BEHAVIOR OF NON-METALLIC CRYSTALS IN STRONG ELECTRIC FIELDS

L. V. KELDYSH

P. N. Lebedev Physics Institute, Academy of Sciences, U.S.S.R.

Submitted to JETP editor April 30, 1957

J. Exptl. Theoret. Phys. (U.S.S.R.) **33**, 994-1003 (October, 1957)

An expression is obtained for the number of electron-hole pairs generated in a semiconductor by a uniform electric field. The derivation is made for an arbitrary crystal. The result differs from the usually-employed formula in that it contains an explicit angular dependence and a slightly different dependence on the field. A particularly essential fact is that in the absence of electron-phonon collisions and collisions between electrons themselves the magnitude of the effective potential barrier is determined not by the width of the forbidden band, but by the lower edge of optical absorption (the internal photoeffect), which, as a rule, is considerably greater. This circumstance should lead to an essential increase in the critical fields.

THE presence of a strong electrical field ($E \sim 10^5$ v/cm), as is known, produces in semiconductors additional carriers, the number of which increases sharply with increasing field. The most probable mechanisms causing this fact are, first, shock ionization and second, direct knocking out of valence electrons by the field in the conduction band. This latter mechanism, analogous in a certain sense to cold electron emission from a metal surface, was first considered by Zener[1] in the quasi-classical approximation, which is natural for such a problem. The best expressions obtainable by this method for the number of electrons n passing into the conduction band per unit volume per unit time, is given apparently in the work by McAfee et al.[2] and has the form

$$n = N\frac{eEd}{2\pi\hbar}\exp\left\{-\frac{\pi}{2e\hbar E}\sqrt{2m^*}\Delta^{3/2}\right\}, \tag{1}$$

where N is the number of valence electrons per unit volume, d the crystal lattice period, m^* the effective mass of the electron, Δ the width of the forbidden zone, and e the electron charge.

This formula is somewhat indefinite, since the values of the effective mass of the electron in the valence band and in the conduction band are in general different and it is not clear exactly which value is contained in the exponent. In addition, expression (1) was obtained by solving the unidimensional problem. As will be shown below, the correct allowance for the three-dimensionality leads to certain quite substantial qualitative changes. Finally, a particularly important point, no account is taken in the derivation of Eq. (1) of the scattering of electrons by thermal lattice vibrations, which, as will be shown below, is of decisive significance for this problem.

We shall calculate in this work the probability of the passage of a valence electron into the conduction band by a method already used for this purpose by Houston[3] and representing essentially a method commonly used in perturbation theory to calculate the transition probability per unit time. The entire analysis will be carried out in the so-called single-electron approximation, i.e., the interaction between electrons will be disregarded, with the exception of that portion of the interaction included in the general self-consistent field of the crystal.

The Hamiltonian of the system has in this case the following form:

$$\hat{H} = \hat{H}_{0e} + \hat{H}_{0L} + \hat{H}_{eL} + e\mathbf{E}\mathbf{r}; \tag{2}$$

$$\hat{H}_{0e} = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla\right)^2 + W(\mathbf{r});\ \hat{H}_{0e}\psi_{0j}(\mathbf{p},\mathbf{r}) = \varepsilon_j(\mathbf{p})\psi_{0j}(\mathbf{p},\mathbf{r}); \tag{3}$$

$$\hat{H}_{0L} = \frac{1}{2}\sum_{|\mathbf{k}|\leqslant k_m}\hbar\omega_{\mathbf{k}}(b_{\mathbf{k}}^+ b_{\mathbf{k}} + b_{\mathbf{k}} b_{\mathbf{k}}^+); \tag{4}$$

$$\hat{H}_{eL} = i\sqrt{\frac{\nu}{VN}} \sum_{|\mathbf{k}| \leqslant k_m} \alpha(\mathbf{k}) \{b_{\mathbf{k}}^{+} e^{-i\mathbf{kr}} - b_{\mathbf{k}} e^{i\mathbf{kr}}\}. \tag{5}$$

In these formulas $\mathbf{r}$ is the electron radius vector, $W(\mathbf{r})$ is the periodic potential of the crystal, $\mathbf{p}$ and $\mathbf{k}$ are the electron quasi-momentum and the phonon wave vector, $\omega_{\mathbf{k}}$ and $\epsilon_j(\mathbf{p})$ are functions that determine the dependence of the phonon frequency on the wave vector and of the electron energy on the quasi-momentum in the band with index j, $b_{\mathbf{k}}^{+}$ and $b_{\mathbf{k}}$ are the creation and annihilation operators of a phonon with wave vector $\mathbf{k}$, V is the normalizing volume, ν the number of valence bands, and finally $\psi_{0j}(\mathbf{p}, \mathbf{r})$ the electron wave function in the following Bloch form

$$\psi_{0j}(\mathbf{p}, \mathbf{r}) = \exp(i\mathbf{pr}/\hbar)\, u_j(\mathbf{p}, \mathbf{r}), \tag{6}$$

where $u_j(\mathbf{p}, \mathbf{r})$ is the periodic solution of the equation

$$-\frac{\hbar^2}{2m}\nabla^2 u_j(\mathbf{p}, \mathbf{r}) - \frac{i\hbar}{m}\mathbf{p}\nabla u_j(\mathbf{p}, \mathbf{r}) + \left\{W(\mathbf{r}) + \frac{p^2}{2m} - \varepsilon_j(\mathbf{p})\right\} u_j(\mathbf{p}, \mathbf{r}) = 0. \tag{7}$$

