

# Spectral properties of centers with negative Hubbard energies in glasses

V. G. Karpov

Leningrad Polytechnic Institute

(Submitted 22 February 1983)

Zh. Eksp. Teor. Fiz. **85**, 1017–1028 (September 1983)

The possible origin of the empirically established universality of the spectral properties of centers with negative Hubbard energy in glasses is ascertained. In the described approach these centers are realized when pairs of like charge carriers are self-trapped in the structural units of an amorphous substance containing atomic potentials called critical and having small quasi-elastic constants  $k$ . A significant circumstance is that the values of  $k$  have a probability distribution. The energy parameters of the self-trapped states observed in experiment are found to be related in simple fashion to the mobility gap of the substance. The agreement of the results with the empirical data is discussed.

PACS numbers: 61.40.Df, 72.20.Jv

## §1. INTRODUCTION

The so-called chalcogenide glasses ( $a$ -As<sub>2</sub>S<sub>3</sub>,  $a$ -GeSe, and others) and materials related to them ( $a$ -Se,  $a$ -As, etc.) constitute an extensive group that is of interest in the physics of amorphous structures. The general treatment of the empirical data on these substances leads to a contradictory picture. The basic facts are the following.<sup>1</sup> The Fermi level  $E_F$  is practically insensitive to introduction of impurities and to change of temperature, and is fixed near the middle of the mobility gap, a typical value of which is  $G \sim 1$  to 2 eV; the density of the localized states on the Fermi level is high:  $\rho(E_F) \sim 10^{17} - 10^{19} \text{ cm}^{-1} \cdot \text{eV}^{-1}$ . On the other hand there is an optical gap  $G_0 = G$ , and there is no paramagnetism. The contradictions are resolved by Anderson's postulate that these systems have no centers with negative Hubbard energy (NEC)  $U < 0$ . The negative  $U$  is attributed to the reaction of the atomic subsystem to the carrier localization. There is no paramagnetism because the electrons are paired, and the relation  $G_0 = G$  is due to the appreciable excess of the optical transition energy over the energy of the thermal excitation of the NEC. It is assumed on the basis of empirical data that  $|U| \lesssim 1$  eV. Anderson did not indicate the microscopic nature of such considerable values of  $|U|$ . Attempts by Mott *et al.* (see Ref. 1) to fill this gap were based on the concept of specific defects of the glass structure. According to the experimental data these defects correspond to an energy-level scheme of the type shown in Fig. 1a. It is assumed that the pairing ( $U < 0$ ) is equally effective for electrons and holes. The differences between the energies of the optical and thermal transitions is interpreted in accordance with scheme of Fig. 1b, which is standard for the localization processes. The potential energy of the atomic system depends harmonically on one configuration coordinate,  $V = (1/2)kx^2$ , and the energies ( $n = 1, 2$ ) of the localized carriers are

$$\varepsilon_n(x) = n\varepsilon_0 - nQx + U_c \delta_{n,2},$$

where  $U_c$  describes the Coulomb repulsion in the pair ( $\delta_{n,2}$  is the Kronecker symbol).

A surprising circumstance, which we wish to emphasize, is that for the parameters of the scheme of Fig. 1a the empirical result lead to the relations  $W_1^+ \approx W_2^+ \approx W_1^-$

$\approx W_2^-$  (with certain deviations from exact equalities for different materials).<sup>1</sup> This connection between the observed relaxation energies of the atomic subsystem and the mobility gap was not explained and contradicts the usual self-trapping premises. We present below a theory that establishes such a connection. The results are used to examine the singularities of the possible electronic transitions and for comparison with the observed effects.

The possibility of a large energy gain in self-trapping and the fact that  $U < 0$  are attributed in the present paper to the existence of critical atomic potentials in the systems considered. By definition, the quasi-elastic constants of the latter are anomalously small,  $k \ll k^{(0)} \equiv M\omega_D^2$  ( $M$  and  $\omega_D$  are the characteristic atom mass and the Debye frequency of the material). Self-trapping with participation of critical potentials was considered in Ref. 3, where it was demonstrated that a significant energy gain is possible when the carriers are paired. The question of the relation between the energy parameters of the self-trapped carriers and the size of the mobility gap was not considered there.

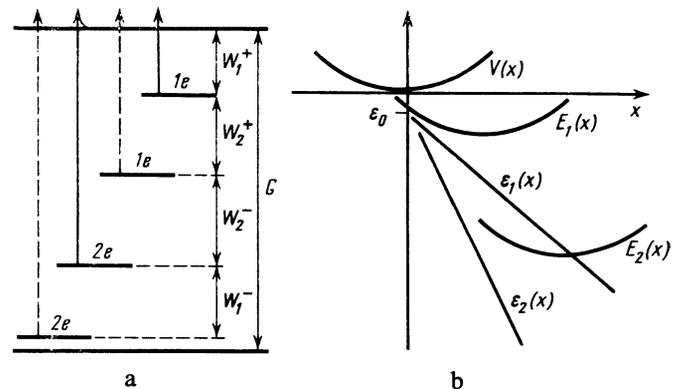


FIG. 1. a) Scheme of energy levels corresponding to electronic transitions with participation of NEC. The solid and dashed lines denote respectively thermal and optical ionization processes with one ( $1e$ ) or two ( $2e$ ) localized electrons. For NEC on which one hole ( $1h$ ) or two holes ( $2h$ ) are localized, the transition scheme is similar ( $e \rightarrow h$ , and the boundaries of the valence and conduction bands change places). b) Energies of the atomic ( $V$ ) and electronic ( $\varepsilon_n$ ) subsystems and the total energy  $E_n$  of NEC with  $n = 1$  or 2 localized carriers, as a function of the configuration coordinate  $x$ .

## §2. CRITICAL ATOMIC POTENTIALS

Problems connected with the description of critical potentials was considered in recent papers.<sup>4</sup> In this section we present the necessary results concerning the critical potentials, together with a brief physical interpretation.

