The Urbach rule and excitonic effects

A. S. Ioselevich

Crystallography Institute, USSR Academy of Sciences
(Submitted 10 September 1981)
Zh. Eksp. Teor. Fiz. **82**, 601–612 (February 1982)

The influence of Coulomb attraction between an electron and a hole on the coefficient of light absorption by a polar crystal is considered under conditions when the Urbach rule holds. It is shown that as the light frequency approaches the exciton binding energy a change takes place in the structure of the final state reached by the electron and hole produced upon absorption of the photon. If the Coulomb interaction is weak enough, the electron and hole are produced in a "quasifree" final state, wherein the regions of their localization are far from each other, and the sign of the screened interaction between them corresponds to repulsion. Enhancement of the Coulomb attraction first decreases the electron-hole distance, but at a certain finite value of the exciton Rydberg R, approximately equal to $3(E_g-\hbar\Omega)m_e/m_h$ (m_e and m_h are the effective masses of the electron and hole, E_g is the gap width, and Ω is the frequency of the light), the effective repulsion gives way to attraction, and the final state acquires an excitonic structure. In all cases, the exciton effect leads to an increase of the absorption. The strongest factor is found to be the renormalization of the absorption edge: the photon energy deficit $\Delta = E_g-\hbar\Omega$ is replaced by $\Delta-R$.

PACS numbers: 71.35. + z

1. INTRODUCTION

In a preceding paper[1] the author derived self-consistent-field equations that describe the absorption of light by a polar crystal with simultaneous absorption of a large number of LO phonons. The essentially multiparticle problem was thus reduced to a certain two-particle (electron and hole) quantum-mechanical problem, already a considerable simplification. However, the obtained nonlinear and nonstationary Schrödinger equation is quite complicated, and the structure of its solution has a number of nontrivial singularities. The self-consistent-field equations were solved in Ref. 1 only for the case of strong fluctuation overscreening, when the bare (unscreened) Coulomb attraction of the electron and hole could be neglected compared with the screened one. This succeeded in reducing the problem in fact to one with a single particle.

The present paper deals with screening of arbitrary strength, so that the resultant screened Coulomb interaction can correspond either to repulsion (overscreened situation) or to attraction (non-overscreened situation).

The appearance of a new physical parameter (on top of the old one), namely of the exciton Rydberg R, increases greatly the number of possible situations and makes the problem, generally speaking essentially two-particle. This necessitates the use of the following simplifications.

1. Although the method developed in Ref. 1 yields the preexponential factor of the absorption coefficient, we confine ourselves to calculation of only the argument of the exponential. We note that excitonic effects can lead to an exponential increase of the absorption coefficient only if

$$R\tau \gg \hbar, \tag{1}$$

where τ is the characteristic absorption time.

2. We forgo a general analysis and study only two limiting cases, quasistatic (when the time τ is short compared with the characteristic time of variation of the form of the fluctuation of ω_0^{-1}, where ω_0 is the frequency of the LO phonons), and dynamic (the opposite case). For details on the meaning of the conditions $\omega_0\tau \gtrless 1$ see Ref. 1.

The effective masses m_e of the electron and m_h of the hole will be assumed to differ greatly. To be specific,

$$\gamma = m_e/m_h \ll 1. \tag{2}$$

The condition (2) allows us in all cases to separate the motions of the electron and the hole, so that we are dealing again with a single-particle problem. We emphasize that in the investigated situation the electron and hole are in a complicated inhomogeneous self-consistent field, so that their relative motion cannot be separated from the mass-center motion at arbitrary γ, a procedure permissible in the free-exciton problem.

In addition to the inequality (2) it is necessary, however, to stipulate that the electronic contribution to the argument (proportional to γ) of the absorption-coefficient exponential must all the same be large compared with unity. Otherwise allowance for the electron (and with it for all the excitonic effects) leads only to a change of the pre-exponential factor.

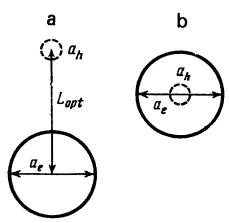

FIG. 1. Structure of the wave function of the electron-hole pair in the overscreened (a) and non-overscreened (b) cases. The electron localization region is shown solid, that of the hole, dashed. a_e and a_h are the characteristic dimensions of the localization regions of the electron and hole, and L_{opt} is the optimal distance between them.

 0038-5646/82/020358-07$04.00

We report below the main results of the study.

2. MAIN RESULTS. QUASISTATIC CASE

In the quasistatic case (when $\omega_0\tau \ll 1$) the absorption coefficient K_a is proportional to the probability of the onset of the optimal fluctuation, i.e., the most probable of all fluctuations such that the binding energy of the electron-hole pair with this fluctuation is $-\Delta = \hbar\Omega - E_g$. In this case, generally speaking, a contribution to the binding energy is made also by the electron-hole interaction, and two cases are possible.

1. If the electron-hole attraction is weak enough ($R < R_c$, where $R_c \approx 3\gamma\Delta$), the resultant interaction is overscreened and corresponds to repulsion.

2. If the attraction is strong ($R > R_c$), the interaction is not overscreened and corresponds to attraction.

It was shown in the preceding paper[1] that in the strongly overscreened situation ($R \ll R_c$) the optimal fluctuation constitutes a pair of fluctuation wells of opposite sign, located far from each other [see Fig. 1(a)]. The optimal distance between the wells L_{opt} is determined by the compeitition between the increase of the fluctuation probability and the decrease of the overlap interval (the probability that the electron and hole turn out to be at the same point of space) with increasing distance L.

