ANNALS OF PHYSICS: 8, 325–342 (1959)

Studies of Polaron Motion

Part I. The Molecular-Crystal Model

T. HOLSTEIN

Westinghouse Research Laboratories,
Pittsburg, Pennsylvania

In this paper is described a model for polaron motion which incorporates, in simplified form, the principal physical features of the problem. The (crystalline) medium, within which a single excess electron (or hole) is contained, is taken to be a one-dimensional molecular crystal, consisting of diatomic molecular sites; each site possesses a single vibrational degree of freedom, represented by the deviation, x_n , of its internuclear separation from equilibrium. The motion of the electron in this medium is treated by a tight-binding approach, in which the wave function is represented as a superposition of local "molecular" functions, $\phi(\mathbf{r} - n\mathbf{a}, x_n)$. In line with the x_n dependence of the ϕ's, it is also assumed that the "local" electronic energy, E_n , (which, in the conventional tight-binding theory, has one and the same value for all sites) depends (linearly) on x_n . This dependence gives rise to electron-lattice interaction; alternatively, it may be regarded as removing the electronic translational degeneracy, characteristic of the undistorted crystal, and thereby providing the possibility for electron trapping.

On the basis of the above-described model, the zeroth order adiabatic treatment of the polaron problem is developed. For values of the parameters such that the linear dimension of the polaron is large compared to a lattice spacing ("large" polaron), an exact solution is obtained; the correspondence between it and Pekar's zeroth-order solution is established. The conditions under which the size of the polaron becomes comparable to a lattice spacing ("small" polaron) are discussed. Finally, by way of exhibiting the relationship of the molecular-crystal concept to the real situation, a description is given of an alternate molecular-crystal model which, in the case of the large polaron, is completely equivalent to the continuum-polarization model of conventional polaron theory.

This note presents a description of a one-dimensional molecular-crystal model which will form the basis for a number of studies of the polaron problem, to be reported subsequently.[1] It is felt that the model, although physically different

[1] As exemplified by the following paper (to be referred to hereafter as II), these studies will be focussed primarily on the case of the "small" polaron, whose linear dimensions are of the order of a lattice-spacing.

from the cases encountered in practice, corresponds conceptually to them in sufficient degree, so as to merit investigation.

Let us first recall the principal features of the real polaron problem. The physical situation is that of a slow electron (or hole) in an insulating crystal. The Hamiltonian of the system may be regarded as a sum of three terms, H_e, H_L, and H_{int}. H_e, the electronic component, consists of the kinetic energy of the electron and the effective one-electron periodic potential. H_L, the lattice component, is the sum of the kinetic energy of the lattice particles and the lattice potential energy; the latter is taken as a known (invariably quadratic) function of the particle displacements from their equilibrium positions. Finally, H_{int}, the electron-lattice interaction, is a function of the electron coordinate and the lattice displacements; the latter dependence is usually assumed to be linear.

In the standard theory of the motion of slow electrons in insulating crystals, H_{int} is treated as a small perturbation, giving rise to electronic transitions (scattering) in which lattice-vibration quanta are simultaneously absorbed or emitted. The basic problem of polaron theory is that H_{int} is often too large to be treated as a perturbation, but must be considered in zeroth order. Probably the most important physical concept of the theory is that of "self-trapping" of the electron. Qualitatively, this may be understood as follows.

Let us imagine that an electron is momentarily fixed at some point of the crystal. As a result of electron-lattice interaction, the surrounding lattice particles will be displaced to new equilibrium positions; according to general mechanical principles, the "induced" displacements will be such as to provide a potential "well" for the electron. If the well is sufficiently deep, the elecrton will occupy a bound state, unable to move unless accompanied by the well, that is to say, by the induced lattice-deformation. The unit consisting of the electron, together with its induced lattice deformation, is called the "polaron."[2]

With these preliminaries out of the way, the molecular-crystal model may now be described. In the absence of the electron, the system is taken to be a linear chain of N identical diatomic molecules, whose orientations and centers of gravity are fixed, but whose internuclear separations are allowed to vary. The "lattice-vibrations" thus consist of the vibrations of the individual internuclear separations. For the time being, the latter will be assumed, in the absence of the electron, to be uncoupled. If, in addition, the potential energy curve of an individual molecule be assumed parabolic, the lattice Hamiltonian reads

$$H_L = \frac{1}{2M}\sum_{n=1}^{N} p_n^2 + \frac{1}{2}\sum_{n=1}^{N} M\omega_0^2 x_n^2, \tag{1}$$

[2] The term owes its origin to the circumstance that the problem was first considered for the case of polar crystals, in which the interaction of the electron with the optical polarization modes is especially strong. However, in principle, the self-trapping phenomenon is not limited to polar crystals, but may apply to any medium in which H_{int} is sufficiently strong, relative to H_e and H_L.

where x_n is the deviation of a typical internuclear separation from its equilibrium value, $p_n \equiv (\hbar/i)\partial/\partial x_n$ is the associated canonical momentum, M the relative mass, and ω_0 the harmonic vibration frequency.[3]