Critical potentials appear in an amorphous phase because of random fluctuations of the short-range parameters (for the valence-angle bonds etc.). In an ensemble of topologically equivalent microscopic structural units, fluctuations of the short-range parameters correspond to fluctuations of the quantities  $k$ ; these fluctuations are characterized hereafter by a distribution  $g(k)$  whose approximate form is shown in Fig. 2. It is recognized here that the mean value

$$\langle k \rangle = \int g(k) k dk / \int g(k) dk$$

should be of the order of  $k^{(0)}$ . In other words, it is assumed that fluctuations with  $k \ll k^{(0)}$  are very infrequent and that a typical unit of the ensemble considered differs little from its crystalline prototype. This agrees, in particular, with the empirical fact that the macroscopic elasticities of the amorphous and crystalline phases do not differ excessively.

The region  $k < 0$  in Fig. 2 corresponds to two-well atomic potentials. Their presence in glasses is confirmed by independent empirical data on the thermal and ultrasonic properties (see Refs. 4–7). In accord with these data, the relative concentration of the two-well potentials is small,  $c \ll 1$ . The available empirical estimates of the values of  $c$  allows us to put here  $c \sim g(0)/g(k^{(0)}) \sim 10^{-4} - 10^{-5}$ . Thus,  $g(k)$  decreases rapidly with decreasing  $k$ . Specifying various  $g(k)$  dependences of exponential type we can estimate the characteristic scale of the decrease of  $g(k)$  at  $\delta k \ll k^{(0)}$ , say  $\delta k \sim 0.1k^{(0)}$  (see Ref. 4).

Since the fluctuations that lead to small values  $k \ll k^{(0)}$  are rare, we can neglect the probability of simultaneous smallness of  $k$  for two or more modes of the atomic motion, and assume that a typical critical potential is effectively single-mode,  $V = V(x)$ . For the same reason, generally speaking, no important role should be played by fluctuations for which not only  $k$  but also the higher (anharmonic) coefficients of the expansion of the atomic potential are small. At the same time, the critical character of the potential requires that its cubic anharmonicity be small. Otherwise  $V(x)$  with  $k \ll k^{(0)}$  contains a deep and a noncritical well. We assume hereafter for simplicity that the cubic anharmonicity of  $V(x)$  is small for systematic reasons.<sup>1)</sup> As a result, a typical critical potential can be described by the expression

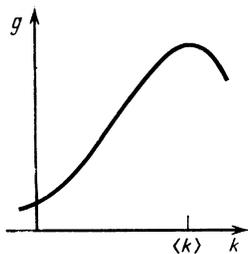


FIG. 2. Probability distribution of quasi-elastic constants of local atomic potentials in an amorphous structure.

$$V(x) \approx \frac{1}{2} k x^2 + C x^4 \quad \text{at} \quad k \ll k^{(0)}, \quad C = \text{const.} \quad (1)$$

Taking into account the character of the  $g(k)$  dependence, we confine ourselves hereafter exclusively to single-well critical potentials ( $k > 0$ ) that exceed substantially in concentration the two-well critical potentials ( $k < 0$ ,  $|k| \ll k^{(0)}$ ). It is convenient to regard the displacements  $x$  as made dimensionless by introducing a corresponding atomic length scale  $a_0 \sim 1 \text{ \AA}$ ; then  $k$  and  $C$  are energy parameters and the characteristic values are  $C \lesssim k^{(0)} \sim M \omega_D^2 a_0^2 \sim 10 \text{ eV}$ . Expression (1) serves as a basis for further description of the relaxation of the atomic subsystem in self-trapping of the carriers.

## §3. ELECTRON-ATOM INTERACTION

We examine the character of the shift of the electronic term  $\varepsilon(x)$  upon relaxation of  $x$  of the atomic subsystem. To describe such an electron-atom interaction it is customary to use the linear approximation  $\varepsilon(x) = \varepsilon_0 - Qx$  (see §1 and also Refs. 1, 2, and 8). A certain distinctive feature of the present problem is that, in accord with the requirement that  $|U|$  be large, the term  $\varepsilon(x)$  should be shifted upon relaxation by a very large amount. Thus, Anderson has proposed that the universality of the properties of NEC may be connected with the condition  $|U| > G$ , which corresponds in Ref. 2 to a shift in which  $\varepsilon(x)$  crosses the mobility gap and lands in the "foreign" band (e.g., the valence band in the case of electron localization). It is clear that, when speaking of appreciable shifts  $|\varepsilon(x) - \varepsilon_0| \sim G$ , account must be taken of the contributions made by both bands—valence and conduction—to the wave function of the state  $|l\rangle$  localized on the term  $\varepsilon$ , and the relative values of these contributions change with changing  $\varepsilon$ . The linear approximation, however, corresponds to an invariant localized wave function for a linear change  $\delta \mathcal{H}(x)$  of the potential, i.e.,

$$\varepsilon(x) = \varepsilon_0 - \langle l | \delta \mathcal{H}(x) | l \rangle = \varepsilon_0 - Qx.$$

A definite role can thus be played in our problem by the nonlinear behavior of  $\varepsilon(x)$ .

When clarifying the character of the  $\varepsilon(x)$  dependence we shall start with the known Anderson Hamiltonian (see Ref. 9)

$$\mathcal{H} = \sum_j \varepsilon_j n_j + \varepsilon'_i n_i + \sum_j v_{ji} c_j^\dagger c_i + v_{ij} c_i^\dagger c_j. \quad (2)$$

Here  $\varepsilon_j$  are the energies of all those electronic states of the system (except  $|l\rangle$ ), the interaction which is described by the quantities  $v_{ji} = \langle j_0 | \delta \mathcal{H}(x) | i_0 \rangle$ . The Hamiltonian (2) is written in the basis of the starting states  $|j_0\rangle$  and  $|i_0\rangle$  of the system (in the absence of relaxation of  $x$ ). In the linear approximation we have for the perturbation  $\delta \mathcal{H}(x)$

$$\varepsilon'_i = \varepsilon_0 - Qx, \quad v_{ji} = \alpha(\varepsilon_j) Qx, \quad (3)$$

$$Q = \langle l_0 | (\partial \mathcal{H} / \partial x)_{x=0} | l_0 \rangle, \quad \alpha(\varepsilon_j) = Q^{-1} \langle l_0 | (\partial \mathcal{H} / \partial x)_{x=0} | j_0 \rangle.$$

