Progress of Theoretical Physics, Vol. 26, No. 1, July 1961

Self-Trapping of an Electron by the Acoustical Mode of Lattice Vibration. I

Yutaka TOYOZAWA

The Institute for Solid State Physics, University of Tokyo, Tokyo

(Received February 24, 1961)

The importance of the acoustical mode of lattice vibration for self-trapping of an electron is emphasized. It is shown that when the coupling constant between the electron and the acoustical mode vibration exceeds a certain critical value, the effective mass of the electron changes discontinuously to such an enormous value that it is practically allowed to take a localized self-trapping state as an eigenstate, in contrast to the case of polaron, in which the effective mass changes continuously with coupling constant. This difference is attributed to the different force range of electron-lattice interaction in the two cases.

§ 1. Introduction

Since Landau[1)] pointed out the possibility of self-trapping of an electron in deformable lattice as early as in 1933, much attention has been focused on the interaction between an electron and polar mode of lattice vibration.[2)~6)] One may well suppose that in ionic crystals the electrostatic field caused by the polarization wave interacts with the conduction electron more strongly than the acoustical mode does. In fact, the interaction in the former case is usually so strong[3),6)] that the perturbation theoretical treatment is not a very good approximation whereas it is in the latter case, as long as one starts from the band picture.

If the interaction of an electron with lattice vibration is very strong, the electron is supposed to 'dig its own hole and be trapped there'.[1),2),7)] On the other hand, true eigenstate of the electron-lattice system must be such that it has translational symmetry with wave vector $\boldsymbol{k}$, that is, it is a plane-wave like linear combination of localized electron states accompanied by surrounding lattice distortion,[8)] even if the localized states themselves are good approximations. When the interaction is not very strong, this way of description is not appropriate. One should rather describe the electron to be virtually emitting and absorbing phonons, suffering recoil each time, and as a whole it moves through the crystal with a certain wave vector $\boldsymbol{k}$.[3)] In any case, the word 'self-trapping' or 'localization' should not be taken so literally, and the problem of the possibility of self-trapping should be replaced by the discussion of the magnitude of the effective mass of this quasi-particle, that is, an electron accompanied by virtual phonons or lattice distortion.

In the case of a slow electron interacting with polar mode of lattice vibration, this quasi-particle has been called 'polaron', and several authors investigated its properties, such as self-energy, effective mass, spatial extension and so on.[3)~6)] According to them, especially to Feynman,[5)] the effective mass of the polaron increases rather rapidly but continuously as the interaction becomes stronger. It proved that neither discontinuity nor sudden increase of the effective mass appears, contrary to the expectation by some people. Thus, the word 'self-trapping' has only a relative meaning here, or has to be taken as a limiting case of very large effective mass of the quasi-particle. Moreover, in real crystals, it is limited to rather few cases that one can expect the effective mass or the time of staying of an electron within a unit cell is so large that the word self-trapping is appropriate.

According to the experimental works carried out hitherto, the electrons in alkali[9)] and silver halides[10)] are mobile though the mobility is much smaller than in the cases of Ge and Si, and through the analysis of the temperature dependence of the mobility it has been shown that the scattering of the electron (or exactly polaron) by polar mode vibration plays the main rôle except at rather low temperature where the scattering by acoustical mode vibration and by impurities becomes more important. As for the holes, the situation is the same in the case of AgBr,[11)] but it is reported that the hole in AgCl is several orders of magnitude less mobile than the electron.[10)] The holes in alkali-halides are now believed to be practically in self-trapped state since the experimental works by Känzig and others[12)] appeared.

This essentially qualitative, difference, or quantitatively quite large difference in the behaviors of holes in the cases of alkali-halides and AgCl on the one hand and of AgBr on the other hand, the other properties of both of which do not seem to have essential difference, suggests that the motion of an electron (or a hole) in deformable lattice is divided rather distinctly into mobile type and practically immobile type (self-trapping state) according as the electron-lattice interaction is smaller or larger than a certain critical value.

The interaction of an electron with polar mode, the importance of which has been emphasized by many authors and in fact has been proved experimentally for mobile electrons, fails to explain this discontinuity, as was mentioned above. In this paper we shall emphasize that the interaction with acoustical mode plays an essential rôle in self-trapping of an electron. The most important difference between the two kinds of interaction is that in the case of polar mode the interaction is of long range (electrostatic Coulomb potential) while it is of short range in the case of acoustical mode (deformation potential). The effect of the force range will be discussed in § 2, where we shall show, in rather intuitive way based on the continuum model, that the electron changes its behavior discontinuously from mobile type to self-trapping type when the strength of the interaction with acoustical mode exceeds certain critical value. In view of the strong localization of the electron and the lattice distortion in the self-trapping state, we shall take into account the discrete structure of the lattice explicitly in § 3, and investigate the eigenstate of the electron-lattice system with wave number $\boldsymbol{k}$, making use of the variation method, with the result that the effective mass of the electron increases discontinuously to an enormous value as the interaction exceeds the critical value. The validity of the approximations used in § 3 are discussed in the final section, more thorough calculations being left to the following paper.

§ 2. Preliminary discussions based on the continuum model

As was mentioned in § 1, the exact eigenstates of the electron-lattice system have the translational symmetry with wave number $\boldsymbol{k}$, and the localized self-trapping state is only an approximate eigenstate. Nevertheless, it seems suggestive in elucidating the effect of the difference in the force ranges of acoustical and polar modes to calculate the adiabatic potential for the localized self-trapping state. We shall take the continuum model for the lattice, for simplicity.