If, remaining within the framework of the overscreened situation, we recognize that the bare attraction is finite, a certain decrease of the total repulsion is obtained, and as a consequence also a decrease in the optimal distance between the wells:

$$L_{opt}(R) = L_{opt}(0)\,[1-(R/R_c)^{1/2}]^{1/2}. \tag{3}$$

If R is now increased (as the measure of the Coulomb attraction), then L_{opt} vanishes formally at $R = R_c$. Actually, however, the validity of the "two-well" approximation is violated somewhat earlier, when L_{opt} becomes comparable with the width a_e of the well. The expression for the absorption coefficient at $R < R_c$ and $(\Delta/\hbar\omega_0)^{1/2}/\alpha^* N \ll 1$ is

$$\ln K_a = -\frac{1.17}{\alpha^* N}\left(\frac{\Delta}{\hbar\omega_0}\right)^{3/2} - 7\left[\frac{\gamma}{\alpha^* N}\left(\frac{\Delta}{\hbar\omega_0}\right)^{3/2}\left(1-\left(\frac{R}{R_c}\right)^{1/2}\right)\right]^{1/2}$$

$$+1.73\left(\frac{R}{R_c}\right)^{1/2}\ln\left\{\frac{12\gamma}{\alpha^* N}\left(\frac{\Delta}{\hbar\omega_0}\right)^{3/2}\left[1-\left(\frac{R}{R_c}\right)^{1/2}\right]\right\}, \tag{4}$$

$$\alpha^* = \frac{1}{2}(\varepsilon_\infty^{-1}-\varepsilon_0^{-1})\frac{e^2}{\hbar\omega_0}\left(\frac{2(m_e+m_h)\omega_0}{\hbar}\right)^{1/2} = (\alpha_e^2+\alpha_h^2)^{1/2} \approx \alpha_h\left(1+\frac{\gamma}{2}\right).$$

Here α_e and α_h are respectively the electron and hole polaron coupling constants, and $N = [\exp(\hbar\omega_0/T)-1]^{-1}$ is the phonon occupation number.

Thus, in the overscreened case allowance for the Coulomb attraction leads to an increase of the absorption coefficient. The principal (first) term in the expression (4) for K_a, which depends only on the internal structure of the wells, is not altered by the Coulomb interaction. The second term, which describes the interaction of the electron and hole and the overlap of their wave functions, decreases in absolute value because of the decrease of the effective repulsion and the mutual approach of the wells. Moreover, an additional increase of the overlap intergral appears, namely the third term (Coulomb logarithm). In the case of strong overscreening ($R \ll R_c$) Eq. (4) agrees with the one previously obtained.[1]

Equation (4) ceases to hold when the second term becomes comparable with unity; this corresponds to the condition $L_{opt} \sim a_e$ or

$$0 < R_c - R \sim \alpha^* N \hbar\omega_0 (\hbar\omega_0/\Delta)^{1/2} \ll R_c.$$

The region in which the electron-hole wave function becomes restructured is thus very narrow.

With further increase of R ($R > R_c$) the repulsion between the electron and the hole gives way to attraction and the structure of the optimal fluctuation is radically changed.

The fluctuation becomes spherically symmetrical and constitutes a broad electron well, with a narrow hole well at its center [Fig. 1(b)]. The hole is localized in the hole well.[1)] The electron localization region is considerably broader, so that from the point of view of the electron the hole can be regarded as fixed in a point. The electron localization is due to the joint action of the (screened) attraction to the pointlike hole and to the electron self-action (the electron well).

At $R \gg R_c$ the electron energy is determined mainly by its attraction to the hole; the screening effect, as well as the interaction with the electron well proper, can be neglected in this case, and the excitonic effects reduce to a renormalization of the edge—to a decrease of Δ by an amount R.

At $R \gtrsim R_c$ all the indicated effects (the bare attraction, the screening, and the self-action) are of the same order. The electron energy is therefore in this case, too, of the order of R and the size of the state is of the same order of magnitude as the exciton Bohr radius. The absorption coefficient takes at $R > R_c$ the form

$$\ln K_a = -\frac{1.17}{\alpha_h N}\left(\frac{\Delta-R}{\hbar\omega_0}\right)^{3/2}\left[1+2\left(\frac{3\gamma R}{\Delta-R}\right)^{1/2}-2\gamma\right]. \tag{5}$$

Comparing (5) with (4) we see that both equation yield the same value at $R = R_c$.

Thus, whereas at large values of R it is expedient for the electron and hole "to sit one on the other" and form a spherically symmetrical configuration, at $R = R_c$ a peculiar spontaneous violation of the spherical symmetry takes place: the electron breaks away from the hole. At the point $R = R_c$ the probabilities of the two configurations are equal. At $R < R_c$ the fluctuation has a dipole moment (there is a degeneracy with respect to its direction), and at $R > R_c$ the dipole moment is identically equal to zero.

It is seen from (5) that the strongest action (which reduces to a shift of the edge by an amount R) is exerted by the excitonic effects on the absorption at $\Delta \gtrsim R$. At $\Delta \gg R$ the R-dependent terms in (5) and in (4), although large compared with unity (thus indicating that they are important for the value of K_a), are small compared with the principal R-independent term, and therefore cannot alter the character of the frequency dependence of K_a.

If the hole-phonon coupling is very strong, the quasistatic approximation is valid also at low temperature. All the arguments of Ref. 1 remain in force in this case, and Eqs. (4) and (5) must be modified by replacing N with $N+\frac{1}{2}$.