To determine that eigenvalue of the Hamiltonian (2) which corresponds to the state  $|l\rangle$  we obtain the equation

$$\varepsilon = \varepsilon_0 - Qx + Q^2 x^2 \sigma(\varepsilon), \quad \sigma(\varepsilon) = \sum_j \frac{|\alpha(\varepsilon_j)|^2}{\varepsilon - \varepsilon_j}. \quad (4)$$

From here on we shall take  $\varepsilon$  to mean its real part, and take accordingly the integrals contained in  $\sigma(\varepsilon)$  in the sense of the principal value. In the case of a disordered system the determination of  $\sigma(\varepsilon)$  is made difficult by the lack of detailed information on the spectrum and structure of the states  $\varepsilon_j$ . For our purposes it suffices to make clear the approximate form of  $\sigma(\varepsilon)$ . We start here from the accepted premises that the mobility-gap density of states has tails that fall off rapidly with increasing distance from the band edge (the scale of the decrease is  $w \ll G$ ).

Let us ascertain first the character of the influence of different sections of the  $\varepsilon_j$  spectrum on the behavior of  $\varepsilon(x)$ . The spectrum  $\varepsilon_j$  can have a finite low density even deep in the interior of the mobility band. The corresponding states are strongly localized, so that close-lying levels correspond to centers that are considerably separated in space and to values of  $\alpha(\varepsilon_j)$  that differ exponentially. It suffices then to take into account in  $\sigma(\varepsilon)$  only several terms corresponding to the closest centers. The random energies of these centers differ noticeably, and the interaction of a term with them can be considered separately in the "paired" approximation. Thus, the usual quasicrossing of individual levels  $\varepsilon_j$  by the term  $\varepsilon(x)$  takes place in the interior of the mobility gap. The situation changes when dealing with levels  $\varepsilon_j \approx \varepsilon_z$  near the edge  $\varepsilon_z$  of the mobility band. As  $|\varepsilon_j - \varepsilon_z| \rightarrow 0$  the density  $\rho_0(\varepsilon_j)$  of the localized states increases noticeably, as do also their radii, which become infinite at  $\varepsilon_j = \varepsilon_z$ .<sup>1</sup> Owing to the appreciable spatial overlap of such states, a comparable contribution to the interaction is made by many levels that are close in energy, i.e., what is effectively realized is interaction of the term  $\varepsilon$  with the continuous spectrum. The transition from the effectively discrete to the continuous section of the spectrum  $\varepsilon_j$  takes place in an interval  $\delta\varepsilon_j \ll G$  whose smallness is due to a rapid (exponential) decrease of the factor  $f(\varepsilon_j) = \rho_0(\varepsilon_j)|\alpha(\varepsilon_j)|^2$  as  $\varepsilon_j$  moves away from the edges  $\varepsilon_z$  into the interior of the mobility gap. The essentially nonlinear behavior of  $\varepsilon(x)$  can be due precisely to the interaction of the term with the states of the continuous spectrum.

We use first a very rough approximation, assuming that the continuous spectrum terminates abruptly at  $\varepsilon_j = \tilde{\varepsilon}_j$ . In this case

$$\sigma(\varepsilon) \approx f(\tilde{\varepsilon}_j) \ln |( \tilde{\varepsilon}_j - \varepsilon ) / G | \quad \text{at} \quad | \tilde{\varepsilon}_j - \varepsilon | \ll G, \quad (5)$$

and the term  $\varepsilon(x)$  cannot cross the boundary of the continuous spectrum. The singularity of  $\sigma(\varepsilon)$  in (5) is due precisely to the jump of  $f(\varepsilon_j)$ , owing to which the divergences of the integrals over the regions  $\varepsilon_j > \varepsilon$  and  $\varepsilon_j < \varepsilon$  in the expression for  $\sigma(\varepsilon)$  do not cancel each other.

In a real situation, regardless of the choice of  $\tilde{\varepsilon}_j$ , which is a certain arbitrary boundary of the continuous spectrum, the level density differs from zero on both its sides and the function  $\sigma(\varepsilon)$  should not be singular. However, although the spectrum does indeed have no sharp boundary,  $f(\varepsilon_j)$  can nevertheless change substantially over a small energy interval  $\delta\varepsilon_j \ll G$ . As the term  $\varepsilon$  passes through this interval,  $\sigma(\varepsilon)$  increases and the shift of  $\varepsilon(x)$  is effectively slowed.

An approximate description of the growth of  $\sigma(\varepsilon)$  when  $\varepsilon$  lands in a rapidly growing section of  $f(\varepsilon_j)$  can be obtained

by approximating on this section

$$f(\varepsilon_j) \approx f(\varepsilon) + (\varepsilon_j - \varepsilon) f'(\varepsilon).$$

The lower limit of integration in  $\sigma(\varepsilon)$  is then determined from the condition that the approximating function vanishes, and the upper limit is of the order of the length  $\delta\varepsilon_j$  of the section itself. As a result we have

$$\sigma(\varepsilon) \approx f(\varepsilon) \ln [ f'(\varepsilon) \delta\varepsilon_j / f(\varepsilon) ] + \sigma_1(\varepsilon). \quad (6)$$

The nonsingular function  $\sigma_1(\varepsilon)$  describes here the interaction of the term with the more remote regions of the continuous spectrum  $|\varepsilon - \varepsilon_j| \sim D$ . We bear it in mind here that outside the mobility gap  $f(\varepsilon_j)$  increases smoothly into the interior of allowed bands of width  $D \gg G$ . Since  $\sigma(\varepsilon_0) = 0$  in the chosen basis and in this case  $\sigma(\varepsilon_0) \approx \sigma_1(\varepsilon_0)$ , for the important small  $|\varepsilon - \varepsilon_0| \lesssim G$  we can write

$$\sigma_1(\varepsilon) \approx (\varepsilon - \varepsilon_0) \left. \frac{d\sigma_1}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0} \sim \frac{\varepsilon - \varepsilon_0}{D^2}.$$