Acoustical mode of lattice vibration is nothing but the elastic wave as long as the wavelength is large compared with the interatomic distance. Let us assume that the crystal has deformed in such a way that the volume dilation $\Delta(\boldsymbol{r})$ is given by

$$\Delta(\boldsymbol{r})=\begin{cases}\Delta & (r<R),\\ 0 & (r>R).\end{cases} \tag{2·1}$$

The elastic potential energy for this deformation is given by

$$U=\frac{1}{2}C\frac{4\pi}{3}R^3\Delta^2, \tag{2·2}$$

where C is an appropriate elastic constant.*

According to the deformation potential theory,[13] an electron in the conduction band suffers the potential energy

$$E_1\Delta(\boldsymbol{r}) \tag{2·3}$$

due to this deformation, where E_1 is the deformation potential constant. If the

---

* It is to be noted that the local dilation (2·1) induces shearing strain field outside the sphere such that the medium is compressed along the radial direction and stretched in directions normal to it. Since the displacement of the medium at $r=R$ is equal to $\Delta R/3$, the displacement outside the sphere is given by $(\Delta R/3)\ (R/r)$, and the elastic energy stored outside the sphere is proportional to $\Delta^2R^4\int_R^\infty\left(\nabla\frac{1}{r}\right)^2 r^2\mathrm{d}r\propto\Delta^2R^3$. This energy is to be included in (2·2). C, therefore, is a somewhat complicated combination of elastic constants. As for the electron-lattice interaction, the shearing strain has no first order interaction with the electron in the case where the lattice is cubic and the minimum of the electronic energy band is located at $k=0$.[13] Hence we have the form (2·3) for the interaction energy.

sign of Δ is such that $E_1\Delta<0$, we have a three-dimensional square well potential problem for the electron. Let us denote the band effective mass of the electron by m^*. When the depth of the well is sufficiently large, that is, when

$$\delta\equiv-\frac{2m^*}{\hbar^2}R^2E_1\Delta>\left(\frac{\pi}{2}\right)^2, \tag{2·4}$$

there is a bound state, with binding energy

$$-E=\frac{\hbar^2}{2m^*R^2}(\xi\cot\xi)^2, \tag{2·5}$$

where ξ is the root of

$$\left(\frac{\xi}{\sin\xi}\right)^2=\delta\quad\left(\frac{\pi}{2}<\xi<\pi\right). \tag{2·6}$$

If (2·4) is not valid, there is no bound state and

$$E=0 \tag{2·7}$$

for the ground state of the electron.

The adiabatic potential for the ground state of the electron is now written as

$$W(\Delta, R)=U+E$$
$$=\begin{cases}\dfrac{\hbar^2}{2m^*}\dfrac{\delta^2}{R_0R} & \left(\delta<\left(\dfrac{\pi}{2}\right)^2\right),\\[2ex] \dfrac{\hbar^2}{2m^*}\left[\dfrac{\delta^2}{R_0R}-\dfrac{\delta\cos^2\xi}{R^2}\right] & \left(\delta>\left(\dfrac{\pi}{2}\right)^2\right),\end{cases} \tag{2·8}$$

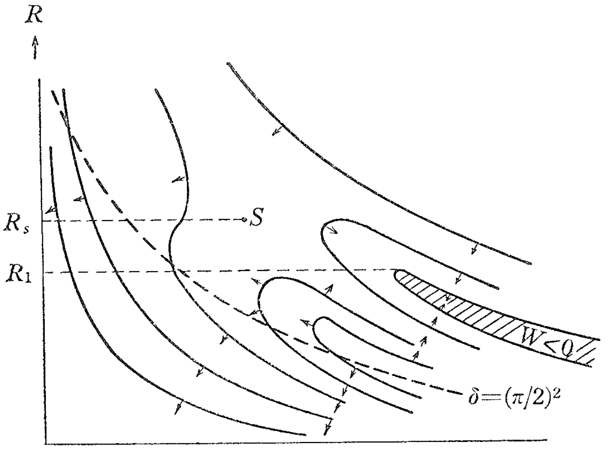


Fig. 1. Schematic representation of the equipotential lines for the adiabatic potential of the electron-acoustical mode system on the elastic continuum model. R is the radius of the sphere in which there is a uniform volume dilation Δ. The arrows give the direction of *decreasing* energy.

where

$$R_0=\frac{3m^*|E_1|^2}{\pi\hbar^2C}. \tag{2·9}$$

Taking R constant, (2·8) increases monotonously with δ for larger values of R, while it takes a maximum and a minimum, and then increases monotonously for smaller values of R. The equipotential line of (2·8) on the (Δ, R)-plane is shown schematically in Fig. 1. When δ is larger than $(\pi/2)^2$ (the right side of the broken line), a bound state for the electron appears, and when

$$R<R_1\equiv R_0\cos^2(2\xi_1)=0.0476R_0 \tag{2·10}$$

where $\xi_1=2.25$ is the root of

$$\frac{\tan(2\xi)}{2\xi}=1 \qquad \left(\frac{\pi}{2}<\xi<\pi\right),$$

W takes negative values for appropriate values of $\varDelta$ (the shaded region in Fig. 1).

If the crystal were perfectly continuous elastic medium, the stable state of the system corresponds to infinitesimally small R and infinitely large $\varDelta$, with infinitely deep adiabatic potential. Due to the atomic structure of the crystal, R must be larger than a' where a' is of the same order of magnitude with interatomic distance a.

Consequently, we can conclude as follows. If R_1 is larger than a', that is, if

$$\frac{m^*|E_1|^2}{\hbar^2 Ca'}>22.0, \tag{2·11}$$

the stable state of the system is such that the electron is trapped at the lattice distortion which takes as small a spatial extension as is compatible with atomic structure of the crystal, while the state in which the electron moves through the lattice without lattice distortion is metastable. The latter statement is evident because these two states are completely different from each other being separated by a *barrier* of adiabatic potential.