The motion of a single electron (or hole) in the above-described one-dimensional molecular-crystal medium will here be formulated in terms of the tight-binding approximation. The details of this formulation are presented in Appendix I; the principal features are as follows. The state of the system is expressed as a linear superposition

$$\psi(\mathbf{r}, x_1, \cdots, x_n) = \sum a_n(x_1, \cdots, x_n)\phi(\mathbf{r} - n\mathbf{a}, x_n), \tag{2}$$

of "molecular" electron wave functions, $\phi(\mathbf{r} - n\mathbf{a}, x_n)$, each localized about a particular (nth) molecular site ($\mathbf{a}$ is a unit lattice vector) and depending upon the internuclear coordinate, x_n, of that site. The coefficients of the superposition, $a_n(x_1, x_2, \cdots, x_N)$, are each functions of all the internuclear coordinates, x_1, x_2, $\cdots$, x_N. The equations which they obey are obtained from the time-dependent Schrödinger equation of the system by a standard "projection" procedure. With approximations appropriate to the tight-binding case (smallness of overlap and nonorthogonality integrals) and to the large mass ratio of electrons and nuclei, they take the form (see I-16)

$$\left[i\hbar\frac{\partial}{\partial t} - \sum_m \left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x_m{}^2} + \frac{1}{2}M\omega_0{}^2 x_m{}^2\right) - E(x_n) - W(x_1, \cdots, x_n)\right] \times a_n(x_1, \cdots, x_n) = \sum_{(\pm)} J(x_n, x_{n\pm1})a_{n\pm1}(x_1, \cdots, x_n), \tag{3}$$

where the sum on the right-hand side goes over the two neighbors of a given nth site. The quantities $E(x_n)$, $W_n(x_1, \cdots, x_N)$, and $J(x_n, x_{n\pm1})$ are defined by appendix equations (I-3), (I-14), and (I-15), respectively. Physically, $E(x_n)$ is the electron energy of an isolated "molecular-ion," i.e., of the system of the electron and an isolated molecule; in conformance with standard concepts of molecular theory, it is to be regarded as a function of the internuclear displacement coordinate of the molecule. $W_n(x_1, \cdots, x_N)$ represents the perturbation on $E(x_n)$ due to the presence of the other molecules. Finally, the $J(x_n, x_{n\pm1})$ are overlap integrals, depending on the ϕ's of typical neighbor sites n, $n \pm 1$, and are hence functions of the internuclear coordinates x_n and $x_{n\pm1}$ of these sites.

Let us at this point note that, if the internuclear coordinates are all "frozen"

[3] In some studies (e.g., II), coupling terms of the form

$$H_c = \tfrac{1}{2} M\omega_1{}^2 \sum_{n=1}^{N} x_n\, x_{n+1},$$

will be introduced into (1). It may be remarked here that the neglect of coupling is a simplification of a type comparable to that often employed in treatments of polaron theory, in which the actual optical frequency spectrum of a polar crystal is replaced by a single frequency (Einstein model).

at one and the same value, say x, the system becomes equivalent to the conventional atomic tight-binding case, in which the localized states, $\phi(\mathbf{r} - n\mathbf{a}, x)$ form a degenerate set. Then, by virtue of the translational symmetry of the system, the stationary solutions of (3) take the standard form

$$a_n^{(k)} = e^{ikn}, \tag{4}$$

with eigenvalues

$$E_k = E(x) + W + (N/2)M\omega_0^2 x^2 - 2J \cos k, \tag{5}$$

[where $W = W(x, x, \cdots)$ and where $J(x_n, x_{n\pm1})$ is set equal to a negative constant, $-J$, in recognition of the fact that the potential, $U(\mathbf{r} - n\mathbf{a}, x_n)$, which enters into the overlap integral, (I-15), is intrinsically negative], and group velocities

$$V_K = \frac{a}{\hbar}\frac{\partial E_k}{\partial k} = -\frac{2Ja}{\hbar}\sin k. \tag{6}$$

The system thus exhibits typical band character; in particular, at the band extrema ($k = 0, \pi$), the motion may be described in terms of an effective mass

$$m^* \equiv \left(\frac{a}{\hbar}\frac{\partial v_k}{\partial k}\right)^{-1} = \pm\hbar^2/2Ja^2. \tag{7}$$

In proceeding further, it is desirable to introduce a number of simplifications into (3). These are:

(a) The neglect of the energies, $W_n(x_1, \cdots, x_N)$.

(b) The neglect of the x-dependence of the $J(x_n, x_{n\pm1})$; this simplification means that the $J(x_n, x_{n\pm1})$ are all to be taken equal to a single constant, $-J$.

(c) The x-dependence of $E(x_n)$ is taken to be linear, i.e.,

$$E(x_n) = -Ax_n \tag{8}$$

(the sign being chosen negative, and the additive energy constant being set equal to zero, both in the interests of notational convenience).

The physical significance of (a) is that, when the electron is localized on a given molecular site, the perturbation of its expectation-energy due to the potentials of other molecules is assumed small—an assumption which is appropriate to the tight-binding case. Its employment has the effect of restricting the x-dependence of the expectation-energy of a localized electron to a single coordinate, x_n—that of the occupied site. This restriction appears to be a special feature of the molecular-crystal model; in particular, it does not apply to the real situation, in which the energy of a localized electron depends upon lattice displacements at a number of more or less neighboring sites. However, at the present time, it is not at all established whether this difference is significant, as far as physical results are concerned. In fact, the point of view of the present studies is

that such questions will best be answered by a systematic consideration of various hypothetical cases, starting with the simplest. In the spirit of this approach, $W_n(x_1, \cdots, x_N)$ will be neglected, as a suitable first step.

Similarly, the procedure of taking the J's equal to a constant may be inapplicable in some cases. In principle, the J's do depend on the x's mainly via the local wave functions, $\phi(\mathbf{r} - n\mathbf{a}, x_n)$; however, as will be seen, e.g., in II, the primary physical property of the J's is their order of magnitude relative to other energies. Hence, at least in the initial phases of the program, their x-dependence will be neglected. Finally, with regard to assumption (c), it may be pointed out that the term $E(x_n)$ plays the role of electron-lattice interaction; its approximation by (8) leads to a linear dependence of this interaction of the lattice-vibration coordinates, a feature which is characteristic of all treatments of the problem (known to the present author).