We note also that the structure of the optimal fluctuation and its symmetry properties are preserved also at $\gamma \sim 1$. For example, in the derivation of (4) we have in fact not used the condition (2). Therefore expression (4) can be directly generalized to include the case of arbitrary γ. In this case $R_c = 3\gamma\Delta/(1+\gamma)^2$.

At $R > R_c$ and $\gamma \sim 1$, however, we have essentially a two-particle problem, and the calculation of the binding energy is difficult. Equation (5) cannot be generalized in simple fashion to include the case $\gamma \sim 1$. Incidentally, at $\gamma \sim 1$ the frequency region where the Coulomb interaction is not overscreened is already very narrow. At $\gamma = 1$, for example,

$$R < \Delta < {}^1/_2 R,$$

and this region is hardly of interest.

3. MAIN RESULTS. DYNAMIC CASE

In the dynamic case (in contrast to the quasistatic) the time of the absorption process is long compared with the time of the restructuring of the fluctuation well, therefore the screening can change strongly during the course of the process. Specifically, if the evolution is described in terms of imaginary time (see Ref. 1) the widths of the fluctuation wells increase exponentially with increasing imaginary time t, and their depths and the screening decrease exponentially. Three different situations are accordingly possible.

1. The Coulomb interaction is overscreened during the entire time τ of the process:

$$\tau = \omega_0^{-1} \ln\left[\frac{4.56}{\alpha^* N}\left(\frac{\bar{\Delta}}{\hbar\omega_0}\right)^{1/2}\right] \gg \omega_0^{-1}, \tag{6}$$

where $\bar{\Delta} = \Delta - E_p^h$, where E_p^h is the polaron shift for the hole and is significant in the case of strong coupling ($\alpha_h \gg 1$).

This fully overscreened situation is realized for very strong overscreening, when

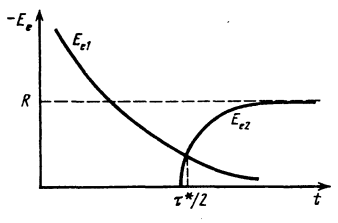

FIG. 2. Dependence of the electron binding energy E_e on the time. E_{e1} is binding energy in the electron well, and E_{e2} is the energy of binding with the hole. The condition $E_{e2} = 0$ corresponds to reversal of the sign of the interaction and to the onset of an excitonic state, while the transition takes place at the instant $\tau^*/2$, when $E_{e1} = E_{e2}$.

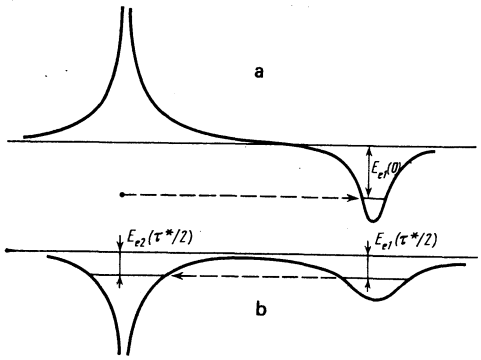

FIG. 3. Dependence of the potential energy of the electron on the coordinate at the instant $t = 0$ of the direct transition (a) and at the instant $t = \tau^*/2$ of the inverse transition (b).

$$\frac{R_c}{R} > \frac{4.56}{\alpha^* N}\left(\frac{\bar{\Delta}}{\hbar\omega_0}\right)^{1/2} \gg 1. \tag{7}$$

In this case the electron and hole are far from each other all the time, each in its own nonstationary fluctuation well. The excitonic effect influences only the terms that describe the interaction between the electron and the hole. Just as in (3), the distance L_{opt} between wells decreases. By virtue of the inequality (7), however, this decrease is insignificant and can be neglected. Thus, in the fully overscreened case the excitonic effects are not important and the following equation is valid for the absorption coefficient (see Ref. 1)

$$\ln K_e = -\frac{\bar{\Delta}}{\hbar\omega_0}\ln\left[\frac{2.77}{\alpha^* N}\left(\frac{\bar{\Delta}}{\hbar\omega_0}\right)^{1/2}\right] - 8.55\left(\frac{\gamma\bar{\Delta}}{\hbar\omega_0}\right)^{1/2}. \tag{8}$$

2. The Coulomb interaction is overscreened on 1 part of the time τ. This partially overscreened situation is possible under the condition

$$\frac{4.56}{\alpha^* N}\left(\frac{\bar{\Delta}}{\hbar\omega_0}\right)^{1/2} > \frac{R_c}{R} > 1, \tag{9}$$

and the time during which the electron and hole are far from each other is

$$\tau^* = \omega_0^{-1} \ln(6.93 R_c/R) < \tau. \tag{10}$$

At the start of the process the screening is strong enough, so that the effective interaction corresponds to repulsion. It is expedient for the electron and hole to be far from each other in this case. In the course of time, the screening decreases exponentially, and at a certain instant the repulsion gives way to attraction and a bound state of the electron (with energy E_{e2}) is produced near the hole. It is not expedient for the electron, however, to go over immediately after its onset into this state, since it is initially too shallow. The transition takes place when the binding energies of both states—the one with the hole (E_{e2}) and the one with the electron well (E_{e1})—become equal (see Fig. 2). At this instant the radius of the electron well is still small compared with the distance between wells. Consequently the transition is of the tunneling type. This must be taken into account when L_{opt} is calculated.