The interaction Hamiltonian (5) should generally speaking also contain functions of the Bloch type instead of the plane waves. However, in the calculation of the matrix elements for transitions between states (6) allowance for this circumstance gives only a corrective factor on the order of unity, which is insignificant, since we do not specify the form of $W(\mathbf{r})$ anyhow, nor do we consequently specify $\psi_{0j}(\mathbf{p}, \mathbf{r})$.

No assumptions whatever are made concerning the form of the function $\alpha(\mathbf{k})$, with the exception of the obvious property $\alpha(\mathbf{k}) \sim k^{1/2}$ for small k.

It is well known that the presence of a homogeneous electric field $\mathbf{E}$ leads to a uniform increase in the quasi-momentum of the electron with time in accordance with the law $\mathbf{p} = \mathbf{p}_0 - e\mathbf{E}t$. In other words, if at the instant $t = 0$ the electron is described by a wave function $\psi_{0j}(\mathbf{p}_0, \mathbf{r})$ then in subsequent instants its wave function will have in the zeroth approximation the form

$$\psi_j(\mathbf{p}_0, \mathbf{r}, t) = \exp\left\{-\frac{i}{\hbar}\int_0^t \varepsilon_j(\mathbf{p}_0 - e\mathbf{E}x)\, dx\right\} \psi_{0j}(\mathbf{p}_0 - e\mathbf{E}t, \mathbf{r}). \tag{8}$$

By virtue the periodic dependence (with the period of the reciprocal lattice) of $\epsilon_j(\mathbf{p})$ and $\psi_{0j}(\mathbf{p}, \mathbf{r})$ on the quasi-momentum $\mathbf{p}$, a uniform increase in the latter means that the electron vibrates within the confines of a single band, if the field is directed along one of the principal crystallographic axes (with a period $2\pi\hbar/eEd$ for the case of a simple cubic lattice). If the field is not aligned with any of the reciprocal-lattice vectors, the motion can have a complex aperiodic character, but is still confined to the same band. The exact wave function should, naturally, contain also terms connected with the transitions into other bands, but these terms will obviously be small. It is therefore natural to seek a solution of the Schrödinger equation

$$i\hbar\, \partial\Psi/\partial t = \hat{H}\Psi \tag{9}$$

in the form of a superposition of products of functions of the form (8) (which already include the fundamental effect of the field — uniform acceleration) and of the phonon wave functions

$$\Psi = \sum_{\substack{j,\, \mathbf{p}_0 \\ \{N_{\mathbf{k}}\}}} c_j([N_{\mathbf{k}}], \mathbf{p}_0, t) \exp\left\{e\mathbf{E}\int_0^t \gamma_j(\mathbf{p}_0 - e\mathbf{E}x)\, dx\right\} \psi_j(\mathbf{p}_0, \mathbf{r}, t) \prod_{\mathbf{k}} (b_{\mathbf{k}}^{+})^{N_{\mathbf{k}}} \Phi_0. \tag{10}$$

The symbol $[N_{\mathbf{k}}]$ denotes the set of occupation numbers $N_{\mathbf{k}}$ corresponding to all possible values of $\mathbf{k}$; Φ_0 is the wave function of the lowest energy state of the lattice; summation over $\mathbf{p}_0$ is carried out over all physically-different states, i.e., over the volume of the first Brillouin zone

$$\gamma_j(\mathbf{p}) = \int_{\Omega_0} u_j^{*}(\mathbf{p}, \mathbf{r})\, \mathrm{grad}_{\mathbf{p}} u_j(\mathbf{p}, \mathbf{r})\, d\tau. \tag{11}$$

The integral in (11) is taken over the volume of the elementary cell. The quantity (11) is pure imaginary, since by virtue of the condition

$$\int_{\Omega_0} u_j^{*}(\mathbf{p}, \mathbf{r})\, u_j(\mathbf{p}, \mathbf{r})\, d\tau = 1 \text{ we have } \mathrm{Re}\,\gamma_j(\mathbf{p}) = 0.$$