With the simplifications discussed above—namely, the neglect of the W's, the equating of the J's to a constant $(-J)$, and the use of (1.13), the equations of "motion," (1.3) reduce to

$$\left[i\hbar \frac{\partial}{\partial t} - \sum_m \left(- \frac{\hbar^2}{2M} \frac{\partial^2}{\partial x_m{}^2} + \frac{1}{2} M\omega_0{}^2 x_m{}^2 \right) + Ax_n \right] a_n(x_1, \cdots, x_n) + J(a_{n+1} + a_{n-1}) = 0. \tag{9}$$

Equation (9) will form the basis for a number of studies of the polaron problem (the first of which is presented in II). These studies will be focussed primarily on the case of the "small" polaron, the linear dimensions of which are of the order of a lattice spacing. For introductory purposes, it will be convenient at this time to discuss the alternate case of the "large" polaron. In this case, the linear dimensions of the polaron are sufficiently large compared to a lattice spacing so that the "wave function," $a_n(x_1, \cdots, x_n)$ may be taken to a be a continuous, differentiable function of a continuous position variable, with, e.g.,

$$a_{n+1} = a_n + \frac{\partial a_n}{\partial n} + \frac{1}{2} \frac{\partial^2 a_n}{\partial n^2}. \tag{10}$$

The present discussion will be concerned with the adiabatic theory of the large polaron as developed by Pekar (*1*). Specifically, it will be found that the "molecular-crystal" analog of the Pekar-type adiabatic solution (see Ref. *1*, Chapter II)—i.e., that solution which, with neglect of the vibrational kinetic energy, minimizes the sum of electronic and vibrational potential energies—may be obtained exactly in closed form.

The starting-point is the system of equations

$$E(x_1, \cdots, x_N)a_n(x_1, \cdots, x_N) = (\tfrac{1}{2}\sum_m M\omega_0{}^2 x_m{}^2 - Ax_n)a_n(x_1, \cdots, x_N) - J(a_{n-1} + a_{n-1}), \tag{11}$$

which, subject to the neglect of the vibrational kinetic energy operator,

$$-\frac{\hbar^2}{2M}\sum_m \frac{\partial^2}{\partial x_m{}^2},$$

is the stationary equivalent of (9). Here, the energy eigenvalue, $E(x_1, \cdots, x_N)$, is a parametric function of the vibration coordinates, $x_1, \cdots, x_N$; one seeks those values, $x_N^{(0)}, \cdots, x_1^{(0)}$, for which $E(x_1, \cdots, x_N)$ is a minimum.

Upon multiplying (11) by a_n^* and summing over n, one obtains

$$E(x_1, \cdots, x_N) = \tfrac{1}{2}\sum_m M\omega_0{}^2 x_m{}^2 - \sum_n A x_n \mid a_n \mid^2 + \sum_n J(a_{n-1} + a_{n+1})a_n^* \tag{12}$$

where use has been made of the normalizing condition

$$\sum_n \mid a_n \mid^2 = 1. \tag{13}$$

Upon differentiating (12) with respect to a given vibrational coordinate, x_p [account being taken of (13) by the introduction of a Lagrange multiplier, λ] and rearranging the summations in the J-proportional terms, one obtains

$$\begin{aligned}\frac{\partial E(x_1, \cdots, x_N)}{\partial x_p} &= M\omega_0{}^2 x_p - A \mid a_p \mid^2 \\ &+ \sum_n \frac{\partial a_n^*}{\partial x_p}[-Ax_n a_n - J(a_{n-1} + a_{n+1}) + \lambda a_n] \\ &+ \sum_n \frac{\partial a_n}{\partial x_p}[-Ax_n a_n^* - J(a_{n-1}^* + a_{n+1}^*) + \lambda a_n^*]\end{aligned} \tag{14}$$

(plus "boundary" terms, proportional to products of the a_n's at $n = 0$ and $n = N$, which, as verified below, may be assumed to vanish). One now substitutes (11) and its complex conjugates into the first and second square brackets of (14), respectively, obtaining for the last two terms of (14) the expression.

$$\left[E(x_1, \cdots, x_N) - \frac{1}{2}\sum_m M\omega^2 x_m{}^2 + \lambda\right]\sum_n \left(a_n \frac{\partial a_n^*}{\partial x_p} + a_n^* \frac{\partial a_n}{\partial x_p}\right),$$

which is obviously zero because of (13). Equation (14) then reduces to

$$\frac{\partial E(x_1, \cdots, x_N)}{\partial x_p} = M\omega_0{}^2 x_p - A \mid a_p \mid^2. \tag{15}$$

Upon setting the right-hand side of (15) equal to zero for all p, one obtains for the $x_p^{(0)}$ the relations

$$x_p^{(0)} = \frac{A \mid a_p^{(0)} \mid^2}{M\omega_0{}^2} \tag{16}$$

where $a_p^{(0)}$ is that solution of (11) which corresponds to $x_n = x_n^{(0)}$. Upon substituting (16) into (11), and introducing the quantity

$$-\epsilon \equiv E(x_1, \cdots, x_N) - \tfrac{1}{2} \sum_m M\omega_0^2 x_m^{(0)^2} + 2J \tag{17}$$

one obtains

$$\left(\frac{A^2}{M\omega_0^2} \mid a_n \mid^2 - \epsilon\right) a_n + J(a_{n+1} + a_{n-1} - 2a_n) = 0, \tag{18}$$

where the superscript, "(0)," has been dropped with the understanding that a_n will henceforth denote that solution of (11) for which $E(x_1, \cdots, x_N)$ is a minimum.

At this point it is desirable to make use of the "continuum" approximation, (10) (the applicability of which, as discussed above, characterizes the case of the large polaron). The introduction of (10) into (18) yields

$$\left(\frac{A^2}{M\omega_0^2} \mid a_n \mid^2 - \epsilon\right) a_n + J \frac{\partial^2 a_n}{\partial n^2} = 0. \tag{19}$$

Upon multiplying (19) and its complex conjugate with $\partial a_n^*/\partial n$ and $\partial a_n/\partial n$, respectively, and adding, one obtains

$$\frac{A^2}{2M\omega^2} \mid a_n \mid^4 - \epsilon \mid a_n \mid^2 + J \frac{\partial a_n}{\partial n} \frac{\partial a_n^*}{\partial n} = C_1 \tag{20}$$

where C_1 is a constant to be determined presently.