Taking as an estimate  $x = (\varepsilon - \varepsilon_0)/Q$ , we find that  $Q^2 x^2 |\sigma_1(\varepsilon)| \ll |\varepsilon - \varepsilon_0|$ , and allowance for  $\sigma_1(\varepsilon)$  in (4) leads only to corrections that are small in terms of the parameter  $G^2/D^2$ , without a change in the form of  $\varepsilon(x)$ . At the same time, the first term in (6) can be substantial. Differentiation of (4) yields

$$\frac{d\varepsilon}{dx} = -Q \frac{1 - 2Q^2 x \sigma(\varepsilon)}{1 - Q^2 x^2 \sigma'(\varepsilon)}. \quad (7)$$

At  $|x| \sim |\varepsilon - \varepsilon_0|/Q \lesssim G/Q$  we have

$$|Q^2 x^2 \sigma'(\varepsilon) / 2Q^2 x \sigma(\varepsilon)| \sim G / \delta\varepsilon_j \gg 1.$$

Therefore, considering the strong slowing down of the term (when  $G/\delta\varepsilon_j \gg -Q^2 x^2 \sigma'(\varepsilon) \gg 1$ ) we can neglect the quantity  $2Q^2 x \sigma(\varepsilon)$  in the right-hand side of (7). Determining  $Q^2 x^2$  from (4), we get

$$\frac{d\varepsilon}{dx} \approx - \frac{Q}{1 - \Delta\varepsilon \sigma'(\varepsilon) / \sigma(\varepsilon)}, \quad \Delta\varepsilon = \varepsilon - \varepsilon_0 + Qx. \quad (8)$$

Equation (8) is convenient for iterations. In the zeroth approximation we assume that the deviation  $\Delta\varepsilon(x)$  from the linear law is zero at  $x \leq x_0$  ( $\sim G/Q$ ), and that at  $x > x_0$  total stopping takes place,  $\varepsilon(x) = \text{const}$ , i.e.  $\Delta\varepsilon(x) = Q(x - x_0)$ . Making the substitutions  $\sigma'(\varepsilon)/\sigma(\varepsilon) \approx f'(\varepsilon)/f(\varepsilon) \approx 1/\delta\varepsilon_j$ , we obtain after integrating (8)

$$\varepsilon(x) \approx \varepsilon(x_0) - \delta\varepsilon_j \ln [ 1 + Q(x - x_0) / G ] \quad \text{at} \quad x > x_0. \quad (9)$$

The corrections introduced by the subsequent iterations are small in the parameter  $\delta\varepsilon_j/G \ll 1$ . From (9) it can be seen directly that when  $\varepsilon$  lands in the section of rapid growth of  $f(\varepsilon_j)$  the linear law  $d\varepsilon/dx = -Q$  is replaced by the much weaker  $d\varepsilon/dx \approx -Q\delta\varepsilon_j/G$ . It is important in what follows that this effect manifests itself only at  $\Delta_0 = |\varepsilon - \varepsilon_0| - G \ll G$ . It can thus be assumed that as  $x$  relaxes the term  $\varepsilon(x)$  shifts practically linearly all the way to energies close to the opposite edge of the mobility gap, and

then shifts much more slowly (Fig. 3). We shall see that the specific form of the nonlinearity plays no role.

In concluding this section we note that there is one other type of nonlinearity of  $\varepsilon(x)$ , due to the higher powers of the expansion

$$\mathcal{H}(x) = \mathcal{H}(0) + \left. \frac{\partial \mathcal{H}}{\partial x} \right|_{x=0} x + \frac{1}{2!} \left. \frac{\partial^2 \mathcal{H}}{\partial x^2} \right|_{x=0} x^2 + \dots$$

It is assumed next that the corresponding corrections to the law  $\varepsilon(x) = \varepsilon_0 - Qx$  are small at the important not too large values of  $x$  (see below).

#### §4. SPECTRAL CHARACTERISTICS OF SELF-TRAPPED STATES

The equilibrium energy  $E_n^{(0)} \equiv E_n(x_n)$  of the system in self-trapping of  $n = 1, 2$  carriers is the result of minimization of the functional (see also §1):

$$E_n(x) = V(x) + n\varepsilon(x) + U_c(x) \delta_{n,2} \equiv V(x) + \varepsilon_n(x). \quad (10)$$

As shown in §3, effects connected with the change of the type of wave function of a localized state  $|l\rangle$  are negligibly small in the course of the shift of a term, all the way to values  $|\varepsilon(x) - \varepsilon_0| \approx G$ . The quantity  $U_c(x)$  in (8), which describes the carrier interaction in a pair, can then be regarded as practically constant,  $U_c(x) \approx U_c = \text{const}$ . The results of a numerical calculation<sup>8</sup> of  $U_c(x)$  for a strongly localized pair of carriers in an actual crystalline structure confirm this conclusion.

The equilibrium energies  $E_n^{(0)}$  depend parametrically on the quantities  $\varepsilon_0, k, Q, C$ , and  $U_c$ . The position of the starting term  $\varepsilon_0$  can vary in principle in a very wide range inside and outside the mobility gap. However, self-trapping effects are suppressed for  $\varepsilon_0$  outside the mobility gap, and inside the gap they have a very low relative density. It is therefore natural to assume that the term  $\varepsilon_0$  is located in the mobility gap at a distance  $\sim w \ll G$  from that gap boundary that corresponds to the type of localized carriers—electrons or holes. The corresponding states are characterized by a localization radius of the order of the “unit cell” size.<sup>1</sup> We confine ourselves hereafter to just these most effective terms and neglect for simplicity the variance of their distribution. Such a simplification is justified if, as we shall for the most part hereafter, we deal with strong self-trapping whose energy scales exceed  $w$  appreciably. We note that if necessary the probability distribu-

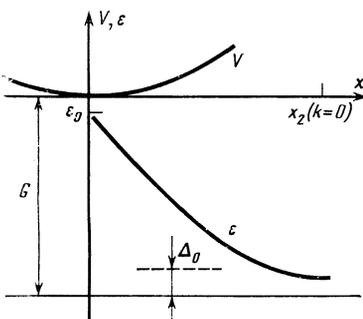


FIG. 3. Dependence of the shift of the electron term  $\varepsilon$  on the configuration coordinate  $x$ .

tion of the terms  $\varepsilon_0$  can be easily accounted for and no new qualitative peculiarities are encountered.