The advisability of separating out the exponential factor in (10) will be seen from the following. Substituting this expansion into (9), we obtain in the usual manner a system of equations for the coefficients:

$$i\hbar \frac{\partial c_j([N_{\mathbf{k}}], \mathbf{p}_0, t)}{\partial t} = i\hbar \sum_{j' \neq j} \frac{\mathbf{E}\mathbf{J}_{jj'}(\mathbf{p}_0 - e\mathbf{E}t)}{\varepsilon_j(\mathbf{p}_0 - e\mathbf{E}t) - \varepsilon_{j'}(\mathbf{p}_0 - e\mathbf{E}t)} c_{j'}([N_{\mathbf{k}}], \mathbf{p}_0, t) Q^0_{jj'}(\mathbf{p}_0, t) + i\sqrt{\frac{\nu}{VN}} \sum_{j', |\mathbf{k}'| \leqslant k_m}$$

$$\times \{M_{jj'}(\mathbf{p}_0 - e\mathbf{E}t, \hbar\mathbf{k}') \sqrt{N_{\mathbf{k}'}}\, c_{j'}([N_{\mathbf{k}}] - 1_{\mathbf{k}'}, \mathbf{p}_0 + \hbar\mathbf{k}', t) Q^+_{jj'}(\mathbf{p}_0, \mathbf{k}', t)$$

$$- M_{jj'}(\mathbf{p}_0 - e\mathbf{E}t, -\hbar\mathbf{k}') \sqrt{N_{\mathbf{k}'} + 1}\, c_{j'}([N_{\mathbf{k}}] + 1_{\mathbf{k}'}, \mathbf{p}_0 - \hbar\mathbf{k}', t) Q^-_{jj'}(\mathbf{p}_0, \mathbf{k}', t)\}, \tag{12}$$

where

$$M_{jj'}(\mathbf{p}, \hbar\mathbf{k}) = \alpha(\mathbf{k}) \int_{\Omega_0} u_j^*(\mathbf{p}, \mathbf{r}) u_{j'}(\mathbf{p} + \hbar\mathbf{k}, \mathbf{r})\, d\tau, \quad \mathbf{J}_{jj'}(\mathbf{p}) = \frac{ie\hbar}{2m} \int_{\Omega_0} \{u_{j'}(\mathbf{p}, \mathbf{r}) \nabla u_j^*(\mathbf{p}, \mathbf{r}) - u_j^*(\mathbf{p}, \mathbf{r}) \nabla u_{j'}(\mathbf{p}, \mathbf{r})\}\, d\tau,$$

$$Q^{\pm}_{jj'}(\mathbf{p}, \mathbf{k}, t) = \exp\left\{\frac{i}{\hbar} \int_0^t [\varepsilon_j(\mathbf{p} - e\mathbf{E}x) - \varepsilon_{j'}(\mathbf{p} - e\mathbf{E}x \pm \hbar\mathbf{k}) \pm \hbar\omega_{\mathbf{k}}]\, dx\right.$$

$$\left. - e\mathbf{E} \int_0^t [\gamma_j(\mathbf{p} - e\mathbf{E}x) - \gamma_{j'}(\mathbf{p} - e\mathbf{E}x \pm \hbar\mathbf{k})]\, dx\right\}, \quad Q^0_{jj'}(\mathbf{p}, t) \equiv Q^+_{jj'}(\mathbf{p}, 0, t).$$

In the derivation of (12) use is made of the identity

$$\mathbf{J}_{jj'}(\mathbf{p}) / e\{\varepsilon_j(\mathbf{p}) - \varepsilon_{j'}(\mathbf{p})\} = \int_{\Omega_0} u_j^*(\mathbf{p}, \mathbf{r}) \operatorname{grad}_{\mathbf{p}} u_{j'}(\mathbf{p}, \mathbf{r})\, d\tau.$$

The diagonal term in the first sum over j' drops out virtue of the presence of an exponential factor in the expansion (10). The symbols $[N_{\mathbf{k}}] \pm 1_{\mathbf{k}'}$ denote that in the set of numbers $[N_{\mathbf{k}}]$ the quantity $N_{\mathbf{k}'}$ is replaced by $N_{\mathbf{k}'} \pm 1$, and all the remaining quantities remain the same.

To obtain an expression analogous to (1) we must discard from the system of equations (12) those terms containing collisions between electrons and phonons. Taking it into account that at $t = 0$

$$c_j(\mathbf{p}, 0) = \delta_{jV}$$

(the indices V and c are necessary in what follows to denote quantities pertaining to the valence and conduction bands respectively), introducing by way of a new variable the vector $\mathbf{p} = \mathbf{p}_0 - e\mathbf{E}t$, and resolving this vector into a component $\mathbf{p}_\parallel$ parallel to the field and a vector $\mathbf{p}_\perp$ perpendicular to the field, we obtain

$$c_c(\mathbf{p}_0, t) = \int_{p_{0\parallel}}^{p_{0\parallel} - eEt} \frac{\mathbf{E}\mathbf{J}_{Vc}(\mathbf{p})}{\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p})} \exp\left\{i \int_{p_{0\parallel}}^{p_\parallel} [\varepsilon_c(\mathbf{p}') - \varepsilon_V(\mathbf{p}')] \frac{dp'_\parallel}{e\hbar E}\right.$$

$$\left. + \int_{p_{0\parallel}}^{p_\parallel} \mathbf{n}\gamma_{Vc}(\mathbf{p}')\, dp'_\parallel\right\} \frac{dp_\parallel}{eE}, \quad \gamma_{Vc}(\mathbf{p}) \equiv \gamma_c(\mathbf{p}) - \gamma_V(\mathbf{p}). \tag{13}$$

Here $\mathbf{n}$ is a unit vector in the direction of the field.