In proceeding further, it is expedient to write a_n in the form

$$a_n = b_n e^{i\beta_n} \tag{21}$$

where b_n and β_n are real (and b_n positive, as well). Introducing (21) into (19), and equating real and imaginary parts to zero, one finds for the imaginary part

$$2 \frac{d\beta_n}{dn} \frac{db_n}{dn} + b_n \frac{d^2\beta_n}{dn^2} = 0, \tag{22}$$

a first integration of which yields

$$\frac{d\beta_n}{dn} b_n^2 = C_2 \tag{23}$$

where C_2 is another constant, which will now be shown to vanish.

Introducing (21) and (22) into (20), one has

$$\frac{A^2}{2M\omega_0^2} b_n^4 - \epsilon b_n^2 + J\left[\left(\frac{db_n}{dn}\right)^2 + \frac{C_2^2}{b_n^2}\right] = C_1.$$

At this point, let us impose the condition that b_n differs from zero only in a limited

region of the specimen, so that as $| n | \rightarrow \infty$, $b_n \rightarrow 0$ and $db_n/dn \rightarrow 0$. It is then immediately seen that, in order for this condition to be obeyed, it is necessary that both C_2 and C_1 vanish. One thus arrives at the result that β_n is equal to a constant (which may for convenience be set equal to zero, so that $b_n = a_n$) and (23) reduces to

$$\frac{A^2}{2M\omega_0^2} a_n^4 - \epsilon a_n^2 + J\left(\frac{da_n}{dn}\right)^2 = 0. \tag{24}$$

Introducing the notations

$$\alpha^2 \equiv \frac{A^2}{2M\omega_0^2 J}, \tag{25}$$

$$\gamma^2 \equiv \epsilon/J, \tag{26}$$

one may write (24) in the form

$$\frac{da_n}{dn} = \pm a_n(\gamma^2 - \alpha^2 a_n^2)^{1/2}. \tag{27}$$

The integration of (27) is straightforward, the result being

$$a_n = \frac{\gamma}{\alpha} \operatorname{sech}[\gamma(n - n_0)] \tag{28}$$

where n_0 is a constant of integration, indicative of the fact that the centroid of the polaron may be located at an arbitrary point in the specimen.

The application to (28) of the normalizing condition, which in the continuum approximation, reads

$$\int a_n^2 \, dn = 1,$$

yields the relation

$$\gamma = \alpha^2/2 = \frac{A^2}{4M\omega_0^2 J} \tag{29}$$

so that

$$a_n = \left(\frac{A^2}{8M\omega_0^2 J}\right)^{1/2} \operatorname{sech}\left[\left(\frac{A^2}{4M\omega_0^2 J}\right)(n - n_0)\right] \tag{30}$$

and

$$\epsilon = \frac{1}{16J}(A^2/M\omega_0^2)^2. \tag{31}$$

Inserting (31) and (16) into (17), one has for the total energy (exclusive, of

course, of the vibrational kinetic energy)

$$\begin{aligned} E(x_1^{(0)}, \cdots, x_N^{(0)}) \equiv E^{(0)} &= -2J - \frac{1}{16J}(A^2/M\omega_0^2)^2 \\ &\quad + \frac{1}{2}M\omega_0^2\left(\frac{A^2}{M\omega_0^2}\right)^2 \int a_n^4\, dn \\ &= -2J - \frac{1}{16J}\left(\frac{A^2}{M\omega_0^2}\right)^2 + \frac{1}{24J}\left(\frac{A^2}{M\omega_0^2}\right)^2 \\ &= -2J - \frac{1}{48J}\left(\frac{A^2}{M\omega_0^2}\right)^2. \end{aligned} \tag{32}$$

It should at this point be noted that (20) has an alternate solution, namely, the one which was excluded by the requirement [stated in the paragraph subsequent to (23)] that $a_n \to 0$ as $|n| \to \infty$. If this condition be withdrawn, (20) is also solved by

$$a_n = N^{-1/2} e^{ikn} \tag{33}$$

with

$$\epsilon = -Jk^2 \tag{34}$$

(plus the term $(1/N)(A^2/M\omega_0^2)$ which is to be discarded, since it is of order N^{-1}). Inserting (34) into (17), and continuing to discard terms of order N^{-1}, one has

$$E(x_1^{(B)}, \cdots, x_N^{(B)}) \equiv E^{(B)} = -2J + Jk^2 \tag{35}$$

where the superscript, (B), is used in recognition of the fact that (33) is an unbounded, "band-type," solution of the form discussed earlier [see Eq. (4) and subsequent].

The minimum energy for the band-type solution is attained when $k = 0$, and is

$$E_{\min}^{(B)} = -2J. \tag{36}$$

Comparing (36) with (32), we observe that the energy of the bound state is lower than that of the free state by an amount

$$E_p = \frac{1}{48J}(A^2/M\omega_0^2)^2. \tag{37}$$

E_p may be considered to be the zeroth order binding energy of the large polaron, i.e., the binding energy obtained in the approximation in which the vibrational kinetic energies are neglected.[4]

[4] For formal purposes, it is convenient to consider the vibrational mass, M, as a variable

With the aid of (30), explicit formulas may be obtained for a number of physical properties of the large polaron state.

