AUGUST 1, 1937 PHYSICAL REVIEW VOLUME 52

The Structure of Electronic Excitation Levels in Insulating Crystals

GREGORY H. WANNIER
*Princeton University, Princeton, New Jersey**

(Received May 13, 1937)

In this article, a method is devised to study the energy spectrum for an excited electron configuration in an ideal crystal. The configuration studied consists of a single excited electron taken out of a full band of N electrons. The multiplicity of the state is N^2. It is shown that because of the Coulomb attraction between the electron and its hole $N^{8/5}$ states are split off from the bottom of the excited Bloch band; for these states the electron cannot escape its hole completely. The analogy of these levels to the spectrum of an atom or molecule is worked out quantitatively. The bottom of the Bloch band appears as "ionization potential" and the Bloch band itself as the continuum above this threshold energy.

FOR several years, there have been two competing pictures in use to describe the behavior of electrons in crystals. The one adopted in most theoretical calculations and especially successful for metals describes each electron by a running wave, but Frenkel has shown that in many cases the more elementary atomic picture may be the better approximation.[1] This apparent contradiction has been removed by Slater and Shockley,[2] who showed with a simplified model that the two types of states actually coexist in a crystal. It is the purpose of the present paper to treat this question in a quantitative way, starting out from the actual Hamiltonian of the system.

We shall restrict ourselves in this article to insulators containing one electron in the lowest excited state, and we shall study the energy spectrum of this single configuration, neglecting perturbations arising from other configurations. As to the method we shall proceed in the following way:

(1) We shall construct orthogonal "atomic" wave functions and express the energy matrix in this vector system.

(2) The energy matrix contains many terms having the periodicity of the lattice and a few which have not; we shall develop a method which takes them both into account.

(3) We shall derive some general results and discuss their consequences.

1. BASIC WAVE FUNCTIONS AND ENERGY MATRIX

It would no doubt be more satisfactory for insulating crystals, to discuss the Hamiltonian using atomic functions rather than Bloch functions. But this line of attack has been hampered by the fact that atomic functions are not orthogonal. We can, however, build up orthogonal functions having all the advantages of atomic ones by starting out from a Bloch approximation. Let us assume then that a Bloch or Fock method has given us functions $b(\mathbf{k}, \mathbf{x})$ of energy $W(\mathbf{k})$. Then the required functions are

$$a(\mathbf{x}-\mathbf{n}) = 1/(N)^{\frac{1}{2}} \cdot \sum_{\nu} \exp\left[-i\mathbf{k}_\nu\mathbf{n}\right] b(\mathbf{k}_\nu, \mathbf{x}),^{3} \quad (1)$$

where N is the number of cells in the crystal and the $\mathbf{k}$'s are as usual determined by some boundary condition.

Formula (1) applies to any set of Bloch functions, but it might be interesting to get some insight into the structure of the a's. For this purpose let us first make the *ad hoc* assumption (valid for free electrons) that b is of the form

$$b(\mathbf{k}_\nu, \mathbf{x}) = \exp\left[i\mathbf{k}_\nu\mathbf{x}\right] \cdot b(\mathbf{x}),$$

where the periodic factor $b(\mathbf{x})$ is independent of $\mathbf{k}$. Then we find explicitly:

---

* I want to express my thanks to Princeton University for the grant of its Swiss-American Exchange Fellowship for the year 1936–37.

[1] J. Frenkel, Phys. Rev. **17**, 17 (1931); Physik. Zeits. Sowjetunion **9**, 158 (1936); Physik. Zeits. Sowjetunion **8**, 185 (1935).

[2] J. C. Slater and W. S. Shockley, Phys. Rev. **50**, 705 (1936).

[3] The unit of length adopted in this article is the elementary translation in the direction of each of the crystal axes. In some deductions the crystal is assumed to be simple cubic, but this could easily be removed.

$$a(\mathbf{x}-\mathbf{n})=\frac{\sin\pi(x_1-n_1)\sin\pi(x_2-n_2)\sin\pi(x_3-n_3)}{\pi^3(x_1-n_1)(x_2-n_2)(x_3-n_3)}b(\mathbf{x}),$$

which gives us the desired concentration around n_1, n_2, n_3. Their orthogonality is immediately seen because

$$\int_{-\infty}^{+\infty}\frac{\sin\pi x}{x}\frac{\sin\pi(x+n)}{x+n}|b(x)|^2dx$$
$$=(-)^n/n\cdot\int_{-\infty}^{+\infty}\sin^2\pi x|b(x)|^2\left(\frac{1}{x}-\frac{1}{x+n}\right)dx.$$
$$=0.$$