For simplicity and clarity let us first calculate the integral (13) for the simplest case, when the field is directed along one of the principal crystalline axes of a simple cubic lattice with a period d, and then generalize the results to include the case of a lattice of any symmetry and an arbitrarily oriented field.

In this particular case, as indicated above, the motion of the electron in the band is periodic with a period $2\pi\hbar/eEd$. A natural characteristic of the infiltration is therefore the probability of passing through the conduction band during one period

$$D_0(\mathbf{p}_\perp) = \left| \int_{-\pi\hbar/d}^{\pi\hbar/d} \frac{\mathbf{n}\mathbf{J}_{Vc}(\mathbf{p})}{e[\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p})]} \exp\left\{i \int_{p_{0\parallel}}^{p_\parallel} \frac{\varepsilon_c(\mathbf{p}') - \varepsilon_V(\mathbf{p}')}{e\hbar E} dp'_\parallel + \int_{p_{0\parallel}}^{p_\parallel} \mathbf{n}\gamma_{Vc}(\mathbf{p}')\, dp'_\parallel\right\} dp_\parallel \right|^2. \tag{14}$$

At a fixed value of $\mathbf{p}_\perp$ the functions $\epsilon_c(\mathbf{p})$ and $\epsilon_V(\mathbf{p})$ are different branches of the same infinitely-valued analytic function $\epsilon(\mathbf{p})$ of complex variable $p_\parallel$, since they represent different roots of the same eigenvalue problem. Since these functions are close to each other on the real axis, there should be located somewhere near the real axis in the complex plane a branch point $p_\parallel = q$, in which

$$\varepsilon_c(q) = \varepsilon_V(q).$$

Obviously, q depends on $p_\perp$. In the vicinity of the point q the branching is into two bands and $\epsilon_c(p_\parallel) - \epsilon_V(p_\parallel) \sim \sqrt{p_\parallel - q}$. The possibility of branching into a larger number of bands,* as well as the multiplicity of the inverse function $p_\parallel(\epsilon)$ in the vicinity of the point q (which corresponds to $\epsilon_c(p_\parallel) - \epsilon_V(p_\parallel) \sim (p_\parallel - q)^{(2n+1)/2}$) would appear only accidently, under particular selection of the potential W(r), and will therefore not be considered below.

By using the general properties of solutions of second-order differential equations it is possible to show (see Supplement) that in the vicinity of the point q the quantity $n\gamma_{Vc}(p_\parallel)$ behaves as $(p_\parallel - q)^{-1/2}$ and the factor ahead of the exponent in (14) has at the same point a simple pole with a universal value of the residue i/4. The latter circumstance makes it possible to calculate the integral in Eq. (14). For this purpose we introduce into (14) a new variable $y = y(p_\parallel)$, satisfying the following conditions: (1) $dy(p_\parallel)/dp_\parallel$ is real and positive for all real $p_\parallel$ and has the same period as the reciprocal lattice; (2) $dy(p_\parallel)/dp_\parallel$ behaves in the vicinity of the point $p_\parallel = q$ as $(p_\parallel - q)^{-1/2}$ and has neither zeros nor singularities lying closer to the real axis than q.

It is easy to check that the integrand in (14) is single-valued in the vicinity of the point $y = y(q)$ and has at this point a pole with residue i/2. An example of such a function is the integral

$$\int_0^{p_\parallel} \frac{dp'_\parallel}{\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p})}.$$

Shifting the contour of integration in (14) to the upper half of the y plane, we obtain

$$D_0(\mathbf{p}_\perp) = \pi^2 \left| \exp\left\{ \int_0^q \left[i\,\frac{\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p})}{e\hbar E} + n\gamma_{Vc}(\mathbf{p}) \right] dp_\parallel \right\} \right|^2 \tag{15}$$

with accuracy to terms that are exponentially small compared with fundamental term. Equation (15) discards, in addition, terms connected with the limits of integration, and therefore not increasing with time.†

The probability of passage after ℓ periods is,

$$D_l(\mathbf{p}_\perp) = D_0(\mathbf{p}_\perp) \left| \sum_{r=0}^{l} e^{irs_0} \right|^2 = \frac{\sin^2[(l+1)s_0/2]}{\sin^2(s_0/2)} D_0(\mathbf{p}_\perp),$$

where

$$s_0(\mathbf{p}_\perp) = \int_{-\pi\hbar/d}^{\pi\hbar/d} \left\{ i\,\frac{\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p})}{e\hbar E} + n\gamma_{Vc}(\mathbf{p}) \right\} dp_\parallel$$

is a rapidly oscillating function of the field ($s_0 \sim 1/E$).