The saddle point of this barrier is given by

$$\left.\begin{aligned} R_s&=2\left(\frac{\sin 2\xi_s}{2\xi_s}\right)^2 R_0=0.0578R_0 \\ \delta_s&=\left(\frac{\xi_s}{\sin\xi_s}\right)^2 \\ W_s&=\frac{\hbar^2}{8m^*R_0^2}\,\frac{\xi_s}{\sin^6\xi_s\cos^2\xi_s} \end{aligned}\right\}, \tag{2·12}$$

where $\xi_s=1.929$ is the root of

$$\frac{1}{\xi}-\operatorname{ctn}\xi+\frac{1}{3}\tan\xi=0 \qquad \left(\frac{\pi}{2}<\xi<\pi\right).$$

It is important to note that at the saddle point a bound state is already split off from the conduction band, and that the adiabatic approximation is valid there.

The existence of the potential barrier between the two types of states is due to the fact that for the square well potential the bound state appears only when the depth or the radius is larger than a finite value. Before the energy gain due to the binding appears, the lattice must suffer finite distortion with finite elastic energy. However, one may well ask why one has to confine the distortion to the shape given by (2·1). Were there no path in the functional space of $\varDelta(\boldsymbol{r})$ through which one can get to the negative W region starting from the undistorted state without any potential barrier? In order that the bound state appears for any small distortion, $\Delta(\boldsymbol{r})$ must decay as r^{-1} or more slowly with increasing r. This makes the elastic energy $U=(C/2)\int\Delta(\boldsymbol{r})^2 d\boldsymbol{r}$ diverge at $r=\infty$. The energetically favorable distortion is such that $\Delta(\boldsymbol{r})$ is more or less localized. That is the reason why we took the form (2·1).

The situation is quite different in the case of polar mode. Denote by $\boldsymbol{P}(\boldsymbol{r})$ the local electric polarization due to the relative displacement of positive and negative ions. The space charge $\rho(\boldsymbol{r})$ due to this polarization is given by $\rho(\boldsymbol{r})=-\operatorname{div}\boldsymbol{P}(\boldsymbol{r})$, while the electrostatic potential energy for the electron is given by $-e\varphi(\boldsymbol{r})$ where $\Delta\varphi(\boldsymbol{r})=-4\pi\rho(\boldsymbol{r})$.

Let us take the polarization field such that

$$-e\varphi(\boldsymbol{r})=\begin{cases}-\dfrac{Ze^2}{r} & r>b,\\[2ex] -\dfrac{Ze^2}{b} & r<b,\end{cases}\tag{2·13}$$

where Z and b are parameters. $\boldsymbol{P}(\boldsymbol{r})$ is then given by

$$\boldsymbol{P}(\boldsymbol{r})=\begin{cases}-\dfrac{Ze}{4\pi}\dfrac{\boldsymbol{r}}{r^3} & r>b,\\[2ex] 0 & r<b,\end{cases}\tag{2·14}$$

and the electrostatic energy stored in the lattice is

$$U=\text{const.}\int\boldsymbol{P}(\boldsymbol{r})^2 d\boldsymbol{r}=\frac{1}{2K}\frac{Z^2}{b},\tag{2·15}$$

where K is an appropriately defined polarizability of the lattice. It is important to note that in the present case it is possible to choose a pattern of lattice distortion such that the lattice energy integral converges and yet there are bound states for any small magnitude of the distortion. The binding energy of the ground state is given by

$$-E=\frac{Ze^2}{2r_0},\quad r_0=\frac{\hbar^2}{m^*Ze^2},\tag{2·16}$$

if $r_0\gg b$ so that the flattening of the potential (2·13) at $r<b$ has no sizable effect, while if $r_0\ll b$ there is no effect of uncertainty principle and $-E$ is given by Ze^2/b itself. Combining these limiting cases we can write

$$-E=\frac{Ze^2}{2r_0+b}\tag{2·17}$$

for qualitative purpose.

The equipotential line of the adiabatic potential $W=U+E$ on the (Z, b)-plane is shown in Fig. 2. There is a negative energy region (shaded in the Figure) which begins from the undistorted state and there is only one minimum point,

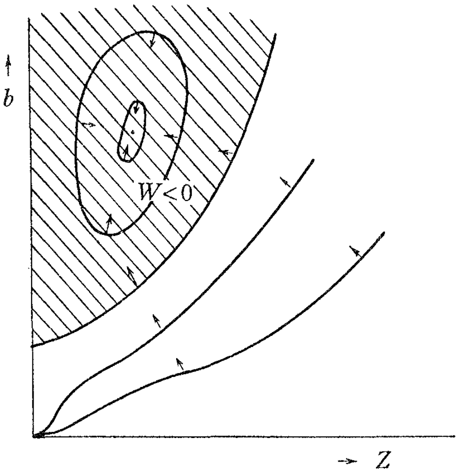


Fig. 2. Schematic representation of the equipotential line for the adiabatic potential of the electron-polar mode system on the dielectric continuum model b is the radius of the sphere on the surface of which there is a uniform charge, $+Ze$ in total, due to the polarization field surrounding the sphere. Z is a parameter representing the strength of polarization.

$$Z=\frac{8}{9}e^2K,\qquad b=\frac{9}{2}\frac{\hbar^2}{m^*e^4K}. \tag{2·18}$$

There is no potential barrier between this stable state and the undistorted state.

As the coupling constant K increases, the polarization Z of the stable state increases, and the radius b of the space charge and also the radius of electron orbit decreases. That is, the stable configuration changes continuously with coupling constant, in contrast to the case of acoustical mode where the stable state changes abruptly from the undistorted configuration into the strongly distorted configuration.

It is easy to see from the above discussions that the difference between these two modes of lattice vibration results from the different force ranges of electron-lattice interaction in the two cases. As is given by (2·3), a local dilation induces the potential for the electron at the same position only, that is, the interaction is a short-range force. The local space charge due to the electric polarization, on the other hand, gives rise to electrostatic potential at a distant point, too. A long-range potential, such as Coulomb potential, can always accommodate an electron, while a short-range potential, such as square well potential can bind an electron only when the depth of the potential becomes sufficiently large, at the cost of finite energy of the lattice.