These new functions provide us also with a method to orthogonalize standard atomic functions without doing very much harm to them. For this purpose let us take atomic functions $\psi(\mathbf{x}-\mathbf{m})$ and build up from them running waves:

$$b(\mathbf{k},\mathbf{x})=A(\mathbf{k})/(N)^{\frac{1}{2}}\sum_m\exp[i\mathbf{km}]\cdot\psi(\mathbf{x}-\mathbf{m}).\qquad(2)$$

The formula without A gives orthogonal b's, but they are not normalized, if the ψ's are not orthogonal. We find for $1/A^2$:

$$1/A^2(\mathbf{k})=\sum_n\exp[i\mathbf{kn}]\cdot\int\psi^*(\mathbf{x})\psi(\mathbf{x}-\mathbf{n})d\tau.\qquad(3\mathrm{a})$$

From this we can find in principle the value of $A(\mathbf{k})$:

$$A(\mathbf{k})=\sum_n\exp[-i\mathbf{kn}]\cdot\mathrm{a}(\mathbf{n}),\qquad(3\mathrm{b})$$

where the a's depend on the integrals in (3a). Putting (3b) and (2) into (1), we find

$$a(\mathbf{x}-\mathbf{m})=\sum_n\mathrm{a}(\mathbf{n}-\mathbf{m})\psi(\mathbf{x}-\mathbf{n}),\qquad(4)$$

which is the required result. If, in particular, the integrals are small and fall off slowly, then we can write

$$a(\mathbf{x}-\mathbf{m})=\psi(\mathbf{x}-\mathbf{m})$$
$$-\tfrac{1}{2}\sum_{n\neq m}\psi(\mathbf{x}-\mathbf{n})\int\psi^*(\mathbf{x}-\mathbf{n})\psi(\mathbf{x}-\mathbf{m})d\tau;$$

but if they fall off rapidly, the integral with nearest neighbors determines the coefficient of all ψ's.

As an example let us carry out the orthogonalization (4) explicitly for a set of atomic functions whose integrals are

$$\int\psi^*(\mathbf{x})\psi(\mathbf{x}-\mathbf{n})d\tau$$
$$=\exp[-\beta_1|n_1|-\beta_2|n_2|-\beta_3|n_3|].$$

Then we get

$$A(\mathbf{k})=C(\beta_1,k_1)\cdot C(\beta_2,k_2)\cdot C(\beta_3,k_3),$$

where

$$C(\beta,k)=\left(\frac{1-2e^{-\beta}\cos k+e^{-2\beta}}{1-e^{-2\beta}}\right)^{\frac{1}{2}},$$

from which it follows that the expansion coefficients (4) are

$$\mathrm{a}(\mathbf{n})=c(\beta_1,|n_1|)\cdot c(\beta_2,|n_2|)\cdot c(\beta_3,|n_3|),$$

where

$$c(\beta,0)=1/(1-e^{-2\beta})^{\frac{1}{2}}F(-\tfrac{1}{2},-\tfrac{1}{2},1,e^{-2\beta}),$$

$$c(\beta,n)=-1/(1-e^{-2\beta})^{\frac{1}{2}}\frac{1}{2n-1}\frac{(2n-1)!!}{(2n)!!}$$
$$\times e^{-n\beta}F(-\tfrac{1}{2},n-\tfrac{1}{2},n+1,e^{-2\beta}).$$

By $m!!$ is meant $m(m-2)(m-4)\cdots$ down to 1 or 2. F is the hypergeometric function

$$F(a,b,c,x)=1+\frac{ab}{1c}x+\frac{a(a+1)b(b+1)}{12c(c+1)}x^2+\cdots.$$

The F's used here are very slowly varying functions. When x increases from 0 to 1, $F(-\frac{1}{2},-\frac{1}{2},1,x)$ increases monotonely from 1 to $4/\pi=1.273\cdots$ and $F(-\frac{1}{2},n-\frac{1}{2},n+1,x)$ decreases monotonely from 1 to

$$[(2n)!!/(2n+1)!!](4/\pi).$$

In most cases one may replace them by 1.

These two examples should show that the new functions defined by (1) go over into atomic ones for large separation and are their natural substitute at smaller distances.