For the states considered with small localization radius we can estimate  $Q$  in the tight-binding approximation at  $Q \sim z\delta I / \partial x \sim zI$ , where  $z$  is the coordination number and  $I$  is the “transport integral.” Recognizing that in this approximation  $2zI \sim D$ , we get  $Q \sim D/2$ . Thus,  $Q$  is of the order of several electron volts (say,  $Q \sim 3$  eV) and at any rate  $Q > G$ . The parameter  $U_c$  remains one of the most indeterminate. By way of estimate we can use the numerous empirical data on the values of  $U_c$  for deep (multiply charged) centers in semiconductors (see Ref. 10) as well as the result of a numerical calculation.<sup>8</sup> This leads to an estimate  $U_c \lesssim 0.2-0.5$  eV.

Proceeding to minimization of the functional (8), we note that the energies  $|E_n^{(0)}|$  and  $|\varepsilon_n^{(0)}| \equiv |\varepsilon_n(x_n)|$  are maximal in the case  $k = 0$ . We assume furthermore that at  $k = 0$  the equilibrium value  $x_2$  of the coordinate  $x$  at  $n = 2$  corresponds to the nonlinear section of the  $\varepsilon(x)$  plot in Fig. 3. This is equivalent to assuming that terms cross the opposite edge of the mobility gap in the hypothetical case of the purely linear relation  $\varepsilon = \varepsilon_0 - Qx$  ( $\sigma = 0$ ) and under the condition that the self-trapping conditions are maximal ( $k = 0, n = 2$ ). The corresponding conditions on the parameters

$$\gamma = \frac{Q}{G} \left( \frac{Q}{2C} \right)^{1/2} > 1 + \frac{U_c}{2G} \quad (11)$$

agree with the estimates given for them above. With increasing  $k$ , the value of  $|x_2|$  decreases. This decrease, however, is negligibly small so long as the potential  $V(x)$  is effectively biharmonic ( $Cx_2^4 \gg \frac{1}{2}kx_2^2$ ). A noticeable change of  $|x_2|$  takes place when  $V(x)$  becomes effectively a harmonic potential, and only then can the value of  $x_2$  land on the linear section of the  $\varepsilon(x)$  plot. We arrive at the conclusion that under the condition (11) the linear region of  $\varepsilon(x)$  corresponds to harmonic  $V(x)$ . In that case the following relations hold:

$$E_n^{(0)} = n\varepsilon_0 + U_c \delta_{n,2} - n^2 Q^2 / 2k, \quad \varepsilon_n^{(0)} = E_n^{(0)} - n^2 Q^2 / 2k, \quad (12)$$

$$U \equiv E_2^{(0)} - 2E_1^{(0)} = -Q^2/k + U_c, \quad x_n = nQ/k.$$

The condition for linearity of  $\varepsilon(x)$  is

$$\varepsilon(x) = \varepsilon_0 - Qx \lesssim \varepsilon_0 - (G - \Delta_0), \quad \Delta_0 \ll G.$$

Putting here  $x = x_2 = 2Q/k$ , we find that relations (12) are valid at

$$k \gg k_G \equiv 2Q^2 / (G - \Delta_0) \approx 2Q^2 / G. \quad (13)$$

They are violated at an insignificant decrease of  $k < k_G$  to the extent that  $k_G - k \ll k_G$  is small. The inaccuracy introduced in (12) by neglecting the biharmonic component  $V(x)$  is small relative to the parameter  $\frac{1}{2}\gamma^3 \ll 1$ . It will be shown below that the NEC corresponding to  $k$  that are much smaller than  $k_G$  have practically no effect on the properties of the glasses considered, so that there is no need to minimize (8) for small  $k$  and for nonlinear  $\varepsilon(x)$ .

We introduce for the paired NEC states ( $n = 2$ ) the effective single-particle energies  $\bar{E}_2 = E_2^{(0)}/2$ . Their distribution density

$$\rho(\bar{E}_2) = ng \{k(\bar{E}_2)\} |dk(\bar{E}_2)/d\bar{E}_2| \quad (14)$$

decreases rapidly in the interior of the mobility gap; here  $k(\tilde{E}_2)$  is the inverse of  $\tilde{E}_2(k)$ ;  $n$  is the density of the starting terms whose variance is negligible. The corresponding scale of the decrease is

$$\delta\tilde{E}_2 \sim \delta k(\varepsilon_0 + U_c/2 - \tilde{E}_2)^2/Q^2 \quad \text{at} \quad k \gg k_G. \quad (15)$$

For  $k \lesssim k_G$  we have  $|\tilde{E}_2| \lesssim G/2$  and  $\delta\tilde{E}_2 \lesssim \delta k(G/D)^2$ . It can be assumed on the basis of §2 that  $\delta k \lesssim G$  ( $\sim 1$  eV), so that the scale of variation of  $\rho(\tilde{E}_2)$  is small in the sense that  $\delta\tilde{E}_2 \lesssim G(G/D)^2 \ll G$ .

In accord with the general formulation of the problem (see §1) we assume that the NEC is realized for both types of carrier—electrons and holes—and the arguments advanced above are applicable to them to an equal degree. In particular, we can consider the distribution densities  $\rho_e(\tilde{E}_{2e})$  and  $\rho_h(\tilde{E}_{2h})$  corresponding to paired states of electrons ( $e$ ) and holes ( $h$ ) on the NEC (Fig. 4). Owing to the rapid decrease of  $\rho_e(\tilde{E}_{2e})$  and  $\rho_h(\tilde{E}_{2h})$  in the interior of the mobility gap the overwhelming majority of the carriers should be localized on NEC with  $k$  close to the maximum value  $k_{\max}$  allowed by the condition for filling of their paired states  $\tilde{E}_2 - \tilde{E}_2(k)$ . If it is assumed (on empirical grounds for the time being, see §1) that the Fermi level is near the center of the mobility gap,  $E_F \approx G/2$ , then  $k_{\max}$  is determined by the conditions  $\tilde{E}_{2e} = -E_F$  and  $\tilde{E}_{2h} = -G + E_F$  and, in accordance with (12), is equal to