However, taking it into account that s_0 is a function of $\mathbf{p}_\perp$ and averaging over the narrow region $\Delta p_\perp$, we obtain

$$D_l(\mathbf{p}_\perp) = lD_0(\mathbf{p}_\perp). \tag{16}$$

Equations (14)—(16) were obtained, as already remarked, for a field directed along one of the principal crystallographic axes of a simple cubic lattice, when the motion of the electron has a simple peri-

*A distinction must be made between the branching of $\epsilon(p_\parallel)$ for fixed $p_\perp$, considered here, and the frequently-encountered band degeneracy by virtue of the crystal symmetry. In the latter case the equivalent states belonging to the various zones correspond to different directions of the quasi-momentum, i.e., to different $p_\perp$.

†A formula analogous to (14) was also used by Franz.[4] However, he did not take into account thereafter the presence of a pole in the integrand, and as a result his final results differed strongly from (15) and the equations that follow. These include the value of the factor ahead of the exponent at the point q, which as shown above, is infinite. If this factor is nevertheless replaced by the most sensible value $\sim d/2\pi\hbar$, the resultant values are several orders of magnitude smaller than (15), and contain a factor ahead of the exponent that is dependent on E ($\sim E^{4/3}$).

odic character. In general, however, these cannot be used, since the motion is generally speaking aperiodic and the probability of infiltration changes from period to period. But this circumstance, which is significant in the calculation of the probability of passage of each electron, does not play any role for the total effect when the valence electrons fully fill the band, since the place of the departing electron is taken during the next period by another electron, which will infiltrate with the same probability. Still another difference lies in the fact that the time interval, to which formula (15) pertains, is no longer equal to $2\pi\hbar/eEd$, but is determined by the length, divided by eE, of the segment of the straight line lying within the first Brillouin zone parallel to the field and passing through point $\mathbf{p}_0$, viz.: $T(\mathbf{p}_0) = \Delta p_{\parallel}(\mathbf{p}_0)/eE$. But even this fact is of no significance for the summary effect, because all the electrons having initial states along this segment have an equal probability of passage, and their number is proportional to $\Delta p_{\parallel}(\mathbf{p}_0)/2\pi\hbar$. The number of electrons passing in the conduction band per unit time is proportional to $\Delta p_{\parallel}(\mathbf{p}_0)/T(\mathbf{p}_0)2\pi\hbar = eE/2\pi\hbar$ and is independent of $\Delta p_{\parallel}$, i.e., of the form of the Brillouin band. The equations that will be obtained below are therefore valid for lattices of all symmetries and for all field orientations.

The total number of electrons passing in the conduction band per unit time and per unit volume is given by the expression

$$n = \frac{eE}{2\pi\hbar}\int D_0(\mathbf{p}_\perp)\frac{d^2p_\perp}{(2\pi\hbar)^2} = \pi^2\frac{eE}{2\pi\hbar}\int\left|\exp\left\{\int_0^{q(\mathbf{p}_\perp)}\left[i\frac{\varepsilon_c(\mathbf{p})-\varepsilon_V(\mathbf{p})}{eE\hbar}+\mathbf{n}\gamma_{Vc}(\mathbf{p})\right]dp_\parallel\right\}\right|^2\frac{d^2p_\perp}{(2\pi\hbar)^2}. \tag{17}$$

The principal contribution to the integral (17) is made by the narrow region near that value of $\mathbf{p}_\perp$, which corresponds to the minimum of the function $\epsilon_c(\mathbf{p}) - \epsilon_V(\mathbf{p})$. Actually, for all specified values of $p_\perp$ this function reaches a certain relative minimum $\epsilon_{\min}(\mathbf{p}_\perp)$ as $p_\parallel$ is varied. The greater $\epsilon_{\min}(\mathbf{p}_\perp)$, the farther is the branch point $q(\mathbf{p}_\perp)$ from the real axis. At values of $\epsilon_{\min}(\mathbf{p}_\perp)$ that are not too large, $q(\mathbf{p}_\perp) \sim \epsilon_{\min}^{1/2}(\mathbf{p}_\perp)$. The exponent therefore contains a large negative quantity, proportional to $\epsilon_{\min}^{3/2}(\mathbf{p}_\perp)$ and consequently, the integrand has a sharp maximum in the region where the function $\epsilon_c(\mathbf{p}) - \epsilon_V(\mathbf{p})$ has an absolute minimum.