The building up of many electron wave functions from this basis is discussed elsewhere in full detail (e.g. reference 2, appendix). There is a unique wave function for a full band, but for a state having one excited electron and a hole in its full band, three basic types can be constructed:

(a) The atomic wave functions $A(\mathbf{m}, \mathbf{n})$ built up from our "atomic" one-electron functions $a_{\text{ground}}(\mathbf{x}-\mathbf{s})$ and $a_{\text{exc}}(\mathbf{x}-\mathbf{s})$. They contain no electron in the mth cell and an extra excited electron in the nth cell.

(b) The Bloch wave functions $B(\mathbf{k}_p, \mathbf{k}_q)$ built up from one-electron Bloch functions. They contain no electron of wave vector $\mathbf{k}_p$ and an excited electron of wave vector $\mathbf{k}_q$. Between the two there exists the relation:

$$\begin{aligned} &B(\mathbf{k}_p, \mathbf{k}_q) \\ &\quad = 1/N \cdot \sum_n \sum_m \exp\left[i\mathbf{k}_q\mathbf{n} - i\mathbf{k}_p\mathbf{m}\right] \cdot A(\mathbf{m}, \mathbf{n}), \\ &A(\mathbf{m}, \mathbf{n}) \\ &\quad = 1/N \cdot \sum_p \sum_q \exp\left[-i\mathbf{k}_q\mathbf{n} + i\mathbf{k}_p\mathbf{m}\right] \cdot B(\mathbf{k}_p, \mathbf{k}_q). \end{aligned} \tag{5}$$

(c) The excitation waves $E(\mathbf{K}_\alpha, \boldsymbol{\beta})$ introduced by Frenkel.[1] They describe like Bloch functions the motion of the center of gravity of electron and hole by a wave vector $\mathbf{K}_\alpha$; but as in (a), the two bodies have no relative motion and their distance is given by a space vector $\boldsymbol{\beta}$. Their relation to A and B is

$$\begin{aligned} &E(\mathbf{K}_\alpha, \boldsymbol{\beta}) \\ &\quad = 1/(N)^{\frac{1}{2}} \cdot \sum_n \exp\left[i\mathbf{K}_\alpha\mathbf{n}\right] \cdot A(\mathbf{n} - \tfrac{1}{2}\boldsymbol{\beta}, \mathbf{n} + \tfrac{1}{2}\boldsymbol{\beta}), \end{aligned} \tag{6}$$

$$\begin{aligned} &B(\mathbf{k}_p, \mathbf{k}_q) \\ &= 1/(N)^{\frac{1}{2}} \cdot \sum_\beta \exp\left[i\tfrac{1}{2}(\mathbf{k}_p + \mathbf{k}_q)\boldsymbol{\beta}\right] \cdot E(\mathbf{k}_q - \mathbf{k}_p, \boldsymbol{\beta}). \end{aligned} \tag{7}$$

The energy matrix is most easily calculated in terms of the A's. We start out from the Hamiltonian

$$H = \sum_i T_i + \sum_{i<k} \frac{e^2}{d \cdot |\mathbf{x}_i - \mathbf{x}_k|} - e \cdot \sum_{i,\, m} V(\mathbf{x}_i - \mathbf{m}) \tag{8}$$

$$+\text{terms independent of } \mathbf{x}_i,$$

where T_i is the kinetic energy of the ith electron, d the length of the elementary crystal-translation in cm, and $\mathbf{x}_i$ the coordinates of the ith electron. The energy is written here in such a form, that only the electrons of the band under consideration are treated as such; the influence of the others is only taken into account through the screened potential $V(\mathbf{x}-\mathbf{m})$ arising from the mth cell.

If one works out the energy matrix $(\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}')$ one finds a very great number of terms, which depend on $\mathbf{m}, \mathbf{n}, \mathbf{m}', \mathbf{n}'$ only through the combination $\mathbf{m}-\mathbf{m}'$ and $\mathbf{n}-\mathbf{n}'$. The physical reason for this fact is that a transition of the excited electron (or the hole among the unexcited ones) from one cell to another takes place in a field which is very nearly periodic, i.e., the transition $\mathbf{n} \rightarrow \mathbf{n}'$ and $\mathbf{n}+\mathbf{s} \rightarrow \mathbf{n}'+\mathbf{s}$ must have the same matrix element. In addition, one finds for those periodic terms that if both vectors $\mathbf{m}-\mathbf{m}'$ and $\mathbf{n}-\mathbf{n}'$ are different from zero, the matrix element vanishes; one can therefore distinguish between transitions of the excited electron and transitions of its hole. There remain, however, two integrals which do not follow these two rules; if we write them out explicitly H takes the form:

$$\begin{aligned} (\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}') &= \delta_{mm'} H_{\text{exc}}(\mathbf{n}-\mathbf{n}') \\ &- \delta_{nn'} H_g(\mathbf{m}-\mathbf{m}') - \int\!\!\int a_g^*(\mathbf{x}-\mathbf{m}) a_e^*(\mathbf{x}'-\mathbf{n}) \\ &\times \frac{e^2}{d \cdot |\mathbf{x}-\mathbf{x}'|} a_g(\mathbf{x}-\mathbf{m}') a_e(\mathbf{x}'-\mathbf{n}') d\tau d\tau' \\ &+ \int\!\!\int a_g^*(\mathbf{x}-\mathbf{m}) a_e^*(\mathbf{x}'-\mathbf{n}) \frac{e^2}{d \cdot |\mathbf{x}-\mathbf{x}'|} \\ &\times a_g(\mathbf{x}'-\mathbf{m}') a_e(\mathbf{x}-\mathbf{n}') d\tau d\tau'. \end{aligned} \tag{9}$$

These integrals evidently come from the Coulomb attraction between the electron and its hole; they depend on the relative position of the two particles.

There is, however, no term depending on their absolute position in space. We can separate out therefore the motion of their center of gravity and go over to excitation waves.

2. Discussion and Approximate Solution of the Matrix Equation

There are two limiting cases in which (9) is diagonalized immediately. The one is the "infinite separation," where only diagonal terms remain appreciable and either atomic wave functions or excitation waves may be used. The second, which we may call "negligible correlation," occurs when the two integrals can be omitted; then Bloch functions are correct.

This last knowledge can be exploited further, because it means that a Bloch approximation can take care of all terms H^+ in H, except the two integrals. Let us assume in the future that such a calculation has been carried out and that we know the energy as a function of the wave vector for both bands:

$$
\begin{aligned}
&(\mathbf{k}_p, \mathbf{k}_q | H^+ | \mathbf{k}_{p'}, \mathbf{k}_{q'}) \\
&\qquad = \delta_{pp'}\delta_{qq'}[W_e(\mathbf{k}_q) - W_g(\mathbf{k}_p)], \qquad (10)
\end{aligned}
$$

and therefore, transforming with (5) and comparing with (9):

$$
\begin{aligned}
H_e(\mathbf{n}-\mathbf{n}') &= 1/N\cdot\sum_q \exp[i\mathbf{k}_q(\mathbf{n}-\mathbf{n}')]W_e(\mathbf{k}_q), \\
H_g(\mathbf{m}-\mathbf{m}') & \\
&= 1/N\cdot\sum_p \exp[-i\mathbf{k}_p(\mathbf{m}-\mathbf{m}')]W_g(\mathbf{k}_p).
\end{aligned}
$$

Equation (9) is now identical with Eq. (21) in Slater's paper,[4] but looks somewhat simpler, because the spin of the electron is assumed to be unchanged in all states involved, and the last term in his expression arising from the change in the self-consistent field is absorbed in W_e and W_g, but a further transformation is in some cases very useful. Let H_e operate on an arbitrary function $F(\mathbf{n})$ satisfying our boundary conditions:

$$F(\mathbf{n}) = \sum_\kappa \exp[i\boldsymbol{\kappa}\mathbf{n}]G(\boldsymbol{\kappa}).$$

We then get

$$
\begin{aligned}
&\sum_{n'} H_e(\mathbf{n}-\mathbf{n}')F(\mathbf{n}') \\
&= 1/N\cdot\sum_q\sum_\kappa\sum_{n'} \exp[i\mathbf{k}_q\mathbf{n}+\mathbf{n}'(\boldsymbol{\kappa}-\mathbf{k}_q)]W_e(\mathbf{k}_q)G(\boldsymbol{\kappa}) \\
&= \sum_\kappa W_e(\boldsymbol{\kappa}) \exp[i\boldsymbol{\kappa}\mathbf{n}]G(\boldsymbol{\kappa}) \\
&\qquad\qquad = W_e\left(-i\frac{\partial}{\partial n_1}, -i\frac{\partial}{\partial n_2}, -i\frac{\partial}{\partial n_3}\right)\cdot F(\mathbf{n}).
\end{aligned}
$$