$$\begin{aligned} k_{\max,e} &= \frac{Q_e^2}{\varepsilon_{0e} + U_{ce}/2 + E_F} \\ &\approx \frac{2Q_e^2}{2\varepsilon_{0e} + U_{ce} + G}, \\ k_{\max,h} &= \frac{Q_h^2}{\varepsilon_{0h} + U_{ch}/2 + G - E_F} \\ &\approx \frac{2Q_h^2}{2\varepsilon_{0h} + U_{ch} + G} \end{aligned} \quad (16)$$

for the electrons and holes, respectively. Here and below the subscripts  $e$  and  $h$  designate the type of carrier to which the pertinent parameters apply. It can be seen from a comparison of (16) with (13) that at the parameter values estimated above we have  $k_{\max,e} \approx k_{G,e}$  and  $k_{\max,h} \approx k_{G,h}$ . The carriers

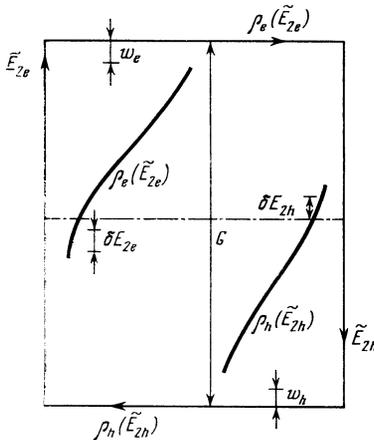


FIG. 4. Densities of effective single-particle levels of two-charge NEC states in the mobility gap. The dash-dot horizontal line is the level  $E_F$ .

are therefore localized mainly on NEC with  $k \approx k_G$ , for which the simple relations (12) hold, and the NEC with  $k \ll k_G$  (for which relations (12) are markedly violated) can be disregarded because of their low density. We note that for NEC with  $k \approx k_G$  the absolute value of the correlation energy  $|U| = -U \approx G/2 - U_c/2$  is high enough and agrees with the empirically determined situation.

Substituting in (12)  $k = k_{\max,e}$  or  $k = k_{\max,h}$  we easily obtain, given  $E_F \approx G/2$ , all the energy parameters. With the same accuracy with which the relation  $k_{\max,e}$  or  $k_{\max,h} = k_{G,h}$  is satisfied, we have

$$|E_1^{(0)}| \approx G/4, \quad |E_1^{(0)}| \approx G/2, \quad |E_2^{(0)}| \approx G, \quad |E_2^{(0)}| \approx 2G \quad (17)$$

for both types of carrier. The quantities  $W_n^\pm$  defined in §1 can be easily expressed in terms of (12). As a result we can write, with the same accuracy as in (17),

$$W_1^+ \approx W_2^+ \approx W_1^- \approx W_2^- \approx G/4. \quad (18)$$

Thus, the observed NEC energy parameters turn out to be connected with the size of the mobility gap of the material, as called for by the experimental data.

It was assumed above from empirical considerations that  $E_F \approx G/2$ . In the described approach  $E_F$  is determined from the electroneutrality condition (see Fig. 4)

$$\int_{-\infty}^{E_F} \rho_e(\tilde{E}_{2e}) d\tilde{E}_{2e} = \int_{-\infty}^{G-E_F} \rho_h(\tilde{E}_{2h}) d\tilde{E}_{2h}. \quad (19)$$

The densities  $n_e$  and  $n_h$  of the starting terms  $\varepsilon_{0e}$  and  $\varepsilon_{0h}$  in (19) can differ noticeably. This circumstance, however, can be neglected if  $E_F$  is located near the center of the mobility gap and the ratio of  $n_e$  and  $n_h$  is substantially smaller than  $\rho(w)/\rho(E_F) \sim g(k^{(0)})/g(k_G) \gg 1$ . Recognizing, in accord with the prevailing notions,<sup>1</sup> that the densities of states on the tails of the valence and conduction bands do not differ very greatly (say by not more than one or two orders of magnitude), we put  $n_e = n_h$ . Transforming in the integrands of (19) to the distribution  $g(k)$  and equating the integration limits, we obtain

$$E_F \approx \frac{G}{1+\lambda} \left( 1 + \frac{\lambda\varepsilon_{0e} - \varepsilon_{0h}}{G} + \frac{\lambda U_{ce} - U_{ch}}{2G} \right), \quad \lambda = \left( \frac{Q_h}{Q_e} \right)^2. \quad (20)$$

The error introduced by the approximation  $n_e = n_h$  is small to the extent that  $|\delta\tilde{E}_{2e} + \delta\tilde{E}_{2h}| \ll 2G$  is small. At  $\lambda \approx 1$  and at the characteristic values indicated above for the remaining parameters, expression (20) yields  $E_F \approx G/2$  in agreement with the experimental results. On the other hand, if  $\lambda$  differs substantially from unity,  $E_F$  is located near one of the edges of the mobility gap.

The condition  $\lambda \approx 1$ , which agrees within the framework of the described approach with the empirical data, calls for special comments. It is obvious that although  $Q_e$  and  $Q_h$  are of the same order, they are in general still noticeably different (say, by 1.5–2 times, in analogy with the constants of the deformation potential). Therefore the condition  $\lambda \approx 1$  can hold only in special cases. This, in particular, is the situa-

tion if the electrons and holes are localized on terms of opposite parity, antibinding and binding. The data on the band structure of chalcogenide glasses<sup>11</sup> suggests that such a case is quite realistic. The condition  $\lambda \approx 1$  is satisfied also if both types of carrier are localized on identical terms. In the latter case the filling of a term by two electrons corresponds qualitatively to formation of an assembled valence bond, whereas filling with two holes corresponds to its complete destruction. In this case the presence of one electron on the bond is not energetically profitable. The second of the noted possibilities agrees with the premise whereby the NEC is identified with glass-structure defects (see §1), for which either formation of a fully assembled valence band or its absence is energetically favored. We shall hereafter not particularize the analysis by choosing a specific situation that ensures the relation  $\lambda \approx 1$ , but confine ourselves on the basis of the foregoing to the statement that such situations are perfectly feasible.

