[1]D. Emin and T. Holstein, Phys. Rev. Lett. 36, 323 (1976).

[2]M. H. Cohen and E. N. Economou (unpublished).

[3]M. H. Cohen, E. N. Economou, and C. M. Soukoulis, J. Non-Cryst. Solids 59-60, 15 (1983); Phys. Rev. Lett. 51, 1202 (1983).

[4]R. J. Bell and P. Dean, Discuss. Faraday Soc. 50, 55 (1970); see also J. M. Ziman, *Models of Disorder* (Cambridge University Press, Cambridge, 1979).

[5]P. W. Anderson, Phys. Rev. Lett. 34, 953 (1975); B. K. Chakraverty and C. Schlenker, J. Phys. (Paris) Colloq. 37, C4-353 (1976); S. Lakkis *et al.*, Phys. Rev. B 14, 1429 (1976); B. K. Chakraverty, M. J. Sienko, and J. Bonnerot, *ibid.* 17, 3781 (1978); B. K. Chakraverty, J. Phys. (Paris) Lett. 40, L99 (1979); C. Schlenker, *et al.* J. Phys. C 12, 3503 (1979); A. Alexandrov and J. Ranninger, Phys. Rev. B 23, 1796 (1981); 24, 1164 (1981); J. L. Brédas, R. R. Chance, and R. Silbey, Phys. Rev. B 26, 5843, (1982); 26, 6000 (1982); G. Crecelius *et al.* Phys. Rev. Lett. 50, 1498 (1983); J. L. Brédas, B. Thémans, and J. M. André, Phys. Rev. B 27, 7827 (1983).