The evolution of the state in imaginary time can be described in the following manner.

The thermodynamic fluctuation gives rise to two polarization wells of opposite sign separated by a distance L_{opt}. The electron and hole are produced at one point (inside the hole well). The electron tunnels next rapidly (compared with the entire time of the process) into its own well, while the hole remains in place. The electron and hole "live" then in their wells, whose depths decrease exponentially in this time (and the width increases). At some instant the binding energy of the electron in the well becomes so small that tunneling back to the hole is favored, and the electron remains the rest of the time near the hole in a state of excitonic type (Fig. 3).

The described picture takes place in the partially overscreened situation; in the case of complete overscreening there is no tunneling back and the electron remains in its well to the very end.

Expression (8) (as seen from the calculation in Sec. 5) is determined mainly by times $t \sim \omega_0^{-1}$. Therefore if $\tau^* \gg \omega_0^{-1}$ (this condition is satisfied at $R \ll R_c$) the contribution of the time during which the electron and hole are spatially separated practically coincides with (8). During the rest of the time the hole is an excitonic state with energy of the orderof R.

The final expression for K_a is of the form

$$\ln K_a = -\frac{\tilde{\Delta}}{\hbar\omega_0}\ln\left[\frac{2.77}{\alpha^* N}\left(\frac{\tilde{\Delta}}{\hbar\omega_0}\right)^{1/2}\right] \qquad (11)$$
$$-8.55\left(\frac{\gamma\tilde{\Delta}}{\hbar\omega_0}\right)^{1/2}\left[1-0.38\left(\frac{R}{R_c}\right)^{1/2}\ln 1.72\frac{R_c}{R}\right]^{1/2}+\frac{R}{\hbar\omega_0}\ln\left[\frac{0.083}{\alpha^* N}\left(\frac{\Delta}{\hbar\omega_0}\right)^{1/2}\frac{R}{R_c}\right].$$

The last term in (11) is the contribution of the exciton. Accurate to the number under the logarithm sign, its value is $R(\tau-\tau^*)/\hbar$. Expression (11) is exact at $\omega_0^{-1} \ll \tau^* \ll \tau$, i.e.,

$$\frac{4.56}{\alpha^* N}\left(\frac{\tilde{\Delta}}{\hbar\omega_0}\right)^{1/2} \gg \frac{R_c}{R} \gg 1.$$

At $R \lesssim R_c$ the distance between the wells changes substantially and τ^* turns out to be of the order of ω_0^{-1}. Therefore the first term in (8) also begins to change (α^* is replaced by α_h at $R > R_c$).

3. At $R > R_c$ the Coulomb interaction is (always) fully overscreened, and the electron wave function is spherically symmetrical during the entire time τ. In this case

$$\ln K_a = -\frac{\tilde{\Delta}-R}{\hbar\omega_0}\ln\left[\frac{2.77}{\alpha_h N}\left(\frac{\Delta-R}{\hbar\omega_0}\right)^{1/2}\right]. \qquad (12)$$

Equation (12) is exact at $R \gg R_c$. At $R \gtrsim R_c$, however (this corresponds to $\Delta \sim R/\gamma \gg R$), another term of the order of $R/\hbar\omega_0$ appears in (12) and when combined with the term $(R/\hbar\omega_0)\ln(\ldots)$ it changes the argument of the logarithm in the latter. We shall not calculate this change.

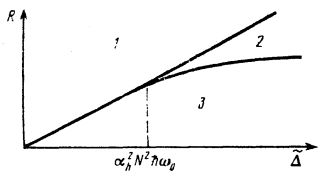

FIG. 4. Regions, on the $R\tilde{\Delta}$ plane, corresponding to different screening strengths. 1) non-overscreened situation, 2) partially screened situation, 3) full overscreening.

Thus, just as in the quasistatic case, the principal excitonic effect is the shift of the edge of the band by an amount R, a shift particularly substantial in the region $\tilde{\Delta} \sim R$.

Figure 4 shows, in the $(R\tilde{\Delta})$ plane, the regions of validity of the expressions presented above. The value $\tilde{\Delta} \sim \alpha^{*2} N^2 \hbar\omega_0$ separates the quasistatic ($\tilde{\Delta} \ll \alpha^{*2} N^2 \hbar\omega_0$) and the dynamic ($\tilde{\Delta} \gg \alpha^{*2} N^2 \hbar\omega_0$) regions. At $R > 3\gamma\Delta$ the Coulomb interaction is always overscreened. In the dynamic case the condition $R = 0.66\,\alpha^* N(\hbar\omega_0\tilde{\Delta})^{1/2}$ separates the partially and fully overscreened regions.

It must be emphasized that in the dynamic and quasistatic cases the values of R are different: in the quasistatic case

$$R = R_\infty = e^4 m_e/2\hbar^2\varepsilon_\infty^2, \qquad (13)$$

and in the dynamic

$$R = R_0 = e^4 m_e/2\hbar^2\varepsilon_0^2. \qquad (13a)$$

This difference is natural, since ε_0 manages to be formed within a time τ only if $\omega_0\tau \gg 1$.

We derive below all these expressions on the basis of the self-consistent-field method proposed in Ref. 1.