Of course, the above discussion is not fully satisfactory in that we have not taken into account the translational symmetry of the eigenstate, and in that we have made use of the adiabatic approximation. Moreover, we found that for the self-trapping state in the case of acoustical mode, the continuum theory breaks down which compelled us to introduce a cut-off length a'. In the following section we shall explicitly take into account the discrete nature of the lattice which is essential for the self-trapping state, and when we apply a variation principle, we shall choose *ab initio* the trial eigenfunctions which is consistent with the translational symmetry of the lattice, in order to see how the effective mass changes with the coupling constant. We confine ourselves to the case of acoustical mode, because the case of polar mode has been fully explored by a number of authors.

§ 3. Discontinuous change of the effective mass with coupling constant

Let us take tight binding approximation for the electron in the crystal lattice, neglecting the overlap energies other than that for the nearest neighbors: $-V$. Making use of the second quantization formalism for the electron and phonons, we can write the total Hamiltonian as follows:

$$\begin{aligned} H &= K_e + I + H_L \\ &= -V \widehat{\sum_m \sum_n} a_m{}^* a_n \\ &\quad + \sum_m \sum_{\boldsymbol{w}} E_1 \left(\frac{\hbar}{2NMu}\right)^{1/2} iw^{1/2} \{b_{\boldsymbol{w}} \exp(i\boldsymbol{w}\cdot\boldsymbol{m}) - b_{\boldsymbol{w}}{}^* \exp(-i\boldsymbol{w}\cdot\boldsymbol{m})\} a_m{}^* a_m \\ &\quad + \sum_{\boldsymbol{w}} \hbar u w\, b_{\boldsymbol{w}}{}^* b_{\boldsymbol{w}}. \end{aligned} \qquad (3\cdot1)$$

Here a_m^* and a_m are those operators which create and annihilate, respectively, the electron at the m-th lattice, the position vector of which is denoted by $\boldsymbol{m}$. $b_{\boldsymbol{w}}^*$ and $b_{\boldsymbol{w}}$ are those for phonon with wave number $\boldsymbol{w}$. N, M and u are the total number of unit cells, the mass of a unit cell, and sound velocity, respectively. The symbol $\wedge$ over the double sum of the first term means that the summation is to be carried out only when m and n are nearest neighbors to each other. The second term is nothing but the Fourier expansion of the deformation potential given by (2·3), in a discrete lattice space.

We now investigate the lowest eigenstate of (3·1) with wave number $\boldsymbol{k}$, by a variational procedure. We apply the intermediate coupling method[15] which is usually expected to be useful when one wants to see how the weak coupling and strong coupling solutions are connected, and which has been successfully applied to the polaron problem.[4] One important difference between the polaron problem and the present problem is that the total pseudo-momentum operator,

$$\sum_k \hbar \boldsymbol{k}\, a_k{}^* a_k + \sum_w \hbar \boldsymbol{w} b_{\boldsymbol{w}}{}^* b_{\boldsymbol{w}},$$

where $a_k \equiv N^{-1/2} \sum_m \exp(-i\boldsymbol{k}\cdot\boldsymbol{m}) a_m$ is the annihilation operator for the electron in the Bloch state, does not commute with (3·1), because (3·1) includes the matrix elements for the Umklapp-process, as the result of explicit consideration of periodic structure of the lattice for the motion of the electron. It is inconvenient to cancel the electron coordinate by specifying the total momentum $\boldsymbol{k}$, as was done in the polaron problem.[4] Instead, we write the trial wave function of the total system as

$$|\Psi_k\rangle = N^{-1/2} \sum_m \exp(i\boldsymbol{k}\cdot\boldsymbol{m}) S_k(\boldsymbol{m}) T_k(\boldsymbol{m}) |0\rangle, \qquad (3\cdot2)$$

$$T_{\boldsymbol{k}}(\boldsymbol{m})\equiv\sum_{n}\phi_{k}(\boldsymbol{n})\,a_{m+n}^{*},\quad \sum_{n}|\phi_{k}(\boldsymbol{n})|^{2}=1, \tag{3·3}$$

$$S_{k}(\boldsymbol{m})\equiv\exp\Big[\sum_{w}f_{k}^{*}(\boldsymbol{w})\exp(i\boldsymbol{w}\cdot\boldsymbol{m})b_{w} \\ -(\text{hermite conjugate})\Big] \tag{3·4}$$

with $\phi(\boldsymbol{n})$ and $f(\boldsymbol{w})$ are trial functions for the electron and lattice configurations, respectively. $|0\rangle$ means the vacuum state for both electron and phonon.

Making use of the theorem that

$$\text{if } S\equiv\exp(\alpha^{*}b-\alpha b^{*})\quad (\alpha:\ \text{c-number}),$$

$$\text{then } S^{-1}bS=b-\alpha \text{ and } S^{-1}b^{*}S=b^{*}-\alpha^{*}, \tag{3·5}$$

we see that the displacement operator of the l-th atom

$$X(\boldsymbol{l})=\sum_{w}\{h(\boldsymbol{w})b_{w}\exp(i\boldsymbol{w}\cdot\boldsymbol{l})+(\text{h.c.})\}$$

is transformed by $S_{k}(\boldsymbol{m})$ into

$$S_{k}(\boldsymbol{m})^{-1}X(\boldsymbol{l})S_{k}(\boldsymbol{m})=X(\boldsymbol{l})-\sum_{w}\{h(\boldsymbol{w})f_{k}(w)\exp(i\boldsymbol{w}\cdot\boldsymbol{l}-\boldsymbol{m})+(\text{c.c.})\}.$$

in other words, $S_{k}(\boldsymbol{m})$ is the operator which makes the lattice deform, and the fact that the displacement represented by the second term is a function of $(\boldsymbol{l}-\boldsymbol{m})$ only, means that the pattern of distortion generated by $S_{k}(\boldsymbol{m}+\boldsymbol{n})$ is nothing but the pattern obtained as the result of translation of pattern $S_{k}(\boldsymbol{m})$ by lattice vector $\boldsymbol{n}$. This assures the translational symmetry of (3·2) with wave number $\boldsymbol{k}$. The m-th term of (3·2), $S_{k}(\boldsymbol{m})T_{k}(\boldsymbol{m})|0\rangle$, represents the state in which the electron is distributed with amplitude $\phi_{k}(\boldsymbol{n})$ around the m-th lattice point, while the lattice is in the zero point vibrational state around the displaced position corresponding to pattern $S_{k}(\boldsymbol{m})$.