This gives for (9)

$$
\begin{aligned}
(\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}') = \delta_{mm'}\delta_{nn'}\left[W_e\left(-i\frac{\partial}{\partial\mathbf{n}}\right) - W_g\left(-i\frac{\partial}{\partial\mathbf{m}}\right)\right] - \int\int a_g^*(\mathbf{x}-\mathbf{m})a_e^*(\mathbf{x}'-\mathbf{n})\frac{e^2}{d\cdot|\mathbf{x}-\mathbf{x}'|} \\
\times a_g(\mathbf{x}-\mathbf{m}')a_e(\mathbf{x}'-\mathbf{n}')d\tau d\tau' + \int\int a_g^*(\mathbf{x}-\mathbf{m})a_e^*(\mathbf{x}'-\mathbf{n})\frac{e^2}{d\cdot|\mathbf{x}-\mathbf{x}'|}a_g(\mathbf{x}'-\mathbf{m}')a_e(\mathbf{x}-\mathbf{n}')d\tau d\tau'. \qquad (9a)
\end{aligned}
$$

In accordance with the remark at the end of part 1, we now go over to excitation waves as basic wave functions and write our solution in the form

$$\psi = \sum_\beta U_\alpha(\mathfrak{b})E(\mathbf{K}_\alpha, \mathfrak{b}). \qquad (11)$$

Then from (6) and (9a), $U(\mathfrak{b})$ has to satisfy the equation

$$
\begin{aligned}
\Big\{&\exp[-i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{b}]\cdot W_e\left(-i\frac{\partial}{\partial\mathfrak{b}}\right)\cdot\exp[+i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{b}] \\
&-\exp[+i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{b}]\cdot W_g\left(-i\frac{\partial}{\partial\mathfrak{b}}\right)\cdot\exp[-i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{b}] \\
&\qquad\qquad +\phi(\mathfrak{b})\Big\}U(\mathfrak{b}) = E\cdot U(\mathfrak{b}). \qquad (12)
\end{aligned}
$$

$\phi(\mathfrak{b})$ is a potential which is very much like $-e^2/(d\cdot|\mathfrak{b}|)$ for large values of $|\mathfrak{b}|$, but has no singularity at $\mathfrak{b}=0$. For simplicity the non-diagonal Coulomb terms have been omitted, for they do not influence the discussion given below and they are so small that they may be neglected in an approximate solution.

The transformation of difference operators into differential operators, as has been done in our transformation from (9) to (9a) and (12), is not much favored by mathematicians. There exists however a book by Davis[5] in which this point of view is adopted, although it is developed for one independent variable only. With this restriction, the theorems on pp. 100–105 of his book ensure that our operator has a sense if $W(\mathbf{k})$ is an entire function of its argument, i.e. defined and regular for complex values of $\mathbf{k}$. Its equivalence with a difference operator may then be shown from the operational identity

---

[4] J. C. Slater, this issue.

[5] H. D. Davis, *The Theory of Linear Operators* (Bloomington, Ind. 1936).

$$\exp[-i\mathbf{k}\mathfrak{b}]\cdot W\left(-i\frac{\partial}{\partial\mathfrak{b}}\right)\exp[+i\mathbf{k}\mathfrak{b}] = W\left(\mathbf{k}-i\frac{\partial}{\partial\mathfrak{b}}\right), \qquad (13)$$

which, because of the periodicity of W, gives

$$W\left(-i\frac{\partial}{\partial\mathfrak{b}}\right)\exp[2\pi i\beta_\nu] - \exp[2\pi i\beta_\nu]W\left(-i\frac{\partial}{\partial\mathfrak{b}}\right)=0, \qquad (13a)$$

i.e., a solution U may be multiplied with an arbitrary function having the periodicity of the lattice, and still remains a solution.

Before using (12) for an approximation method, let us first find some general properties of the U's and their energy spectrum.