In the vicinity of this point we cannot restrict ourselves to an ordinary quadratic expansion, and must use a somewhat more accurate expression, taking into account the presence of a branch-point surface.

$$\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p}) = \varepsilon_0\left[1+\sum_{i,k}(p_i-p_{0i})(p_k-p_{0k})/m_{ik}\varepsilon_0\right]^{1/2}. \tag{18}$$

On the other hand, the quantity

$$\int_0^{q(\mathbf{p}_\perp)}\mathbf{n}\gamma_{Vc}(\mathbf{p})\,dp_\parallel$$

can be considered constant within this region with a sufficient degree of accuracy, $-\mathbf{n}\gamma_{0Vc} \lesssim 1$.

Elementary integration leads then to the following final result:

$$n = 2\pi\nu\left(\frac{eE}{2\pi\hbar}\right)^2\left(\frac{m_1m_2m_3}{m_\parallel^3}\frac{m_\parallel}{\varepsilon_0}\right)^{1/2}\exp\left\{-\frac{\pi}{2e\hbar E}\sqrt{m_\parallel}\,\varepsilon_0^{3/2}+\mathbf{n}\gamma_{0Vc}\right\}, \tag{19}$$

where $m_\parallel^{-1} = \sum_i[(\cos^2\gamma_i)/m_i]$, m_i^{-1} are the principal values of the tensor m_{ik}^{-1} [see (18)], and γ_i are the angles between the directions of the field and the principal axes of this tensor, which in general do not coincide with the principal crystal axes.

This formula can be presented in a clearer form if one introduces the "average lattice period" d using the equation $d^3 = \Omega_0 = \nu/N$, where Ω_0 is the volume of the elementary cell. Then

$$n = N\frac{\pi}{2}\frac{eEd}{2\pi\hbar}\left(\frac{m_1m_2m_3}{m_\parallel^3}\right)^{1/2}\frac{eEd}{(\pi\hbar/d)(\varepsilon_0/m_\parallel)^{1/2}}\exp\left\{-\frac{\pi}{2e\hbar E}\sqrt{m_\parallel}\,\varepsilon_0^{3/2}+\mathbf{n}\gamma_{0Vc}\right\}. \tag{20}$$

This expression differs from (1) in that the factor in front of the exponent is dependent on the field. The exponential term contains an explicit angular dependence, which can appear also in crystals of cubic symmetry in the presence of degeneracy of the valence band or of the conduction band.

A more substantial difference lies, however, in the meaning that must be ascribed to the quantities ϵ_0 and $m_{\|}$. As can be seen from the above derivation, ϵ_0 is the width of the forbidden band and m_{ik} is the effective mass of the electron (more accurately, the reduced effective mass of the electron and hole) only in that case when the highest states of the valence band and the lowest state of the conduction band correspond to the same value of the quasi-momentum. In practice this never happens, and consequently ϵ_0, coinciding with the red boundary of light absorption for a given crystal, is always substantially greater than the width of the forbidden band. By virtue of this, the values of the critical fields, corresponding to a noticeable infiltration, should be considerably greater than those usually expected on the basis of Eq. (1).

If, however, there exists some interaction that changes the quasi-momentum of the electron, a transition is possible from the highest state of the valence band to the lowest state of the conduction band.

Such interactions may be collisions between electrons and electrons or between electrons and phonons or impurity atoms. The field dependence of the probability of an infiltration involving these processes is given by the same exponential factor as in (19), except that the width of the forbidden band enters in place of ϵ_0. In view of the very strong dependence of (19) on ϵ_0, the role of these processes can turn out to be decisive.

Let us note also that while formula (17) is quite rigorous, formulas (19) and (20) are obtained under the assumption that ϵ_0 is small compared with the widths of the valence and the conduction bands, and therefore expansion (18) is valid everywhere within the confines of the forbidden band. For the opposite case, corresponding to the appproximation of strongly bound electrons, it is also possible to obtain a relatively simple expression. In this approximation, as is known,[5] it is possible to retain only the first terms in the Fourier expansion

$$\varepsilon_c(\mathbf{p}) - \varepsilon_V(\mathbf{p}) = I_0\left[1 + \sum_{\mathbf{g}} \alpha_g e^{i\mathbf{pg}/\hbar}\right], \tag{21}$$

where $\mathbf{g}$ are the vectors of the crystal lattice. In this case $I_0 \approx \epsilon_0 \approx \Delta$ and $\alpha_{\mathbf{g}} \ll 1$. Then

$$n \sim N \frac{eEd}{2\pi\hbar} \frac{eEd}{I_0} \exp\left\{-\frac{I_0}{eEd}\left[q_0 - \sum_{\mathbf{g}} \frac{\alpha_{\mathbf{g}}\hbar}{\mathbf{ng}}\left(\exp\left\{\frac{i\mathbf{p}_m - q_0\mathbf{n}}{\hbar}\mathbf{g}\right\} - \exp\left\{i\frac{\mathbf{p}_m\mathbf{g}}{\hbar}\right\}\right)\right]\right\}, \tag{22}$$

where q_0 is determined by

$$\sum_{\mathbf{g}} \alpha_{\mathbf{g}} \exp\left\{\frac{i\mathbf{p}_m - q_0\mathbf{n}}{\hbar}\mathbf{g}\right\} + 1 = 0.$$

It is easy to verify that the fundamental term in the exponent is always of the order $-(I_0/eEd)\ln(1/\alpha)$, where α is the ratio of the widths of the allowable and forbidden bands and consequently, the transmission coefficient diminishes somewhat slower with increasing ϵ_0 than called for formula (19). In this approximation, the exponential factor (22) is close to that obtained by Feuer[6] for the unidimensional case.