Making use of elementary operational calculus (see the Appendix), we have

$$\langle 0|S_{k}^{*}(\boldsymbol{m})S_{k}(\boldsymbol{m}+\boldsymbol{l})|0\rangle=\exp(-\sigma_{k}(\boldsymbol{l})), \tag{3·6}$$

$$\sigma_{k}(\boldsymbol{l})\equiv\sum_{w}f_{k}^{*}(\boldsymbol{w})f_{k}(\boldsymbol{w})\{1-\exp(-i\boldsymbol{w}\cdot\boldsymbol{l})\}=\sigma_{k}^{*}(-\boldsymbol{l}) \tag{3·7}$$

for the overlap integral between the zero point vibrational states for the pattern $S_{k}(\boldsymbol{m})$ and $S_{k}(\boldsymbol{m}+\boldsymbol{l})$. In the same way we have

$$\langle\Psi_{k}|\Psi_{k}\rangle=\sum_{l}\exp\{i\boldsymbol{k}\cdot\boldsymbol{l}-\sigma_{k}(\boldsymbol{l})\}\sum_{n}\phi_{k}^{*}(\boldsymbol{n})\phi_{k}(\boldsymbol{n}-\boldsymbol{l}),$$

$$\langle\Psi_{k}|K_{e}|\Psi_{k}\rangle=-V\sum_{l}\exp\{i\boldsymbol{k}\cdot\boldsymbol{l}-\sigma_{k}(\boldsymbol{l})\}\sum_{n}\widehat{\sum_{\delta}}\phi_{k}^{*}(\boldsymbol{n})\phi_{k}(\boldsymbol{n}-\boldsymbol{l}+\boldsymbol{\delta}),$$

$$\langle\Psi_{k}|I|\Psi_{k}\rangle=-\sum_{l}\exp\{i\boldsymbol{k}\cdot\boldsymbol{l}-\sigma_{k}(\boldsymbol{l})\}\sum_{n}\phi_{k}^{*}(\boldsymbol{n})\phi_{k}(\boldsymbol{n}-\boldsymbol{l}) \tag{3·8}$$

$$\times\sum_{w}E_{1}\left(\frac{\hbar}{2NMu}\right)^{1/2}iw^{1/2}\{\exp(-i\boldsymbol{w}\cdot\boldsymbol{n})f_{k}(\boldsymbol{w})-\exp(i\boldsymbol{w}\cdot\boldsymbol{n}-\boldsymbol{l})f_{k}^{*}(\boldsymbol{w})\},$$

$$\langle\Psi_k|H_L|\Psi_k\rangle=\langle\Psi_k|\Psi_k\rangle\sum_{w}\hbar uw f_k^*(\boldsymbol{w})f_k(\boldsymbol{w}),$$

where $\frown$ on the summation means that $\boldsymbol{\delta}$ is to run over the nearest neighbours.

As the first approximation, let us neglect the spatial broadening of the electronic wave function and set $\phi(\boldsymbol{n})=\delta_{n,0}$ where $\delta_{n,m}$ is three-dimensional Kronecker's δ-function. (3·8) is greatly simplified then and we can write the expectation value of the total energy as

$$\begin{aligned} E(\boldsymbol{k}) &=\langle\Psi_k|H|\Psi_k\rangle \\ &=-V\widehat{\sum_{\delta}}\exp\{i\boldsymbol{k}\cdot\boldsymbol{\delta}-\sigma_k(\boldsymbol{\delta})\}+\sum_{w}\hbar uw f_k^*(\boldsymbol{w})f_k(\boldsymbol{w}) \\ &\quad -\sum_{w}E_1\left(\frac{\hbar}{2NMu}\right)^{1/2}iw^{1/2}\{f_k(\boldsymbol{w})-f_k^*(\boldsymbol{w})\}. \end{aligned} \tag{3·9}$$

Minimizing (3·9) with respect to $f_k(\boldsymbol{w})$, we have

$$f_k(\boldsymbol{w})=\frac{-iw^{1/2}E_1\left(\dfrac{\hbar}{2NMu}\right)^{1/2}}{V\widehat{\sum_{\delta}}\{1-\exp(-i\boldsymbol{w}\cdot\boldsymbol{\delta})\}\exp\{i\boldsymbol{k}\cdot\boldsymbol{\delta}-\sigma_k(\boldsymbol{\delta})\}+\hbar uw}. \tag{3·10}$$

Inserting this into (3·7) and (3·9), we have

$$\sigma_k(\boldsymbol{\delta})=\sum_{w}\frac{E_1^2\dfrac{\hbar}{2NMu}w\{1-\exp(-i\boldsymbol{w}\cdot\boldsymbol{\delta})\}}{[V\widehat{\sum_{\delta}}\{1-\exp(-i\boldsymbol{w}\cdot\boldsymbol{\delta}')\}\exp\{i\boldsymbol{k}\cdot\boldsymbol{\delta}'-\sigma_k(\boldsymbol{\delta}')\}+\hbar uw]^2},$$

$$\begin{aligned} E(\boldsymbol{k}) &=-V\widehat{\sum_{\delta}}\{1+\sigma_k(\boldsymbol{\delta})\}\exp\{i\boldsymbol{k}\cdot\boldsymbol{\delta}-\sigma_k(\boldsymbol{\delta})\} \\ &\quad -\sum_{w}\frac{E_1^2\dfrac{\hbar}{2NMu}w}{V\widehat{\sum_{\delta}}\{1-\exp(-i\boldsymbol{w}\cdot\boldsymbol{\delta})\}\exp\{i\boldsymbol{k}\cdot\boldsymbol{\delta}-\sigma_k(\boldsymbol{\delta})\}+\hbar uw}. \end{aligned} \tag{3·12}$$

The denominators of (3·11) and (3·12) are real due to (3·7). (3·11) are ν simultaneous equations for ν unknowns $\sigma_k(\boldsymbol{\delta})$, where ν is the number of the nearest neighbors.