4. QUASISTATIC CASE (SELF-CONSISTENT CALCULATION)

It was shown in the preceding paper[1] that in the multiphonon case the absorption coefficient is described by the expressions

$$K_a \sim -\int d^3\mathbf{r}\, e^{-\Delta\tau}\,\mathrm{Im}\,G(\mathbf{r},\mathbf{r};0,0;\tau), \qquad (14)$$

$$G(\mathbf{R}_e', \mathbf{R}_h'; \mathbf{R}_e, \mathbf{R}_h; \tau) = -i\varphi(\mathbf{R}_e', \mathbf{R}_h'; \tau)\varphi^*(\mathbf{R}_e, \mathbf{R}_h; 0)\exp(-\tilde{S}\{\varphi\}), \qquad (15)$$

$$\tilde{S}\{\varphi\} = \int_0^\tau \langle\varphi_1^*|-\frac{\nabla_e^2}{2\gamma}-\frac{\nabla_h^2}{2}-\frac{(2R_\infty/\gamma)^{1/2}}{|\mathbf{R}_{e1}-\mathbf{R}_{h1}|}|\varphi_1\rangle dt_1$$
$$+\frac{1}{2}\int_0^\tau\int_0^\tau dt_1\,dt_2 K(t_1-t_2)\left\langle\!\!\left\langle \varphi_1^*\varphi_2^*\right|\frac{\alpha_h}{\sqrt{2}}\{-|\mathbf{R}_{e1}-\mathbf{R}_{e2}|^{-1}-|\mathbf{R}_{h1}-\mathbf{R}_{h2}|^{-1}$$
$$+|\mathbf{R}_{e1}-\mathbf{R}_{h2}|^{-1}+|\mathbf{R}_{h1}-\mathbf{R}_{e2}|^{-1}\}|\varphi_1\varphi_2\rangle\!\!\right\rangle. \qquad (16)$$

$$K(x) = (N+1)e^{-|x|}+Ne^{|x|}. \qquad (17)$$

The wave function φ of the electron-hole pair is determined from the condition that the action $\tilde{S}$ have an extremum and under the additional normalization condition $\langle\varphi^*|\varphi\rangle = 1$, while the duration τ of the process (in imaginary time) is determined from the extremum condition on K_a:

$$\Delta = -\partial\tilde{S}/\partial\tau.$$

Here and elsewhere we use units in which $\hbar = \omega_0 = m_h = 1$. The condition (2) makes the spatial scales of the electron-localization region much larger than the corresponding scales for the hole. This allows us to factorize the wave function

$$\varphi(\mathbf{R}_e, \mathbf{R}_h; t) = \psi(\mathbf{R}_e, t)\varphi(\mathbf{R}_h, t), \qquad (19)$$

where ψ and φ are the electron and hole wave functions.

Upon substitution of (19) and (16) the action $\tilde{S}$ breaks up into electron and hole terms

$$\tilde{S}=\tilde{S}_e+\tilde{S}_h, \tag{20}$$

$$\tilde{S}_h\{\varphi\}=\int_0^\tau \langle\varphi_1^*|-\frac{\nabla^2}{2}|\varphi_1\rangle dt_1$$

$$+\frac{1}{2}\int_0^\tau\int_0^\tau dt_1dt_2K(t_1-t_2)\langle\langle\varphi_1^*\varphi_2^*|-\frac{\alpha_h}{\sqrt{2}}\mathbf{R}_{h1}-\mathbf{R}_{h2}|^{-1}|\varphi_1\varphi_2\rangle\rangle, \tag{21}$$

$$\tilde{S}_e\{\psi\}=\int_0^\tau \langle\psi_1^*|-\frac{\nabla^2}{2\gamma}-\frac{\varkappa(t_1)}{|\mathbf{R}_{e1}|}|\psi_1\rangle dt_1 \tag{22}$$

$$+\frac{1}{2}\int_0^\tau\int_0^\tau dt_1\,dt_2K(t_1-t_2)\langle\langle\psi_1^*\psi_2^*|-\frac{\alpha_h}{\sqrt{2}}|\mathbf{R}_{e1}-\mathbf{R}_{e2}|^{-1}|\psi_1\psi_2\rangle\rangle,$$

where the screened Coulomb interaction constant, which depends on the imaginary time, is

$$\varkappa(t_1)=\left(\frac{2R_\infty}{\gamma}\right)^{1/2}-2^{-1/2}\alpha_h\int_0^\tau dt_2K(t_1-t_2). \tag{23}$$

$\mathbf{R}_{e1}$ in the first term of (22) is reckoned from the center of the hole well.

In the quasistatic case ($\tau\ll 1$) we obtain (see Ref. 1)

$$\tilde{S}_h=-0.11\alpha_h^2N^2\tau^3, \tag{24}$$

$$\tilde{S}_e=\min_\psi\left\{\tau\int\left(\frac{1}{2\gamma}|\nabla\psi|^2-\frac{\varkappa}{r}|\psi|^2\right)d^3\mathbf{r}\right.$$

$$\left.-2^{-1/2}\alpha_hN\tau^2\int\int d^3\mathbf{r}_1\,d^3\mathbf{r}_2|\mathbf{r}_1-\mathbf{r}_2|^{-1}|\psi(\mathbf{r}_1)\psi(\mathbf{r}_2)|^2\right\}, \tag{25}$$

$$\varkappa=(2R_\infty/\gamma)^{1/2}-\alpha_h2^{1/2}N\tau. \tag{26}$$

If $\varkappa<0$ (the overscreened situation), the electron and hole wells are far apart. Minimizing expression (25) with the aid of a direct variational method and choosing (just as in Ref. 1) a trial function in the form

$$\psi(\mathbf{r})=(14a^3)^{-1/2}(1+\rho/2a)e^{-\rho/2a}, \tag{27}$$

where $\boldsymbol{\rho}=\mathbf{r}-\mathbf{L}$, $\mathbf{L}$ is the distance between the wells, and a is the width of the electron well and is the variational parameter, we obtain (see Ref. 1)

$$\tilde{S}_e=-0.11\alpha_e^2N^2\tau^3+\varkappa\tau/L, \tag{28}$$

$$E_e=-0.33\alpha_e^2N^2\tau^2 \tag{29}$$

(E_e is the energy of the electron in the well).