It must be emphasized that the entire above analysis of the problem of production of electron-hole pairs by the electric field starts out with the far-reaching assumptions on which the band theory of solids is based. It is assumed, in particular, that both the electron states (states of the conduction band) and hole states (valence band) can be obtained by solving a certain single-electron problem with a specified periodic potential, and that the electrons and holes created do not interact with each other. As pointed to the author by L. D. Landau, it would be logical to consider the electrons and holes as two different branches of the excitation spectrum of the crystal, without making any further detailed assumption concerning their nature, and also to take it into account that they are created not free, but interact with each other in accordance with the Coulomb law, which may lead to certain changes in the factor in front of the exponent in (19).

It can be shown that such an analysis will lead in practice to the above results if the Coulomb interaction between the electron and hole are disregarded. There are grounds for hoping that allowance for the latter does not change strongly the final derivations, since the value of the Born parameter $e^2/\mu\hbar v$ (where μ is the dielectric constant and v the relative velocity of the electron and hole), a parameter characteristic of this problem, is of the same order of magnitude as the square root of the ratio of the electron and hole bound-state energy to the width of the forbidden band, and is always small.

In conclusion I thank Professor V. L. Ginzburg and Academician L. D. Landau who made many valuable comments when evaluating the results of this work.

SUPPLEMENT

Let us investigate the behavior of the function

$$J_{cV}(\mathrm{p})/e\,[\varepsilon_c(\mathrm{p})-\varepsilon_V(\mathrm{p})]\equiv\int\limits_{\Omega_0} u_c^*(\mathrm{p},\,\mathrm{r})\,\mathrm{grad}_{\mathrm{p}}\,u_V(\mathrm{p},\,\mathrm{r})\,d\tau. \tag{1S}$$

in the complex plane. For simplicity all further arguments will concern the unidimensional case. Equation (1S), which takes place on the real axis, cannot be extended to the entire complex plane, since its right half contains an essentially non-analytic operation, complex conjugation. We can, however, use the circumstance that on the real axis $u_j^*(p, x) = u_j(-p, x)$ and rewrite (1S) as

$$\frac{J_{cV}(p)}{e\,[\varepsilon_c(p)-\varepsilon_V(p)]}\equiv\int\limits_{\Omega_0} u_c(-p,\,x)\,\frac{\partial u_V(p,\,x)}{\partial p}\,dx. \tag{1'S}$$

In such a form it is possible to continue this equation analytically into the region of complex values of p. Analogously, an analytic continuation of the orthogonality and normalization conditions of the functions $u_j(p, x)$ in the p plane leads to the equation

$$\int\limits_{\Omega_0} u_j(-p,\,x)\,u_{j'}(p,\,x)\,dx=\delta_{jj'}. \tag{2S}$$

To investigate the properties of (1'S) in the vicinity of the energy branch point q it is advisable to change to a new independent (generally speaking, complex) variable, the energy ϵ. Corresponding to each value of ϵ are two values of p of opposite sign, and consequently two functions $u_j(\pm p(\epsilon), x) = u_j^{\pm}(\epsilon, x)$. The function $p(\epsilon)$ is regular in the vicinity of the point q. The index indicating the number of the zone can be omitted hereinafter, since the states of the different zones correspond to different energies. On the other hand, at the point where the energies of both zones coincide, i.e., at the branch point, the functions $u_c^{\pm}(\epsilon, x)$ and $u_V^{\pm}(\epsilon, x)$ tend to values that differ only by a factor that is independent of x. Actually, were this not so, the initial Schrödinger equation would have at this energy value four linearly-independent solutions, an impossibility for a second-order equation.

If we detour the branch point ϵ_q in the ϵ plane, then by virtue of the uniqueness of the solution of the differential equation with the given boundary condition (i.e., with p specified), we should obtain a function that differs from the original one only by a multiplying factor (which in general depends on ϵ). Consequently, $u^{\pm}(\epsilon, x)$ can be represented in the form $C^{\pm}(\epsilon)\,v^{\pm}(\epsilon, x)$ where the $v^{\pm}(\epsilon, x)$ are unique functions of ϵ in the vicinity of the point ϵ_q. Furthermore, they can be considered regular at this point, for if they had, for example, a pole of the mth order, the factor $(\epsilon-\epsilon_q)^{-m}$ could be taken out and included in $C^{\pm}(\epsilon)$.