To the first order with respect to the coupling constant E_1^2, we have

$$E(\boldsymbol{k})=-V\widehat{\sum_{\delta}}\exp(i\boldsymbol{k}\cdot\boldsymbol{\delta})-\sum_{w}\frac{E_1^2\dfrac{\hbar}{2NMu}w}{V\widehat{\sum_{\delta}}\exp(i\boldsymbol{k}\cdot\boldsymbol{\delta})-V\widehat{\sum_{\delta}}\exp(i(\boldsymbol{k}-\boldsymbol{w})\cdot\boldsymbol{\delta})+\hbar uw} \tag{3·13}$$

of which the first term is the electronic band energy of the state $\boldsymbol{k}$, and the second term is the self-energy (at absolute zero of temperature) of this state due to the electron-lattice interaction. On the other hand, for very large coupling constant, we can neglect $\exp(-\sigma_k(\boldsymbol{\delta}))$ compared with unity, and we have

$$E(\boldsymbol{k}) = -\frac{E_1^2}{2Mu^2} \equiv -\frac{1}{2}D. \tag{3·14}$$

It is the main purpose of this section to see how the solution changes from the weak coupling limit (3·13) to the strong coupling limit (3·14) as we increase the coupling constant continuously.

In solving (3·11) we approximate the band energy by a parabolic form:

$$\varepsilon(\boldsymbol{k}) - \varepsilon(0) = -V\widehat{\sum_{\boldsymbol{\delta}}}\{\exp(i\boldsymbol{k}\cdot\boldsymbol{\delta}) - 1\} \approx \frac{\hbar^2 k^2}{2m^*} \tag{3·15}$$

without spoiling the essential feature of the matter. First we solve (3·11) for the state: $\boldsymbol{k}=0$. It is evident that in this case all $\sigma_0(\boldsymbol{\delta})$ are equal:

$$\sigma_0(\boldsymbol{\delta}) = \sigma, \tag{3·16}$$

and (3·11) reduces to a single equation

$$\sigma = 3g\gamma \exp(2\sigma)\int_0^1 \frac{t^3 dt}{\{t+\gamma\exp(\sigma)\}^2}, \tag{3·17}$$

while (3·12) can be written as

$$-\frac{E(0)}{\nu V} = (1+\sigma)\exp(-\sigma) + 3g\gamma\exp(\sigma)\int_0^1 \frac{t^2 dt}{t+\gamma\exp(\sigma)} \tag{3·18}$$

where ν is the number of the nearest neighbors, and

$$\gamma \equiv \frac{\hbar u w_0}{\hbar^2 w_0^2/2m^*}, \tag{3·19}$$

$$g \equiv \frac{D}{2\nu V}, \tag{3·20}$$

w_0 being the Debye cut-off wave number. Hereafter we shall call g the coupling constant. It is important to note that γ, the ratio of maximum phonon energy to the band width, is usually of the order of 1/100~1/10. We shall call it the non-adiabatic parameter, since the lattice vibration can be treated adiabatically in the limit of $\gamma \to 0$.

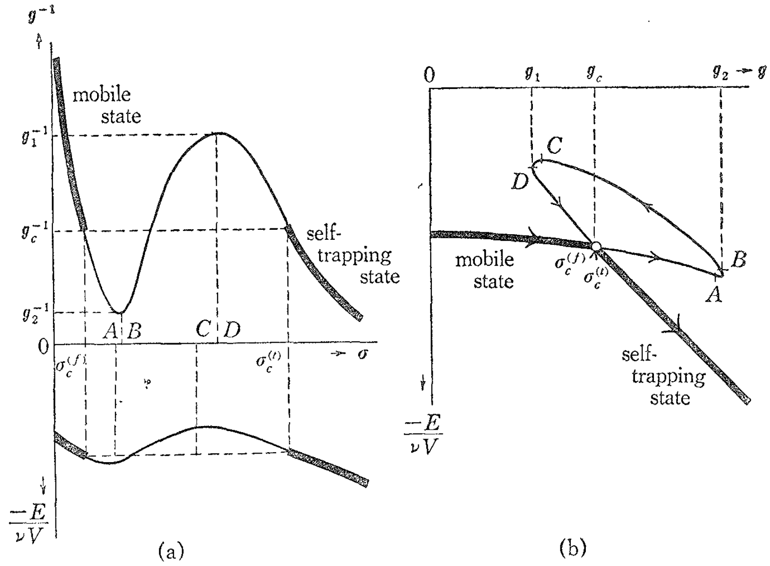
Instead of solving (3·17) with respect to σ as a function of g and then substituting the σ in (3·18), we shall use the parametric representation of g and E with σ as a parameter. In the limiting case of $\gamma\exp(\sigma) \ll 1$, we have

$$g^{-1} \sim \frac{3}{2}\gamma\,\frac{\exp(2\sigma)}{\sigma},$$

$$\frac{-E(0)}{\nu V} \sim (1+2\sigma)\exp(-\sigma),$$

while when $\gamma\exp(\sigma) \gg 1$, we have

$$g^{-1} \sim \frac{3}{4}\gamma^{-1}\frac{1}{\sigma},$$



(a) (b)