To calculate K_a we must know the electron wave function in the hole-localization region, i.e., $\psi(0)$. The trial function (27) cannot be used to find $\psi(0)$, and it is necessary instead to find the correct asymptotic form of the wave function at large distances. We write down for the electron a Schrödinger equation that is valid far from the electron and hole wells

$$\left\{-\frac{1}{2\gamma}\nabla^2-\frac{\varkappa}{r}-\frac{\alpha_h2^{1/2}N\tau}{|\mathbf{r}-\mathbf{L}|}\right\}\psi(\mathbf{r})=E_e\psi(\mathbf{r}). \tag{30}$$

Solution of (30) yields (with exponential accuracy)

$$|\psi(0)|^2\sim\exp\{-2L/a_e+2(R_\infty/|E_e|)^{1/2}\ln(2L/a_e)\},\quad a_e=(2\gamma|E_e|)^{-1/2}. \tag{31}$$

When the argument of the exponential in (31) is calculated, the Coulomb logarithms corresponding to attraction to the electron well and repulsion from the hole well are mutually annihilated. What is left is the Coulomb logarithm corresponding to the unscreened electron-hole attraction.

The optimal fluctuation should ensure a maximum of the quantity $|\psi(0)|^2\exp(-\tilde{S})$. From this we easily obtain L_{opt} with the aid of (31), (28), and (27) (the Coulomb logarithm is disregarded in the determination of L_{opt})

$$L_{opt}(R)=L_{opt}(0)\,[1-(R_\infty/\alpha_e^2N^2\tau^2)^{1/2}]^{1/2}. \tag{32}$$

Substituting (24) and (28) in (14) we get

$$\tau=1.75\Delta^{1/2}/\alpha^*N, \tag{33}$$

with the second term of (28) disregarded (see Ref. 1). With the aid of (33) and (32) we obtain finally expressions (3) for $L_{\text{opt}}(R)$ and (4) for K_a, while the condition $\varkappa<0$ is found to be equivalent to the condition $R_\infty<R_c$.

In the overscreened situation, when $\varkappa>0$, the wave function of the electron is concentrated near the hole. By way of a trial function we use again (27), putting $L=0$. The functional (25) at $\varkappa>0$ recalls the F-center energy functional considered by Pekar and Deigen.[2,3] They also used a trial function in the form (27). Using their result, we can write

$$\tilde{S}_e\approx-\tau(R_\infty^{1/2}-{}^2/_3\alpha_eN\tau)^2. \tag{34}$$

Following Ref. 2, we have rounded off somewhat the numerical coefficients obtained in the exact expression. As shown in Ref. 2, the ensuring error is ~2%.

Substituting (34) and (24) in (18) we obtain

$$\tau=1.75\frac{(\Delta-R)^{1/2}}{\alpha_hN}\left[1+4\left(\frac{\gamma R}{3(\Delta-R)}\right)^{1/2}-2\gamma\right]. \tag{35}$$

We arrive ultimately at expression (5) for K_a.

5. DYNAMIC CASE (SELF-CONSISTENT CALCULATION)

In the dynamic case (see Ref. 1) we obtain

$$\tilde{S}_h=-0.024\alpha_h^2N^2e^{2\tau}-0.11\alpha_h^2\tau, \tag{36}$$

$$\tilde{S}_e=\int_0^\tau dt\int\left(\frac{1}{2\gamma}|\nabla\psi|^2-\frac{\varkappa(t)}{r}|\psi|^2\right)d^3\mathbf{r}$$

$$-\frac{\alpha_hN}{2\sqrt{2}}\int_0^\tau\int_0^\tau dt_1\,dt_2e^{|t_1-t_2|}\iint d^3\mathbf{r}_1d^3\mathbf{r}_2|\mathbf{r}_1-\mathbf{r}_2|^{-1}|\psi_1\psi_2|^2. \tag{37}$$

The second term in (36) corresponds to the polaron energy shift of the hole, which must be taken into account at $\alpha_h\gg 1$ and $N\ll 1$. The polaron shift for the electron [the term with the damped exponential in the expression for $K(t_1-t_2)$] was neglected. Integrating in (23), we get

$$\varkappa(t)=(2R_\infty/\gamma)^{1/2}-\alpha_h2^{1/2}-\alpha_hN2^{-1/2}(e^t+e^{\tau-t})=(2R_0/\gamma)^{1/2}-\alpha_hN2^{-1/2}(e^t+e^{\tau-t}). \tag{38}$$

The maximum of $\varkappa(t)$ is reached at $t=\tau/2$. Therefore if $\varkappa(\tau/2)<0$ or, equivalently, $R_0<\gamma\alpha_h^2N^2e^\tau$, the fully overscreened case is obtained. The electron and hole are separated all the time by a large distance L, and the inner structure of the electron well does not differ from the case when there is no Coulomb interaction at all. This leads to the following expression for $\tilde{S}_e$:

$$S_e=-0.024\alpha_e^2N^2e^{2\tau}+L^{-1}\int_0^\tau \varkappa(t)\,dt=-0.024\alpha_e^2N^2e^{2\tau}$$
$$+L^{-1}[2^{1/2}\alpha_hNe^\tau-\tau(2R_0/\gamma)^{1/2}]. \quad (39)$$