(a) The equilibrium vibrational displacement, $x_n^{(0)}$, given by (17) and (30), is

$$x_n^{(0)} = \frac{A}{M\omega_0^2}\frac{A^2}{8M\omega_0^2 J}\operatorname{sech}^2\left[\frac{A^2}{4M\omega_0^2 J}(n - n_0)\right]. \tag{38}$$

(b) The potential-well, within which the electron is trapped, is described by the expression

$$V(n) \equiv \frac{A^2}{M\omega_0^2}|a_n|^2 = \frac{(A^2/M\omega_0^2)^2}{8J}\operatorname{sech}^2\left[\frac{A^2}{4M\omega_0^2 J}(n - n_0)\right]. \tag{39}$$

Its maximum value, achieved at the polaron centroid, n_0, is

$$V_{\max} = \frac{(A^2/M\omega_0^2)^2}{8J} \tag{40}$$

which is twice the "electronic binding-energy," ϵ, as given by (31), and six times the net binding energy of the polaron, as given by (37).[5]

(c) The linear dimension, L_p, of the large polaron, as may be seen from (30), (38), or (39), is of the order

$$L_p \approx a\left(\frac{4J}{A^2/M\omega_0^2}\right) \tag{41}$$

where a is the lattice spacing. It is in particular to be observed that the condition for the existence of the large polaron, for which $L_p \gg a$, is

$$2J \gg A^2/2M\omega_0^2, \tag{42}$$

i.e., the electronic "bandwidth," $2J$, is to be large compared to $A^2/2M\omega_0^2$.

The physical significance of the quantity, $A^2/2M\omega_0^2$, may be understood by considering briefly the case of the small polaron, for which the opposite condition

$$2J \ll A^2/2M\omega_0^2 \tag{43}$$

parameter. If, in this "thought-variability," the "restoring-force constant," $M\omega_0^2$, be kept fixed, so that the $\omega_0 \propto M^{-1/2}$, E_p gives the polaron binding energy in the limit of $M \to \infty$. It is then apparent that in the further development of the adiabatic theory, higher-order corrections to E_p occur in increasing powers of $1/M$.

[5] It may be remarked in passing that the difference between ϵ and E_p is the energy required to "polarize" the lattice, i.e., it is the work done against the restoring forces, $M\omega_0^2 x_n$, in displacing each x_n from zero to $x_n^{(0)}$. Comparison of (31) and (37) demonstrates that two thirds of the gain in electron-lattice interaction energy (as given by ϵ) has to be returned to the lattice in the form of vibrational potential energy, leaving a net binding energy, $E_p = (\frac{1}{3})\epsilon$. This numerological feature is also characteristic of Pekar's solution (see Ref. *1*, Chapter II).

holds. In this case (18) is to be solved by a perturbation expansion in powers of J. To the first order in J, the self-consistent solutions are found to be

$$a_n = \delta_{n,n_0} + \frac{J}{A^2/M\omega_0^2}[\delta_{n-1,n_0} + \delta_{n+1,n_0}] \tag{44}$$

with

$$\begin{aligned} \epsilon &= A^2/M\omega_0^2 - 2J \\ E^{(0)}(x_1, \cdots, x_N) &= -A^2/2M\omega_0^2 \end{aligned} \tag{45}$$

and n_0 an arbitrary site. Since the minimum energy of the band-type solution is still given by (36), the polaron binding energy is

$$E_p = A^2/2M\omega_0^2 - 2J. \tag{46}$$

From (46) it is seen that the quantity $A^2/2M\omega_0^2$ is the maximum binding energy of the polaron, attained in the limit of an infinitely narrow electronic band width ($J = 0$). Both from (41) and (44), it is clear that the ratio of this maximum binding energy to the electronic bandwidth determines the size of the polaron.

MOLECULAR-CRYSTAL POLARIZATION (MCP) MODEL

It is finally of interest to discuss an alternate molecular-crystal model—to be designated as the "Molecular-Crystal Polarization (MCP) Model"—which is closely related to the standard continuum-polarization model, but in addition takes explicit account of the discreteness of crystal structure.

The host crystal of the MCP model is the three-dimensional analog of the one-dimensional molecular-crystal described at the beginning of this paper. The lattice vibrations consist, as before, of the internuclear vibrations, x_g, of the individual diatomic molecules; the only difference is that the site-index, $\mathbf{g}$, is now a three-dimensional lattice-site vector.[6]

The interaction term of the MCP model is taken to have the form

$$H_{\text{int}} = -A \sum_{g'} x_{g'}/|\mathbf{g} - \mathbf{g}'|^2 \tag{47}$$

where the sum goes over $\mathbf{g}' \neq \mathbf{g}$. This basically long-range interaction, which replaces the short-range interaction

$$H_{\text{int}} = -Ax_n = -A \sum_{n'} \delta_{nn'} x_{n'} \tag{48}$$

of the previously discussed one-dimensional model, in effect, invests the displacements, x_g, with a dipolar character; in line with this property, the interaction

[6] When appearing as subscripts, lattice-site vectors, g, g', etc., will be denoted by light-face roman letters.

is seen to depend on the inverse square of the distance between the electron site, $\mathbf{g}$, and the "dipole" site, $\mathbf{g}'$.

The principal difference between (47) and a real electron-dipole interaction is that (47) is a scalar interaction involving a scalar dipole moment, x_g, in place of the vector-dipole moment of the real interaction. However, as will be seen below, this feature does not invalidate the correspondence of the MCP model with the standard continuum-polarization model.

Finally, it is necessary to generalize the overlap term of (9) so as to apply to the three-dimensional case. The most obvious such generalization is the replacement of $J(a_{n+1} + a_{n-1})$ by

$$J \sum_h a_{g+h}$$

where the relative site-vector, h, goes over nearest neighbors.