For large values of $\mathfrak{b}$, all the Coulomb integrals and therefore ϕ vanish and the solution must then have exponential character with some complex wave vector $\mathbf{k}=\kappa_1+i\kappa_2$. Remembering (13) we get the corresponding energy from (12):

$$E=W_e(\tfrac{1}{2}\mathbf{K}_\alpha+\kappa_1+i\kappa_2)-W_g(\tfrac{1}{2}\mathbf{K}_\alpha-\kappa_1-i\kappa_2). \qquad (14)$$

κ_1 and κ_2 are restricted by the condition that E be real. Now we know in advance that E is real, periodic and bounded for $\kappa_2=0$. Therefore, its gradient vanishes at least at two points within one period and we can conclude from general mapping properties that lines of possible complex values of $\mathbf{k}$ start out from these points. The corresponding energies go upwards from the maxima, and downwards from the minima. Fig. 1 is a simple one-dimensional illustration of this situation. W_e-W_g is taken to be $1-\cos k$. The possible complex values of k are then $n\pi+i\kappa$ and their energy is $-\cosh\kappa+1$, if n is even, and $+\cosh\kappa+1$ if n is odd. However, two restrictions reduce the number of these new levels considerably. First of all, the Coulomb field is an attractive field and it can never raise any level; thus there will be no state above the Bloch band. Secondly, a solution of (12) with complex wave vector will in general be unbounded and the condition of boundedness coincides with the condition that $U(\mathfrak{b})$ vanish exponentially at large distances. This restriction will, for each $\mathbf{K}_\alpha$, select a number of discrete levels out of all possible ones.

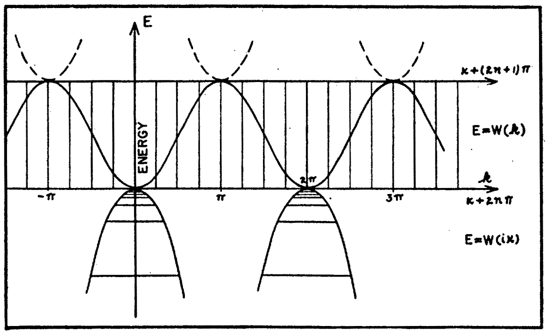

FIG. 1. Energy and wave vector for an excited electron in a crystal.

If we want more detailed information in practical cases we may turn to an atomic picture for the lowest discrete states[6] and to a Bloch picture for the continuum. But for the bulk of the discrete states and the lower continuum there exists a better alternative, namely a local exploration of the Brillouin zone, i.e., if we are interested in a particular region in the Brillouin zone we develop the function $W(\mathbf{k})$ around that point into a power series and take the first two terms or so. The point to be taken in our case is, of course, the bottom of the band where the continuum and the discrete spectrum meet. For practical reasons, let us also assume that the crystal is symmetric enough to allow for a spherically symmetric development of the form

$$\begin{aligned} W_e(\mathbf{k})&=w_e+A_e(k_1^2+k_2^2+k_3^2),\\ W_g(\mathbf{k})&=w_g+A_g(k_1^2+k_2^2+k_3^2). \end{aligned} \qquad (15)$$

Assuming $A_e-A_g>0$,[7] approximating provisionally $\phi(\mathfrak{b})$ by $e^2/d\cdot|\mathfrak{b}|$, and making the substitution

$$U(\mathfrak{b})=\exp\left[\frac{A_e+A_g}{A_e-A_g}i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{b}\right]\cdot u(\mathfrak{b}) \qquad (16)$$

we find for u the differential equation

$$\Delta u(\mathfrak{b})+\left(\mathcal{E}+\frac{2\gamma}{|\mathfrak{b}|}\right)u(\mathfrak{b})=0 \qquad (17)$$

with $\qquad \gamma=e^2/(2d(A_e-A_g)) \qquad (18)$

---

[6] R. Peierls, Ann. d. Physik **13**, 905 (1932). A. v. Hippel, Zeits. f. Physik **101**, 680 (1936).

[7] If this condition is reversed, the minimum energy in the Brillouin zone does not lie at (0,0,0), but e. g. at (π,π,π) and one has to study $W(\boldsymbol{\pi}-i\partial/\partial\mathfrak{b})$. From (13) we see that this gives simply an extra factor -1 for the wave function U, if we pass from one cell to the next.

and $E = w_e - w_g - \mathbf{K}_\alpha{}^2 (A_e A_g)/(A_e - A_g)$

$$+(A_e - A_g)\mathcal{E}. \quad (19)$$

Equation (17) is the Schrödinger equation for the hydrogen atom. It admits for $\mathcal{E}$ all positive eigenvalues and the negative eigenvalues $-\gamma^2/n^2$. One easily verifies that the positive eigenvalues coincide with the Bloch band, while the levels

$$E_n(\mathbf{K}_\alpha) = W_0(\mathbf{K}_\alpha) - \frac{e^4}{4d^2(A_e - A_g)} \frac{1}{n^2} \quad (20)$$

give us the discrete spectrum. The multiplicity of the states belonging to E_n is n^2 as in atomic spectra.