Let us denote by ϵ_1 and ϵ_2 the two values of energy, corresponding to the same value of the quasi-momentum. As p approaches q, both these quantities tend to ϵ_q, and then, to accuracy within terms of higher order, we have $\epsilon_2-\epsilon_q=\epsilon_q-\epsilon_1$. Then

$$v^{\pm}(\varepsilon_2,\,x)=v^{\pm}(\varepsilon_1,\,x)+2\,(\varepsilon_q-\varepsilon_1)\,\partial v^{\pm}(\varepsilon_1,\,x)/\partial\varepsilon+\ldots. \tag{3S}$$

Let us insert this expansion into the normalization condition for $v^{\pm}(\epsilon_2, x)$ and take account at the same time of the orthogonality of $v^{+}(\epsilon_2, x)$ and $v^{-}(\epsilon_1, x)$

$$1=C^{+}(\varepsilon_2)\,C^{-}(\varepsilon_2)\int\limits_{\Omega_0} v^{-}(\varepsilon_2,\,x)\,v^{+}(\varepsilon_2,\,x)\,dx=C^{+}(\varepsilon_2)\,C^{-}(\varepsilon_2)\,2\,(\varepsilon_q-\varepsilon_1)\int\limits_{\Omega_0} v^{-}(\varepsilon_2,\,x)\,\frac{\partial v^{+}(\varepsilon_1,\,x)}{\partial\varepsilon}\,dx+\ldots. \tag{4S}$$

Let us now compare the right half of (4S) with the quantity of interest to us

$$\int\limits_{\Omega_0} u_c(-p,\,x)\,\frac{\partial u_V(p,\,x)}{\partial p}\,dx=C^{-}(\varepsilon_2)\,C^{+}(\varepsilon_1)\,\frac{d\varepsilon_1}{dp}\int\limits_{\Omega_0} v^{-}(\varepsilon_2,\,x)\,\frac{\partial v^{+}(\varepsilon_1,\,x)}{\partial\varepsilon}\,dx. \tag{5S}$$

As a result we obtain the relation

$$\int\limits_{\Omega_0} u_c(-p,\,x)\,\frac{\partial u_V(p,\,x)}{\partial p}\,dx=\frac{d\varepsilon_V(p)}{dp}\,\frac{C^{+}(\varepsilon_V)}{C^{+}(\varepsilon_c)}\,\frac{1}{2\,(\varepsilon_q-\varepsilon_V)}=\frac{C^{+}(\varepsilon_V)}{C^{+}(\varepsilon_c)}\,\frac{1}{4\,(p-q)}. \tag{6S}$$

It is obviously always possible to choose the functions so as to obtain on the real axis $C^+(\epsilon) = C^-(\epsilon)$. Then this equality is always retained, and in the vicinity of the branch point it follows from this equality that $C^+(\epsilon) = C^-(\epsilon) \sim (\epsilon - \epsilon_q)^{-1/2}$ and, consequently

$$\lim_{p \to q} \frac{C^+(\epsilon_V)}{C^+(\epsilon_c)} = \pm i.$$

At first glance it may appear that this deduction is not unique, since $C^{\pm}(\epsilon)$ are determined on the real axis with an accuracy to within an arbitrary phase factor. However, strictly speaking, we are interested not in the quantity (1S), but in its products by the factor

$$\exp\left\{\int_0^p \gamma_{Vc}(p')\,dp'\right\},$$

which is contained in the intergrand of (14) and which, as can be readily verified from the definition of $\gamma_{Vc}(p)$, is in general independent of this phase factor.

We thus have near the branch point

$$\int_{\Omega_0} u_c(-p, x) \frac{\partial u_V(p, x)}{\partial p}\,dx = \pm \frac{i}{4(p-q)} + \ldots \tag{7S}$$

From the definition of the quantity $\gamma_{Vc}(p)$ it is seen that the expansion in powers of $\epsilon - \epsilon_q$ contains no even terms, and therefore the expansion begins with the term $(\epsilon - \epsilon_q)^{-1} \sim (p - q)^{-1/2}$.

[1] C. Zener, Proc. Roy. Soc. A145, 523 (1934).
[2] McAfee, Ryder, Shockley, and Sparks, Phys. Rev. 83, 650 (1951). See also Vol'kenshtein, J. Tech. Phys. (U.S.S.R.) 9, 171 (1939).
[3] W. V. Houston, Phys. Rev. 57, 184 (1940).
[4] W. Franz and L. Tewordt, Halbeiterprobleme 3, 1 (1956).
[5] R. E. Peierls, Quantum Theory of Solids, Oxford, 1955.
[6] P. Feuer, Phys. Rev. 88, 92 (1952).

Translated by J. G. Adashko
197