Introducing these modifications into Eq. (9), one has

$$\left[i\hbar \frac{\partial}{\partial t} - \sum_g \left(-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x_g^2} + \frac{1}{2} M\omega_0^2 x_g^2 \right) + A \sum_g x_{g'} / |\mathbf{g} - \mathbf{g}'|^2 \right] \cdot a_g(x_p, \cdots, x_N) + J \sum_h a_{g+h}(x_1, \cdots, x_N) = 0. \tag{49}$$

In order to exhibit the relationship of (49) to the continuum-polarization model, it is desirable to express the vibration coordinates in terms of normal modes.[7]

$$q_k = \left(\frac{2}{N}\right)^{1/2} \sum_{g=1}^{N} x_g \sin(\mathbf{k}\cdot\mathbf{g} + \pi/4). \tag{50}$$

Introducing (50) into (49), one obtains, after some algebra

$$\left[i\hbar \frac{\partial}{\partial t} - \sum_k \left(-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial g_k^2} + \frac{1}{2} M\omega_0^2 q_k^2 \right) + \left(\frac{2}{N}\right)^{1/2} A_k \sin(\mathbf{k}\cdot\mathbf{g} + \pi/4) \right] \cdot a_g[\cdots q_k \cdots] + J \sum_g a_{g+h}(\cdots q_k \cdots) = 0 \tag{51}$$

where, with $\mathbf{f}$ a relative site-vector,

$$A_k \equiv A \sum_f \cos(\mathbf{k}\cdot\mathbf{f})/f^2 \tag{52}$$

[7] For the sake of straightforwardness of enumeration, the sample will be assumed to have the shape of a cube, each side of which has a length equal to $(2G + 1)$ lattice spacings (with G integral), so that $N = (2G + 1)^3$. The permissible values of the cartesian components of $\mathbf{k}$ are, then, $k_i = 2\pi\kappa_i/(2G + 1)$, where $-G \leqq \kappa \leqq G$. With these specifications, the transformation given by (50) is orthogonal and hence immediately applicable to (49). It may be remarked here that the particular choice of transformation coefficients, $\sin(\mathbf{k}\cdot\mathbf{g} + \pi/4)$, is the same as that employed by Pekar and Buimistrov (*2*).

use having been made of the fact that, for the simple monomolecular cubic crystal under consideration, lattice sums involving odd functions of $\mathbf{f}$ vanish.

It will now be noted that, in the continuum approximation (in which $\mathbf{k}$ is to be restricted to magnitudes small compared to those of reciprocal lattice vectors), the sum in (52) may be approximated by an integral, i.e.,

$$A_k \cong A \int \frac{\cos\mathbf{k}\cdot\mathbf{f}}{f^2} d^3\mathbf{f} = 2\pi^2 A/k. \tag{53}$$

Furthermore, in the spirit of the continuum approximation, one may introduce the three-dimensional analog of the expansion (10), i.e.,

$$a_{g+h} = a_g + \sum_i h_i \frac{\partial a}{\partial g_i} + \frac{1}{2}\sum_{ij} h_i h_j \frac{\partial^2 a}{\partial g_i^2} \tag{54}$$

which, when substituted into the last term of (51), yields $J\nabla_g^2 a_g$. The continuum approximation to (51) is thus

$$\left[ih\frac{\partial}{\partial t} - \sum_k \left(-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial q_k^2} + \frac{1}{2}M\omega_0^2 q_k^2\right) + \sum_k \left(\frac{2\pi^2 A}{k}\right) q_k \left(\frac{2}{N}\right)^{1/2} \sin(\mathbf{k}\cdot\mathbf{g} + \pi/4)\right] \times a_g(\cdots q_k \cdots) + J\nabla_g^2 a_g = 0. \tag{55}$$

In order to bring (55) into a more standard form, it is desirable to introduce dimensionless oscillator coordinates

$$Q_k = \left(\frac{M\omega_0}{\hbar}\right)^{1/2} q_k \tag{56}$$

in terms of which (55) becomes

$$i\hbar\frac{\partial a_g}{\partial t} = Ha_g \tag{57}$$

where

$$H = -\frac{\hbar^2}{2m^*}\nabla_g^2 + \sum_k \frac{\hbar\omega_0}{2}\left(-\frac{\partial^2}{\partial Q_k^2} + Q_k^2\right) + \sum C_k Q_k \left(\frac{2}{N}\right)^{1/2} \sin(\mathbf{k}\cdot\mathbf{g} + \pi/4) \tag{58}$$

with

$$m^* = \hbar^2/2J \tag{59}$$

and

$$C_k = -\frac{2\pi^2 A}{k}\left(\frac{\hbar}{M\omega_0}\right)^{1/2}. \tag{60}$$

Comparison with the continuum-polarization Hamiltonian (given, e.g., in Ref. 6) demonstrates a complete equivalence, provided that the interaction constant, A, hitherto unspecified, be set equal to[8]

$$A = \left(\frac{M\omega_0^2\beta e^2}{\pi^3 a}\right)^{1/2} \tag{61}$$

where $\beta \equiv 1/\epsilon_\infty - 1/\epsilon$ (ϵ_∞ and ϵ representing the high-frequency and static dielectric constants, respectively).

The above-established identity of the MCP and continuum-polarization models demonstrates that, in the degree of approximation commonly employed in polaron theory—i.e., in the continuum approximation—systems of different crystallographic constitution may nevertheless be equivalent. This equivalence, of course, gets lost when the theory is applied to problems in which distances comparable to lattice spacings are important, as is the situation for the small polaron. In the case of a real ionic crystal, it then becomes necessary to consider the coupling of the electron, not only to the longitudinal optical modes, as is done in the continuum-polarization model, but also to the two transverse optical and three acoustical branches. Hence, any small-polaron treatment in which the electron is assumed to interact with but a single branch of vibration modes, will necessarily be based on a model, which, like the MCP-model, differs physically from the real case. However, as stated earlier, the basic point of view of this and the following paper is that the general features of the small-polaron problem are sufficiently independent of the details of the physical system, so that the methods developed for treating the simplified models should eventually be applicable to the real situation.