### §5. ELECTRONIC TRANSITIONS BETWEEN LEVELS IN THE MOBILITY GAP

If account is taken of only the states with  $k \approx k_G$ , which are the most effective with respect to density, the results are summarized by the scheme of Fig. 1a, which is no longer assumed on empirical grounds but is the consequence of the developed approach. At the same time, in certain respects our approach does not reduce to the discrete-level scheme of Fig. 1a. The presence of a continuous distribution of the quasi-elastic constants  $g(k)$  and of a corresponding distribution of the levels leads to a number of distinguishing features.

Consider by way of example certain features of carrier recombination. Let the electron and hole be localized on spatially separated centers with quasi-elastic constants  $k_e$  and  $k_h$  (in particular, one or both carriers can correspond to nonequilibrium NEC states with  $n = 1$ ). Their total energy is

$$E_{eh} = E_{eh}^{(0)} + \frac{1}{2} k_e x_e^2 - Q x_e + \frac{1}{2} k_h x_h^2 - Q x_h, \quad (21)$$

$$E_{eh}^{(0)} = G - e^2 / \kappa R + \varepsilon_{0e} + \varepsilon_{0h},$$

where  $\kappa$  is the dielectric constant and  $R$  is the distance between centers. After the recombination the energy is equal to

$$E_0 = \frac{1}{2} k_e x_e^2 + \frac{1}{2} k_h x_h^2. \quad (22)$$

From this we obtain the value of the recombination activation barrier

$$\Delta E = \frac{\tilde{\kappa}}{2Q^2} \left( E_{eh}^{(0)} - \frac{Q^2}{\tilde{\kappa}} \right)^2 = \frac{(\hbar\omega)^2}{2(E_{eh}^{(0)} - \hbar\omega)}, \quad (23)$$

$$\tilde{\kappa} = (k_e^{-1} + k_h^{-1})^{-1},$$

where  $\hbar\omega$  is the energy of the quantum that might be emitted in this process. It can be seen that the deeper the location of the carriers in the mobility band, the smaller  $\Delta E$  and  $\hbar\omega$  and the more probable is radiative recombination.

It can be concluded from the foregoing that at small  $\hbar\omega$  the luminescence intensity  $I(\hbar\omega)$  decreases rapidly with decreasing  $\hbar\omega$ . In addition,  $I(\hbar\omega)$  decreases with increasing  $\hbar\omega$

in the region of large  $\hbar\omega \lesssim G_0$ , owing to the decrease of the population of the levels as their energies approach the edges of the mobility gap. The reason is that the nonequilibrium carriers produced by the interband light  $\hbar\omega > G_0$  lose their energy rapidly, in transitions of the type of diffusion in energy space, to centers with deeper levels, i.e., with smaller  $k$ . We note also that the density of deep nonequilibrium NEC states with  $k \approx k_G$  and  $n = 1$  builds up as a result of similar transitions. These states arise in the interaction between the doubly filled NEC states with  $n = 2$  and carriers of opposite sign, and are effective recombination states.

The foregoing leads to the following qualitative picture of the processes that occur when glass is irradiated with interband light. If  $\hbar\omega_0$  exceeds considerably the width of the optical gap, the excess energy  $\hbar\omega_0 - G_0 = \Delta\hbar\omega_0$  is sufficient for spatial separation of the produced electron and hole. The subsequent lowering of their energy terminates as a rule in nonradiative recombination, and the luminescence intensity is not high. At small  $\Delta\hbar\omega_0$  the Coulomb interaction hinders the separation of the produced carriers, so that the effectiveness of their independent energy diffusion due to transition in real space is decreased. The pair, however, can move as a whole. One of the pair carriers has a noticeable probability of decreasing in energy and landing at the center with  $k \approx k_G$ , whereas the transitions of the second carrier take place between centers with  $k \approx k^{(0)}$  (the probability of finding simultaneously two centers with sufficiently small  $k$  in the volume occupied by the pair is low). The result is a system that includes a singly populated center with  $k \approx k_G$  and a carrier of opposite sign having an energy close to the edge of the mobility gap. Its radiative recombination corresponds to an energy  $\hbar\omega \approx G/2$  [see (23)]. Thus, as  $\Delta\hbar\omega_0$  decreases the luminescence intensity  $I(\hbar\omega)$  at a given  $\hbar\omega$  first increases because of the suppression of the energy diffusion processes, and then decreases because of the weakening of the absorption at  $\hbar\omega_0 < G$ . In other words, the luminescence excitation spectrum has a maximum at  $\hbar\omega_0 \approx G_0$ . The spectrum of the luminescence itself  $I(\hbar\omega)$  is a band with a maximum at  $\hbar\omega \approx G/2$  and of width determined by the joint effects of the strong relaxation of the atomic subsystem and the distribution probability  $g(k)$ . We note that as the nonequilibrium NEC states  $n = 1$  with  $k \approx k_G$  accumulate in the course of the illumination, the nonradiative recombination connected with the presence of these states becomes more and more effective, i.e., the luminescence intensity should decrease with increasing irradiation (photoluminescence fatigue). In addition, the accumulation of nonequilibrium NEC states  $n = 1$  leads to photon absorption with  $\hbar\omega \approx G/2$  (induced photoabsorption) and to the appearance of an EPR signal (induced EPR).

The foregoing regularities agree with the experimental results of investigations of chalcogenide glasses (see Ref. 1). We emphasize that the quasicontinuous spectrum of the states in the mobility gap manifests itself outwardly in the described processes as an effectively discrete spectrum (e.g., by appearance of photoluminescence or induced-photoconductivity bands with  $\hbar\omega \approx G/2$ , agreeing in this sense with the scheme of Fig. 1a.

The essential role of the quasicontinuous spectrum in-

side the mobility gap is confirmed by results of the investigation of the kinetics of the conductivity of chalcogenide glasses.<sup>12</sup> In Ref. 12 this form of the spectrum was postulated in contradiction to the usually assumed discrete-level scheme of Fig. 1a. This contradiction is eliminated within the framework of our approach.