Fig. 3. Schematic representation of the solution of the variation problem. (a) : Parametric representation of the interaction constant g and the calculated energy E with σ as a parameter which is a measure of the magnitude of distortion. (b) : E as a function of g. The arrows on the curve are towards the direction of *increasing* σ.

$$\frac{-E(0)}{\nu V} \sim \frac{4}{3}\gamma\sigma.$$

Since $\gamma \ll 1$, we get $g^{-1}-\sigma$ and $E-\sigma$ curves as shown in Fig. 3a, from which we get the $E-g$ curve shown in Fig. 3b. It is important to note that σ is a three-valued function of g between g_1 and g_2. As the result, $E-g$ curve has a closed loop with a node at $g=g_c$, σ taking the value $\sigma_c^{(f)}$ and $\sigma_c^{(t)}$ on the two branches. Making use of the fact that $\gamma \ll 1$, we have

$$g_c \sim 1, \quad \sigma_c^{(f)} \sim \frac{3}{2}\gamma, \quad \sigma_c^{(t)} \sim \frac{3}{4}\gamma^{-1}, \tag{3·21}$$

$$g_1 \ll g_c \ll g_2. \tag{3·22}$$

As is evident from (3·7), σ represents the magnitude of lattice distortion around the electron. In fact, the expectation value of the number of phonons induced around the electron is given by

$$\bar{n} = \langle \Psi_k | \sum_w b_w^* b_w | \Psi_k \rangle = \sum_w | f_k(w) |^2, \tag{3·23}$$

which is of the same order of magnitude with σ's.

When the coupling constant g is smaller than g_c, the stable solution (the one with the lowest energy) corresponds to small $\sigma(<\sigma_c^{f)})$ while for $g>g_c$ it corresponds to large $\sigma(>\sigma_c^{(t)})$. There appears discontinuity at $g=g_c$, σ jumping from $\sigma_c^{(f)}(\ll 1)$ to $\sigma_c^{(t)}(\gg 1)$. It is evident that the former solution corresponds to the mobile state and the latter to the (plane wave of) self-trapping state. The second lowest solution for $g_1 < g < g_2$ represents a metastable state which is the self-trapping state or the mobile state according as $g < g_c$ or $g > g_c$. (See the discussion in § 2.) Whether the third solution has a definite physical meaning is not clear at present.

Let us calculate the effective mass of the self-trapped electron. Since $\exp(-\sigma) \ll 1$ in this case, the $\boldsymbol{k}$-dependent term of $\sigma_{\boldsymbol{k}}(\boldsymbol{\delta})$ in (3·11) is of the order of $\exp(-\sigma)$ in comparison with the main term $\sigma_0(\delta) = \sigma$, and we have, from (3·12),

$$E(\boldsymbol{k}) - E(0) = -V \exp(-\sigma) \widehat{\sum_{\delta}} \{\exp(i\boldsymbol{k}\cdot\boldsymbol{\delta}) - 1\} + 0(\exp(-2\sigma)), \quad (3\cdot24)$$

that is, the effective mass of the self-trapped electron is $\exp \sigma$ times the band effective mass. Since $\sigma^{(t)} \geqq \sigma_c^{(t)} = 3/4 \cdot \gamma^{-1}$, the ratio is an enormous quantity of the order of $e^{10} \sim e^{100}$. It is almost nonsense from the practical point of view to speak of the band formation of the self-trapped electron; it is practically immobile at absolute zero of temperature. The situation is quite different for the case of polaron, where neither the discontinuity of the effective mass (as the coupling increases) nor such a large distortion is expected to take place.

Finally, it is to be noted that the large discontinuity and the coexistence of essentially different states, is partly due to the smallness of non-adiabatic parameter γ. In fact, if γ is as large as $\gtrsim 1$, the $g^{-1} - \sigma$ curve of Fig. 3a becomes monotonous, and the discontinuity disappears. The situation is analogous to the problem of gas-liquid transition if we let temperatare correspond to γ.

§ 4. Discussions

We must now discuss the approximations in the preceding section. The first point is concerned with the Ansatz $\phi(\boldsymbol{n}) = \delta_{\boldsymbol{n},0}$. If we treat the case of polar mode in the same way as in § 3, with the same Ansatz, there appears a discontinuity of the same kind as was obtained there. This discontinuity seems to be illusory in view of the results obtained by many authors on the continuum dielectric model. That the Ansatz $\phi(\boldsymbol{n}) = \delta_{\boldsymbol{n},0}$ is inappropriate in the case of polar mode is evident from the discussions in § 2, where we found that there is one and only one optimum broadening of the electron wavefunction, which is usually much larger than the interatomic distance and can be determined independently of the atomic structure of the lattice.

On the other hand this Ansatz seems appropriate or at least tolerable in the case of the acoustical mode, because we found, in § 2, that in the self-trapping state the electron has a tendency to shrink as small as possible within the compatibility with the atomic structure of the lattice. Moreover, the results of § 3 in the weak-coupling case coincide with the results obtained by the conventional perturbation theory which is sufficiently good approximation in the case of acoustical mode.

Thus it seems rather probable that discontinuity obtained in § 3 is a genuine one in the case of acoustical mode, although the approximations used there lead to dangerous results in the case of polar mode. It might be desirable to relax the Ansatz $\phi(\boldsymbol{n})=\delta_{n,0}$, but in that case there appear many $\exp(-\sigma_k(\boldsymbol{l}))$ in the expression for the energy expectation value, and it is practically impossible to solve a number of simultaneous equations for $\sigma(\boldsymbol{l})$'s.

The second point is concerned with the Ansatz (3·4). This operator makes the origin of vibration displace; however, it does not include the effect of change of frequencies and the effects of higher order.