APPENDIX I

The motion of the electron (or hole) in the one-dimensional molecular-crystal medium, described in the text, will here be formulated in terms of the tight-binding approximation. In analogy with the corresponding treatment for atomic sites, one assumes that the one-electron potential may be written as a sum of "molecular" potentials, viz.,

$$V(\mathbf{r}, x_1, \cdots, x_n) = \sum_{n=1} U(\mathbf{r} - n\mathbf{a}, x_n) \tag{I-1}$$

where $\mathbf{a}$ is the (one-dimensional) unit lattice vector and where the molecular

[8] The occurrence in (61) of the lattice spacing, a, arises from the circumstance that, in the present work, k is expressed in units of $1/a$, whereas the wave-vector, κ, of Footnote 7 is given in cgs.

potential, $U(\mathbf{r}, x_n)$ is assumed to be sufficiently short-range so that, e.g.,

$$U(\mathbf{a}, x_n) \ll U(0, x_n). \tag{I-2}$$

The basic electron wave functions, $\varphi(\mathbf{r} - n\mathbf{a}, x_n)$, appropriate to the tight-binding approximation, are obtained as solutions of the equation

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r} - n\mathbf{a}, x_n) \right] \phi(\mathbf{r} - n\mathbf{a}, x_n) = E(x_n)\phi(\mathbf{r} - n\mathbf{a}, x_n) \tag{I-3}$$

with eigenvalues $E(x_n)$, which all depend upon the local internuclear displacement coordinate in the same way. The total wave function of the system is then represented as a superposition

$$\psi(\mathbf{r}, x_1, \cdots, x_n) = \sum_n a_n(x_1, \cdots, x_n)\phi_n(\mathbf{r}, x_n). \tag{I-4}$$

Equations for the $a_n(x_1, \cdots, x_n)$ are obtained by substitution of (I-4) into the Schrödinger equation of the system[9]

$$\begin{aligned} i\hbar \frac{\partial \psi}{\partial t} = \Bigg[-\frac{\hbar^2}{2m} \nabla^2 + \sum_{n=1}^{n} U(\mathbf{r} - n\mathbf{a}, x_n) \\ + \sum_{n=1}^{n} \left(-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x_n^2} + \frac{1}{2} M\omega_0^2 x_n^2 \right) \Bigg] \psi \end{aligned} \tag{I-5}$$

followed by multiplication on the left with $\phi_n^*(\mathbf{r}, x_n)$ and integration over the electron coordinate, $\mathbf{r}$. This procedure gives

$$\begin{aligned} &\sum_m T_{nm} \left[i\hbar \frac{\partial}{\partial t} - \sum_p \left(-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial x_p^2} + \frac{1}{2} M\omega_0^2 x_p^2 \right) - E(x_m) \right] a_m(x_1, \cdots, x_p) \\ &\quad = \sum_{\substack{m \\ p \neq m}} \left[\int \phi^*(\mathbf{r} - n\mathbf{a}, x_n) U(\mathbf{r} - p\mathbf{a}, x_p) \phi(\mathbf{r} - m\mathbf{a}, x_m)\, dV \right] \\ &\quad \cdot a_m(x_1, \cdots, x_p) - \frac{\hbar^2}{2M} \sum_m \int \phi^*(\mathbf{r} - n\mathbf{a}, x_n) \\ &\quad \cdot \left[2 \frac{\partial \phi(\mathbf{r} - m\mathbf{a}, x_m)}{\partial x_m} \frac{\partial a_m}{\partial x_m} + a_m \frac{\partial^2 \phi(\mathbf{r} - m\mathbf{a}, x_m)}{\partial x_m^2} \right] dV \end{aligned} \tag{I-6}$$

where

$$\begin{aligned} T_{nm} &\equiv \int \phi^*(\mathbf{r} - n\mathbf{a}, x_n)\phi(\mathbf{r} - m\mathbf{a}, x_m)\, dV \\ &= \delta_{nm} + S_{nm}. \end{aligned} \tag{I-7}$$

[9] H_{int} is implicitly contained in (I-5) by virtue of the x_n dependence of the U's.

The S_{nm} are the well-known nonorthogonality integrals which will here be assumed to differ from zero only for nearest neighbors and to be small compared to unity; also, with normalized φ's, one has $S_{nn} = 0$.

It is now expedient to multiply (I-6) by the matrix $\check{T}_{nm}$, which is the inverse of T_{nm}, as defined by the relation

$$\sum_p \check{T}_{np} T_{pm} = \delta_{nm} \tag{I-8}$$

and as given to first order in the S_{nm} by the expression

$$\check{T}_{nm} = \delta_{nm} - S_{nm} + \cdots . \tag{I-9}$$

One then obtains in place of (I-6)

$$\begin{aligned}\left[i\hbar \frac{\partial}{\partial t} - \sum_p \left(- \frac{\hbar^2}{2M} \frac{\partial^2}{\partial x_p{}^2} + \frac{1}{2} M \omega_0{}^2 x_p{}^2 \right) - E(x_n) \right] a_n(x_1, \cdots, x_p) \\ = \sum_{\substack{l,m \\ p \neq m}} \check{T}_{nl} \left[\int \phi^*(\mathbf{r} - l\mathbf{a}, x_l) U(\mathbf{r} - p\mathbf{a}, x_p) \phi(\mathbf{r} - m\mathbf{a}, x_m)\, dV \right] \\ \cdot a_m(x_1, \cdots, x_p) - \frac{\hbar^2}{2M} \sum_{l,m} \check{T}_{nl} \int \phi^*(\mathbf{r} - l\mathbf{a}, x_l) \\ \cdot \left[2 \frac{\partial \phi(\mathbf{r} - m\mathbf{a}, x_m)}{\partial x_m} \frac{\partial a_m}{\partial x_m} + \frac{\partial^2 \phi(\mathbf{r} - m\mathbf{a}, x_m)}{\partial x_m{}^2} a_m \right] dV .\end{aligned} \tag{I-10}$$

The rather complicated expressions on the right-hand side of (I-10) may be simplified by the introduction of assumptions appropriate to the case of tight-binding. The principal assumption is that of the smallness of overlap integrals i.e., integrals involving electron wave functions localized at different sites. These integrals are of two types: nonorthogonality integrals, as given by the S_{nm}, and those terms on the right-hand side of (I-10) for which $l \neq m$. In line with the above assumption, all terms containing products of more than one such overlap integral will henceforth be discarded.