An interesting consequence of the concepts expounded is the possibility of temporal evolution of photoluminescence spectra. Carriers excited by a short light pulse recombine during the first instants of their existence and emit photons with  $\hbar\omega \approx E_{eh}^{(0)} \approx G$ . Next, with evolving energy diffusion, the maximum of the photoluminescence spectrum shifts towards smaller  $\hbar\omega$  and becomes fixed in practice at  $\hbar\omega \approx G/2$ . Such a picture was indeed observed in recent experiments.<sup>13</sup>

## §6. CONCLUSION

The essential result of the present paper is the identification of the nature of the observed connection between the energy characteristics of carriers self-trapped on NEC and the mobility gap of the material. It turns out here that the quasicontinuous spectrum of the states inside the mobility gap can manifest itself as effectively discrete, corresponding to the scheme previously established on an empirical basis.

Conceptually, the described approach is closer to Anderson's ideas<sup>2</sup> than to viewpoint of Mott *et al.*,<sup>1</sup> who set the NEC in correspondence with definite defects of the glass structure. This applies in particular to the conclusion that the paired states have a quasicontinuous spectrum in the mobility gap. In contrast to Ref. 2, our approach specifies the nature of this quasicontinuity and attributes it to fluctuations of local quasi-elastic constants of the amorphous material. It is this which explains the empirical relations between the parameters of the carriers localized on the NEC. At the same time, our analysis does not contradict the "defect" approach<sup>1</sup> (see §4), if it is assumed that the quasi-elastic constants are characterized near the corresponding defects by a broad probability distribution of the type shown in Fig. 2. In this case the structural models of defects<sup>1</sup> only specify more concretely the physical meaning of the configuration coordinate  $x$  in our analysis.

We note that the results obtained above can be applicable to a certain degree not only to amorphous substances, but also to crystals in which the critical atomic potentials are realized as defects. Among the various structure faults always present in a crystal, defects with critical potentials have the advantage that part of the energy needed for their formation is cancelled out upon self-trapping of a carrier pair. The nonlinear behavior of  $\varepsilon(x)$ , of the type shown in Fig. 3, can

then play an important role. Because of this behavior, the energy gain due to self-trapping should be a maximum for those defects whose critical potentials admit of a shift of the single-electron terms by precisely the value of the linear section of  $\varepsilon(x)$ , i.e., actually to the boundary of the opposite band, and it is precisely such effects that should occur with maximum densities. The energy spectrum of such defects with  $k \approx k_G$  and  $|\bar{E}_2| \approx G/2$  in the forbidden band of a crystal turns out to be practically discrete. It can be assumed that this situation is realized in chalcogenide crystals. Indeed, the luminescence spectra of amorphous and crystalline  $\text{As}_2\text{S}_3$  and  $\text{As}_2\text{Se}_3$  are similar, but the excitation spectra are different, and no fatigue effect is observed in the crystals.<sup>1</sup> In our approach the latter discrepancies are quite understandable. They reflect the qualitative differences between recombination processes in the quasicontinuous spectrum of glass and in the discrete spectrum of a crystal.

<sup>1</sup>This assumption is not necessary. It is shown in Ref. 4 that for independent fluctuations of the expansion coefficients of  $V(x)$  with  $k \ll k^{(0)}$  the anomalous smallness of the cubic anharmonicity is still substantially more probable than simultaneous smallness of the quasi-elastic constants for two modes of the atomic motion. Thus, in the one-mode model of the critical potential it is legitimate to take into account fluctuations for which the cubic anharmonicity is anomalously small alongside with  $k$ , and the presence of systematic causes of this smallness is not obligatory.

<sup>2</sup>The nonlinear behavior of  $\varepsilon(x)$  in crystals, of the type shown in Fig. 3, can be due to singularities of the band structure. The effect of slowing down  $\varepsilon(x)$  should be particularly noticeable in layered crystals such as  $\text{As}_2\text{S}_3$  and  $\text{As}_2\text{Se}_3$ , where the electron spectrum retains the features of quasi-two-dimensional systems, i.e., the density of states falls off quite steeply near the boundaries of the forbidden band; in this case  $\Delta_0 \sim G(G/D)^\nu$  with  $\nu > 1$ , i.e.,  $\Delta_0 \ll G$  (see (5)).

<sup>1</sup>N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, Oxford, 1971.

<sup>2</sup>P. W. Anderson, *Phys. Rev. Lett.* **34**, 953 (1975).

<sup>3</sup>M. I. Klinger and V. G. Karpov, *Zh. Eksp. Teor. Fiz.* **82**, 1687 (1982) [*Sov. Phys. JETP* **55**, 976 (1982)].

<sup>4</sup>V. G. Karpov, M. I. Klinger, and F. N. Ignat'ev, *Zh. Eksp. Teor. Fiz.* **84**, 760 (1983) [*Sov. Phys. JETP* **57**, 439 (1983)]. *Sol. St. Commun.* **44**, 333 (1982).

<sup>5</sup>P. W. Anderson, B. I. Halperin, and C. M. Varma, *Phil. Mag.* **25**, 1 (1972).

<sup>6</sup>W. A. Phillips, *J. Low Temp. Phys.* **7**, 351 (1972).

<sup>7</sup>*Amorphous Solids. Low Temperature Properties*, W. E. Phillips, ed., Springer, 1981.

<sup>8</sup>G. A. Baraff, O. A. Kane, and M. Schlüter, *Phys. Rev.* **B21**, 5662 (1980).

<sup>9</sup>F. D. M. Haldane and P. W. Anderson, *Phys. Rev.* **B13**, 2553 (1976).

<sup>10</sup>A. Milns, *Deep Impurities in Semiconductors*, Wiley, 1973.

<sup>11</sup>I. Chen, *Phys. Rev.* **B8**, 1440 (1973). H. L. Althaus, G. Weiser, and S. Nagel, *Phys. St. Solid.* **B87**, 117 (1978).

<sup>12</sup>G. Orenstein and M. A. Kastner, *Sol. St. Commun.* **40**, 85 (1981).

<sup>13</sup>G. S. Higashi and M. A. Kastner, *Phys. Rev.* **B24**, 2295 (1981).

Translated by J. G. Adashko