The expressions (19) and (20) are roughly correct as long as approximation (15) holds, i.e., as long as the distance of the level from W_0 is small compared with the width of the whole Bloch band (Fig. 1). Secondly, no negative energy level can lie below the excitation energy of the electron within its own atom; the distance of E from W_0 must therefore also be small compared with the distance of this level. But if one wants to complete these calculations, by a perturbation method, one must remove this second restriction and cut off the Coulomb potential at the origin in a suitable way (Fig. 2). In this case one can at least make the first restriction somewhat less rigid, for one can show that the first order perturbation of the infinite differential operator W exists. It is to be hoped that the perturbation method itself will also converge.

Let us now proceed to establish the chief property of these discrete states, namely that they carry no current.

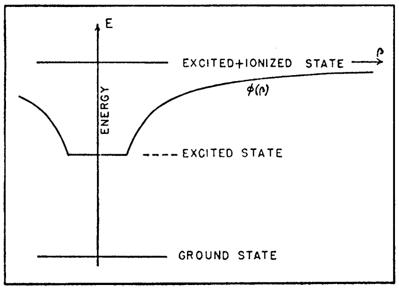

FIG. 2. Potential energy of electron and hole as a function of their distance.

The current operator in our system is obtained by transforming it from its expression in a Bloch system:

$$(\mathbf{k}_p, \mathbf{k}_q | \mathbf{J} | \mathbf{k}_p', \mathbf{k}_q') = -e/\hbar\, \delta_{pp'} \delta_{qq'} [\nabla W_e(\mathbf{k}_q) - \nabla W_g(\mathbf{k}_p)].$$

The same argument that led from (10) to (9a) and (12) gives now

$$\mathbf{J} U(\boldsymbol{\beta}) = -e/\hbar \left\{ \exp\left[-i\tfrac{1}{2}\mathbf{K}_\alpha \boldsymbol{\beta}\right] \cdot \nabla W_e\left(-i\frac{\partial}{\partial \boldsymbol{\beta}}\right) \cdot \exp\left[+i\tfrac{1}{2}\mathbf{K}_\alpha \boldsymbol{\beta}\right] - \exp\left[+i\tfrac{1}{2}\mathbf{K}_\alpha \boldsymbol{\beta}\right] \cdot \nabla W_g\left(-i\frac{\partial}{\partial \boldsymbol{\beta}}\right) \cdot \exp\left[-i\tfrac{1}{2}\mathbf{K}_\alpha \boldsymbol{\beta}\right] \right\} \times U(\boldsymbol{\beta}). \quad (21)$$

It can easily be shown that apart from surface integrals which vanish if they are extended over the whole crystal, $\int_\Omega U^*(\boldsymbol{\beta}) \mathbf{J} U(\boldsymbol{\beta}) d\tau_\beta$ is equal to another volume integral which we denote by $\int_\Omega \{U^* \mathbf{J} U\} d\tau_\beta$. One can easily give rules how to build it up from a given $\mathbf{J}$, but they are somewhat lengthy. In this connection, it is sufficient to know that this new expression satisfies a Green's theorem

$$-e\int_V (U^*(\boldsymbol{\beta}) W U(\boldsymbol{\beta}) - U(\boldsymbol{\beta}) W U^*(\boldsymbol{\beta})) d\tau_\beta = i\hbar \int_S \{U^* \mathbf{J} U\} d\mathbf{s}_\beta,$$

where W stands for the whole differential operator in (12) and the surface integral on the right-hand side is extended over the boundary of the volume on the left. This gives

$$(-e)(d/dt)\int_V U^*(\boldsymbol{\beta}) U(\boldsymbol{\beta}) d\tau_\beta + \int_S \{U^* \mathbf{J} U\} d\mathbf{s}_\beta = 0.$$

As in the case of atoms this law is sufficient to prove that if U vanishes exponentially at infinity, then

$$\int_\Omega U^*(\boldsymbol{\beta}) \mathbf{J} U(\boldsymbol{\beta}) d\tau_\beta = 0$$

if the integral is extended over the whole crystal. The true expression is however $\sum_\beta U^*(\boldsymbol{\beta}) \mathbf{J} U(\boldsymbol{\beta})$; but one can dispose of the arbitrary periodic factor of U in such a way that

$$\sum_\beta U^*(\boldsymbol{\beta}) U(\boldsymbol{\beta}) = \int_\Omega U^*(\boldsymbol{\beta}) U(\boldsymbol{\beta}) d\tau_\beta$$

and $\left|\sum_\beta U^*(\boldsymbol{\beta}) \mathbf{J} U(\boldsymbol{\beta})\right| \leqq \left|\int_\Omega U^*(\boldsymbol{\beta}) \mathbf{J} U(\boldsymbol{\beta}) d\tau_\beta\right| + \epsilon,$

which proves the theorem.