When the terms of (I-10) involving the U's are subjected to this procedure, they reduce to the expression

$$\begin{aligned}\sum_{\substack{m \\ p \neq m}} \left[\int \phi^*(\mathbf{r} - n\mathbf{a}, x_n) U(\mathbf{r} - p\mathbf{a}, x_p) \phi(\mathbf{r} - m\mathbf{a}, x_m)\, dV \right] a_m(x_1, \cdots, x_p) \\ - \sum_{\substack{m \\ p \neq m}} S_{nm} \left[\int \phi^*(\mathbf{r} - m\mathbf{a}, x_m) U(\mathbf{r} - p\mathbf{a}, x_p) \phi(\mathbf{r} - m\mathbf{a}, x_m)\, dV \right] \\ \cdot a_m(x_1, \cdots, x_p).\end{aligned} \tag{I-11}$$

At this stage we may employ a second assumption, also appropriate to the case of tight-binding, which involves the local character of the molecular potential $U(\mathbf{r} - p\mathbf{a}, x_p)$. Specifically, it will be assumed that those terms in (I-11) which

contain products of wave functions and potentials localized at different sites are an order of magnitude smaller than terms for which two (or more) such factors refer to the same site. It then follows that the first member of (I-11) may be approximated by

$$\sum_{m\neq n}\left[\int \phi^*(\mathbf{r}-n\mathbf{a},x_n)U(\mathbf{r}-n\mathbf{a},x_n)\phi(\mathbf{r}-m\mathbf{a},x_n)\,dV\right]a_m(x_1,\cdots,x_p)$$
$$\cdot\left[+\int|\phi(\mathbf{r}-n\mathbf{a},x_n)|^2\Big(\sum_{p\neq n}U(\mathbf{r}-p\mathbf{a},x_p)\Big)\,dV\right]a_n(x_1,\cdots,x_p) \tag{I-12}$$

(i.e., one of the two summation indices gets suppressed), and that the second member of (I-11) may be dropped.[10]

Applying a comparable simplification to the terms of (I-10) which contain derivatives of the φ's with respect to the internuclear coordinates, one obtains

$$-\frac{\hbar^2}{2M}\left[\int\phi^*(\mathbf{r}-n\mathbf{a},x_n)\,\frac{\partial^2\phi(\mathbf{r}-n\mathbf{a},x_n)}{\partial x_n{}^2}\right]a_n(x_1,\cdots,x_n)$$
$$-\frac{\hbar^2}{2M}\sum_{m\neq n}\int\phi^*(\mathbf{r}-n\mathbf{a},x_n)$$
$$\cdot\left[2\,\frac{\partial\phi(\mathbf{r}-m\mathbf{a},x_m)}{\partial x_m}\,\frac{\partial a_m}{\partial x_m}+\frac{\partial^2\phi(\mathbf{r}-m\mathbf{a},x_m)}{\partial x_m{}^2}\,a_m\right]dV \tag{I-13}$$
$$-\frac{\hbar^2}{2M}\sum_m S_{nm}\left[\int\phi^*(\mathbf{r}-m\mathbf{a},x_m)\,\frac{\partial^2\phi(\mathbf{r}-m\mathbf{a},x_m)}{\partial x_m{}^2}\,dV\right]a_m$$

in which use has been made of the fact that, since $\phi(\mathbf{r}-n\mathbf{a},x_n)$ is nondegenerate, and hence, real, the integral $\int\phi^*(\mathbf{r}-n\mathbf{a})(\partial/\partial x_n)\phi(\mathbf{r}-n\mathbf{a})\,dV$ vanishes.

Equation (I-13) is listed here primarily for reference purposes. In line with the standard procedure in which terms involving derivatives of electronic functions with respect to nuclear coordinates are considered to be small (they represent, in essence, the effect of the nuclear kinetic energy on the localized electron wave functions), and also in view of the fact that the molecular-crystal model, itself, constitutes a simplified abstraction of the real problem, (I-13) will herewith be discarded.

With these approximations, the right-hand side of (I-10) reduces to (I-12). Introducing the notations

$$W_n(x_1,\cdots,x_n)\equiv\int|\phi(\mathbf{r}-n\mathbf{a},x_n)|^2[\sum_{p\neq n}U(\mathbf{r}-p\mathbf{a},x_p)]\,dV, \tag{I-14}$$

$$J(x_n,x_m)\equiv\int\phi^*(\mathbf{r}-n\mathbf{a},x_n)U(\mathbf{r}-n\mathbf{a},x_n)\phi(\mathbf{r}-m\mathbf{a},x_m)\,dV \tag{I-15}$$

[10] In particular, it is smaller than the second sum of (I-12) by a factor of the order of a nonorthogonality integral, S_{nm}.

and assuming $J(x_n, x_m)$ appreciable only for nearest neighbors, one may write (I-10) in the form

$$\left[i\hbar\frac{\partial}{\partial t} - \sum_p \left(-\frac{\hbar}{2M}\frac{\partial^2}{\partial x_p^2} + \frac{1}{2}M\omega_0^2 x_p^2\right) - E(x_n) - W_n(x_1, \cdots, x_n)\right] \times a_n(x_1, \cdots, x_n) = \sum_{(\pm)} J(x_n, x_n \pm 1)a_n \pm 1(x_1, \cdots, x_n). \tag{I-16}$$

RECEIVED: May 11, 1959

REFERENCES

1. See, for example, S. I. PEKAR, "Untersuchungen Über die Electronentheorie der Kristalle." Akademie-Verlag, Berlin, 1954.
2. S. I. PEKAR AND V. M. BUIMISTROV, *J. Tech. Phys. U.S.S.R.* **2**, 2478 (1957) (Soviet Physics translation).