L_{opt} is calculated in perfect analogy with the procedure in Ref. 1 or in the preceding section of this paper. The result is

$$L_{opt}(R)=L_{opt}(0)\,[1-\tau(R_0/\alpha_e^2N^2e^{2\tau})^{1/2}]^{1/2}. \quad (40)$$

Substituting (39) and (36) in (18) we obtain expression (6) for τ, and it is easy to verify with its aid that the condition $\varkappa(\tau/2)<0$ is equivalent to condition (7). On the other hand, from (6) and (40) we obtain

$$L_{opt}(R)=L_{opt}(0)\,[1-0{,}38(R_0/R_c)^{1/2}\ln(4.56\Delta^{1/2}/\alpha^*N)]^{1/2}. \quad (41)$$

It is seen now that as a result of the condition (7) the difference between $L_{opt}(R)$ and $L_{opt}(0)$ can be neglected in the fully overscreened case. One can all the more neglect, as can be easily verified, also the Coulomb logarithm. Thus, no excitonic effects appear and Eq. (8) is valid.

In the partially overscreened case we introduce the time τ^* during which the electron is far from the hole (it is equal to that part of the time during which the interaction is overscreened, but differs from it by a value on the order of unity, i.e., much less than τ^* itself). To calculate τ^* we write $\tilde S_e$ in the form

$$\tilde S_e=\tilde S_{e1}+\tilde S_{e2},$$
$$\tilde S_{e1}=2\int_0^{\tau^*/2}dt\int d^3\mathbf{r}\left(\frac{1}{2\gamma}|\nabla\psi|^2-\frac{\varkappa(t)}{\tau}|\psi|^2\right) \quad (42)$$
$$-\frac{\alpha_hN}{\sqrt2}\int_0^{\tau^*/2}\int_{\tau-\tau^*/2}^{\tau}dt_1\,dt_2e^{|t_1-t_2|}\iint d^3\mathbf{r}_1d^3\mathbf{r}_2|\mathbf{r}_1-\mathbf{r}_2|^{-1}|\psi_1\psi_2|^2, \quad (43)$$
$$\tilde S_{e2}=-\int_{\tau^*/2}^{\tau-\tau^*/2}E_{e2}(t)\,dt,\quad E_{e2}(t)=-\left(R_0^{1/2}-\alpha_eN\frac{e^t+e^{\tau-t}}{2}\right)^2, \quad (44)$$

$\tilde S_{e1}$ is the contribution made to the action by the time interval during which the electron is in its own well, and $\tilde S_{e2}$ that near the hole. Making in (43) the substitutions

$$\psi\to(\alpha_eNe^\tau/2)^{3/2}\psi,\quad \mathbf{r}\to(\alpha_eNe^\tau/2)^{-1}\mathbf{r},$$

we obtain

$$\tilde S_{e1}=-\frac{\alpha_e^2N^2e^{2\tau}}{2}I_2(\tau^*)+L^{-1}[2^{1/2}\alpha_hNe^\tau-\tau^*(2R_0/\gamma)^{1/2}], \quad (45)$$
$$I_2(\tau^*)=\max_\psi\left\{-\int_0^{\tau^*/2}dt\int d^3\mathbf{r}\frac{1}{2}|\nabla\psi|^2+2^{-1/2}\int_0^{\tau^*/2}\int_0^{\tau^*/2}dt_1\,dt_2e^{-t_1-t_2}\right.$$
$$\left.\times\iint d^3\mathbf{r}_1d^3\mathbf{r}_2|\mathbf{r}_1-\mathbf{r}_2|^{-1}|\psi_1\psi_2|^2\right\}. \quad (46)$$

To find $I_2(\tau^*)$ at $\tau^*\gg1$ we write down the Schrödinger equation corresponding to the extremum (46) and valid at $t_1\gg1$. We use for this purpose the fact that the characteristic spatial scales of the wave function at the instant $t_1\gg1$ are much larger than the scales of the wave function at the instant $t_2\sim1$. This allows us to neglect $\mathbf{r}_2$ in $|\mathbf{r}_1-\mathbf{r}_2|^{-1}$ compared with $\mathbf{r}_1$ and to integrate with respect to $\mathbf{r}_2$ and t_2. As a result we obtain an equation with a Coulomb potential

$$\left\{-\frac{1}{2}\nabla^2-\frac{2^{1/2}e^{-t}}{r}\right\}\psi(\mathbf{r},t)=\varepsilon_{e1}(t)\psi(\mathbf{r},t), \quad (47)$$

which is easily solved. For the binding energy we have

$$E_{e1}(t)=\alpha_e^2N^2\varepsilon_{e1}(t)/4=-(\alpha_e^2N^2/4)e^{2\tau-2t}. \quad (48)$$

We can now calculate $I_2(\tau^*)$:

$$I_2(\tau^*)=I_2(\infty)-\left\{-\int_{\tau^*/2}^{\infty}dt_1\int d^3\mathbf{r}\frac{1}{2}|\nabla\psi|^2\right.$$
$$\left.+2^{1/2}\int_{\tau^*/2}^{\infty}dt_1\int_0^{\infty}dt_2e^{-t_1-t_2}\iint d^3\mathbf{r}_1d^3\mathbf{r}_2|\mathbf{r}_1-\mathbf{r}_2|^{-1}|\psi_1\psi_2|^2\right\}$$
$$=I_2(\infty)+\int_{\tau^*/2}^{\infty}\varepsilon_{e1}(t_1)\,dt_1=0{,}048-\frac{1}{2}e^{-\tau^*}.$$

We obtain τ^* from the extremum condition on $\tilde S_e$:

$$\frac{\partial}{\partial\tau^*}(\tilde S_{e1}+\tilde S_{e2})=E_{e1}\left(\frac{\tau^*}{2}\right)-E_{e2}\left(\frac{\tau^*}{2}\right)=0.$$

The electron thus goes over from a state in its own polarization well into the excitonic state at the instant when the energies of the two states become equal.