3. General Survey of the Results

The results of the above calculations suggest a rather fundamental change in the picture used to describe the behavior of electrons in crystals.

Since there has been so much talk about discrete levels, one should perhaps first remember that there are no discrete absorption lines to be expected in any crystalline spectrum, because, as in molecules, there is a band structure connected with each electronic energy level. The value of this term in the approximation (19) is

$$-\frac{A_eA_g\mathbf{K}_\alpha^2}{(A_e-A_g)}.$$

The analogy of $\mathbf{K}_\alpha$ to the vibrational quantum number of a molecular band spectrum becomes evident, if we remember Frenkel's selection rule.[1] It states that if a light wave of wave vector $\boldsymbol{\kappa}$ excites a state belonging to the excitation waves of wave vector $\mathbf{K}_\alpha$ the difference $\boldsymbol{\kappa}-\mathbf{K}_\alpha$ must appear as wave vector of a simultaneously absorbed or emitted lattice wave. Since for visible light $\boldsymbol{\kappa}$ is very small, $\mathbf{K}_\alpha$ may be considered as the wave number of this lattice vibration.

The truly electronic spectrum underlying the band spectrum is obtained by putting $\mathbf{K}_\alpha=0$ in the energy expression (20). It is divided into two parts:

First there comes a lower part consisting of discrete states. The electron does not escape its hole and no photocurrent can be observed. The very lowest states are widely spaced and correspond to excitation of an electron within its cell or to some direct neighbor.[6] As one goes higher up this individual character gets lost; the spacing gets narrower and the electron moves in an orbit determined by the Coulomb field of its hole.[8] For a finite crystal with N cells, the number of such states is of the order $N^{3/5}$.

Then there follows a continuum which is just the Bloch band. Electron and hole move independently and a current may be observed. But in these states the electron cannot well approach its hole and optical transitions into them are therefore unlikely. For our finite crystal, the number of such states is of the order N.

As far as can be told, experiment is in agreement with these views.[9]

In the practical application of these results one should keep in mind some serious restrictions which underlie the calculations.

This paper deals with the excited states of a single electron only. It will still be good for the excitation of a number of electrons small compared with the number of cells, but after this qualitative changes may occur; the same is true for conducting crystals.

Secondly these calculations deal with virtual excited states of an unexcited crystal rather than with truly excited states. And the virtual states considered are those compatible with the Frank-Condon principle. Each electronic excitation will however be followed by rearrangements of the lattice and energy dissipation. Many authors believe that this will result in an actual "trapping" of the electron in its new position.[1, 5]

One should think that these facts have no influence on the absorption spectrum of a crystal, but it has been pointed out to me by Morse that, even for this case, the list of levels given above may be incomplete, because there is a possibility of weak absorptions violating the Frank-Condon principle. This is especially true for regions where no other absorption is possible. It is perhaps the explanation for the photoelectric absorption of many crystals on the long wave-length foot of their ultraviolet band.

Finally I want to express my thanks to Professor E. Wigner, who has suggested to me the present problem, and to Professor J. C. Slater, Dr. C. Herring, Dr. F. Smithies and many others for valuable advice. I also want to thank Mr. R. C. Herman for his help in completing this paper.

---

[8] This distinction is well known in the theory of electrolytes. Our lowest levels correspond to the classical theory of weak electrolytes, the higher ones to the theory of Debye-Hückel. This analogy may perhaps be extended if one considers the motion of "trapped" electrons.

[9] R. Hilsch and R. W. Pohl, Zeits. f. Physik **59**, 812 (1930); E. G. Schneider and H. M. O'Bryan, Phys. Rev. **51**, 293 (1937); B. Gudden and R. Pohl, Zeits. f. Physik **17**, 331 (1923); F. C. Nix, Rev. Mod. Phys. **4**, 723 (1932); A. L. Hughes, Rev. Mod. Phys. **8**, 294 (1936).