In the following paper we shall investigate the broadening of electron wave function and the corresponding lattice distortion and frequency change in the self-trapping state more thoroughly, at the cost of neglecting the effect of band formation which, in fact, has been shown to be quite small in § 3 of the present paper. Thereby we shall use the perturbation expansion in V (that is, in g^{-1}) and show that the convergence is fairly rapid, that is, the electronic wave function is in fact almost concentrated on the center atom even when g^{-1} is as large as a few times of g_c^{-1}. This partly justifies the Ansatz $\phi(\boldsymbol{n})=\delta_{n,0}$, and is also consistent with the nature of discontinuity discussed in § 3; that is, there exist two different types of well-defined states (one stable and the other metastable), and as g passes g_c, the only thing to take place is the reversal of energies of the two states, nothing discontinuous occurring in each state.

By the same method we shall also investigate the activated state through which the self-trapping state at some lattice point jumps to its nearest neighbor, and point out that it is when $g>g_c'(>g_c)$ that the electron moves through the lattice by the 'hopping motion',[8] while for intermediate coupling region: $g_c<g<g_c'$, the electron will rather make a 'Schub-type motion', that is, it is once released from the self-trapping state and moves a few lattice distances before it becomes again self-trapped, because the activated state mentioned above has a higher energy than that of the mobile state.

Finally we must answer the most important question: Is it possible for real crystals that the coupling constant is as large as $g_c(\sim 1)$? Assuming simple cubic lattice with nearest neighbor distance a, we have

$$Va^2=\frac{\hbar^2}{2m^*} \tag{4·1}$$

by (3·15), and (3·20) becomes

$$g=\frac{1}{6}\frac{E_1^2 m^*}{\hbar^2 Ca} \quad (C=Mu^2/a^3 \text{ is elastic constant})$$

$$=0.035\frac{E_1(\text{in eV})^2(m^*/m)}{C(\text{in }10^{12}\text{ dyne/cm}^2)a(\text{in Å})}. \tag{4·2}$$

Since E_1 is usually of the order of several electron volts, it is fairly possible in relatively soft crystal with narrow band that the electron or hole becomes self-trapped. Quantitatively, of course, the criterion may vary to some extent if we change the model (for instance, other lattice structures, the consideration of the overlap energy V' for the second neighbor, or of the non-diagonal element $\langle m|I|n\rangle$ of the interaction Hamiltonian which seems important for the Känzig center), or if we improve the approximations of § 3, as we shall do in the following paper.

Appendix

Consider the following two operators;

$$A=\alpha^*b-\alpha b^*,\quad A'=\alpha'^*b-\alpha'b^*$$

where b, b^* are phonon operators and α, α' are complex numbers. The expectation value

$$F(\lambda)=\langle 0|\exp(-\lambda A)\exp(\lambda A')|0\rangle_L$$

can be evaluated by differentiation, for it gives

$$F'(\lambda)=\{\alpha^*(\alpha'-\alpha)-\alpha'(\alpha'^*-\alpha^*)\}\lambda F(\lambda)$$

if one makes use of (3·5). After integration we have

$$F(\lambda)=\exp[\{\alpha^*(\alpha'-\alpha)-\alpha'(\alpha'^*-\alpha^*)\}\lambda^2/2]. \tag{A·1}$$

In the same way we have

$$\begin{aligned}\langle 0|\exp(-\lambda A)b\exp(\lambda A')|0\rangle&=-\lambda\alpha'F(\lambda)\\ \langle 0|\exp(-\lambda A)b^*\exp(\lambda A')|0\rangle&=-\lambda\alpha^*F(\lambda)\\ \langle 0|\exp(-\lambda \mathrm{A})b^*b\exp(\lambda A')|0\rangle&=\lambda^2\alpha^*\alpha'F(\lambda).\end{aligned} \tag{A·2}$$

(A·1) is made use of in deriving (3·6) and (3·7) while (A·2) is used in deriving (3·8).

References

1) L. Landau, Phys. Zeits. d. Sowjetunion **3** (1933), 664.
2) See, for instance, N. F. Mott & R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford, 1940), § 5.
3) H. Fröhlich, Advances in Physics **3** (1954), 325.
4) T. D. Lee, F. E. Low and D. Pines, Phys. Rev. **90** (1953), 297.
T. Yokota, Ann. Inst. Statist. Math. **5** (1954), 107.
M. Gurari, Phil. Mag. **44** (1953), 329.
5) R. P. Feynman, Phys. Rev. **97** (1955), 660.
6) T. D. Shultz, Tech. Rep. M. I. T. No. 9 (1956).
7) J. Markham and F. Seitz, Phys. Rev. **74** (1948), 1014.
8) J. Yamashita and T. Kurosawa, J. Phys. Chem. Solids **5** (1958), 34.
J. Yamashita and T. Kurosawa, J. Phys. Soc. Japan **15** (1960), 802.
T. Kurosawa, J. Phys. Soc. Japan **15** (1960), 1211.
G. L. Sewell, Phil. Mag. **36** (1958), 1361.
9) M. Onuki and H. Kawamura, J. Phys. Soc. Japan **14** (1959), 967.
10) K. Kobayashi and F. C. Brown, Phys. Rev. **113** (1959), 507
D. C. Burnham, F. C. Brown and R. S Knox, Phys. Rev. **119** (1960), 1560.
11) R. C. Hanson and F. C. Brown, J. Appl. Phys. **31** (1960), 210.
12) W. Känzig, Phys. Rev. **99** (1955), 1890.
T. G. Castner and W. Känzig, J. Phys. Chem. Solids **3** (1957), 178.
T. O. Woodruff & W. Känzig, J. Phys. Chem. Solids **5** (1958), 268.
13) J. Bardeen and W. Shockley, Phys. Rev. **80** (1950), 72.
14) L. I. Schiff, *Quantum Mechanics* (McGraw Hill, 1949), Chapt. IV.
15) S. Tomonaga, Prog. Theor. Phys. **2** (1947), 6.