Using (48) and (44) we obtain

$$e^{\tau^*}=\alpha_e^2N^2e^{2\tau}/R_0,$$

which yields, after substitution of (6), the quantity (10) and

$$\tilde S_e=-{}^1/_2\gamma\Delta-R_0(\tau-\tau^*-2)+L^{-1}\{[2\Delta/I_2(\infty)]^{1/2}-(2R_0/\gamma)^{1/2}\tau^*\}.$$

To calculate L_{opt} it is important that this is a tunneling transition, so that in the optimization one must take into account the exponential dependence of its probability W_{tr} on L_{opt}:

$$W_{tr}\sim\exp[-2L(2\gamma|E_e(\tau^*/2)|)^{1/2}]=\exp[-L(2\gamma R_0)^{1/2}]. \quad (49)$$

Using (45), (49), and expression (31) for $\psi(0)$, we find that L_{opt} should be determined from the condition that the quantity

$$L^{-1}\{[2\Delta/I_2(\infty)]^{1/2}-(2R_0/\gamma)^{1/2}\tau^*\}+2L\{\gamma(2\Delta)^{1/2}+(\gamma R_0/2)^{1/2}\},$$

be a minimum, from which we get

$$L_{opt}(R_0)=L_{opt}(0)\,[1-0{,}38(R_0/R_c)^{1/2}(\tau^*+3.2)]^{1/2}.$$

We finally obtain expression (11) for K_e. The Coulomb logarithm is disregarded in the wave function by virtue of the condition $R\ll R_c$.

In the completely overscreened situation the wave function of the electron is concentrated all the time near the hole. If $R\sim R_c$, the autolocalization term in (37) and the screening can be neglected. As a result we obtain the simple answer

$$\tilde S_e=-R_0\tau, \quad (50)$$

which leads directly to Eq. (12). In the intermediate screening region $R\sim R_c$ all three terms in the energy (Coulomb, screening, and autolocalization) turn out to be of the same order at $t\sim1$. This makes it impossible to solve the Schrödinger equation corresponding to the minimum of $\tilde S_e$. The last two energy terms, however, attenuate exponentially in time and contribute only at $t\sim1$, in contrast to the first term, which remains constant during the entire time interval τ. As a result the expression for S_e should differ from (50) at $R\sim R_c$ only by the number under the logarithm sign in τ.

6. CONCLUSION

Despite the great variety of the different physical situations, we can point out the general principal effect of the Coulomb attraction of the electron and hole: when the light frequency approaches the exciton resonance, the type of the final state of the electron-hole produced upon absorption of a photon changes. Whereas this state constitutes far from resonance a pair of particles that are far apart in space and hardly interact (the weak interaction between them corresponds to repulsion), the state near resonance has an excitonic character, the interaction is strong and corresponds to attraction, and the distance between the electron and hole are of the order of the Bohr radius.

When the parameters (such as the frequency of the light) is changed, the change of the type of the final state constitutes practically a "phase transition" (with a very narrow region). It is important that these states have different symmetries: whereas the excitonic state is spherical symmetric and has no dipole moment, the quasifree state has a nonzero dipole moment. This difference must undoubtedly manifest itself when an external electric field is superimposed. The greatest change occurs in the light-absorption coefficient in the immediate vicinity of the excitonic state. The energy of the electron-hole pair should be reckoned in this case not from the edge of the band, but from the ground state of the exciton.

In the dynamic case, the reference level is the binding energy of the localized exciton, and the dielectric constant is ε_0; in the quasistatic case the reference is the energy of the localized exciton, but the dielectric constant is ε_∞ (the reason for the latter is that the absorption time $\tau \ll \omega_0$ is simply too short for polarization to be established). On the other hand, the structure of the final expressions for the absorption coefficient changes insignificantly both in the dynamic and in the quasistatic case.

The author thanks A. G. Aronov for a discussion of the work and for helpful remarks.

[1)]The described structure of the optimal fluctuation changes in the immediate vicinity of the excitonic state (when $\Delta - R \lesssim \gamma R$). The hole is also delocalized in this case. The conditions are more favorable for formation of a "genuine" exciton (with a mobile hole, so that the binding energy is increased by an amount of the order γR) than of a localized exciton. We shall not consider this case, since the condition for the applicability of the quasistatic approximation is very stringent here and can hardly be satisfied.

---

[1]A. S. Ioselevich, Zh. Eksp. Teor. Fiz. **81**, 1508 (1981) [Sov. Phys. JETP **54**, (1981)].
[2]S. I. Pekar and M. F. Deĭgen, Zh. Eksp. Teor. Fiz. **18**, 481 (1948).
[3]S. I. Pekar, Research in Electron Theory of Crystals, US AEC Report AEC-tr-5575 (1963).

Translated by J. G. Adashko