Problems of the electron theory of disordered semiconductors

V. L. Bonch-Bruevich

M. V. Lomonosov Moscow State University Usp. Fiz. Nauk 104, 583-637 (August 1983)

A number of studies is reviewed that has been conducted in recent years on three current problems of the theory of disordered semiconductors: a) The nature of the electron states in systems of different dimensionalities (including the problem of the so-called "power-law" ("weak") localization). b) Some multielectron effects in disordered semiconductors (including the problem of the Coulomb gap). c) Absorption of light with quantum energies smaller than the optical width of the forbidden band (including an analysis of the role of macroscopic inhomogeneity of thin films).

PACS numbers: 71.25.Mg, 71.50. + t

Introduction	664
I. Nature of the electron states in disordered semiconductors of different dimensionalities	665
1. Introduction. Criterion of localization. 2. Scaling theory of localization.	
3. Limiting forms of the functions $\beta(g)$. Thermal and nonohmic effects. 4. Qualitative	
study of Eq. (2.7). Localization of states for $d = 1, 2, 5$. The problem of the minimal	
metallic conductivity $(d=3)$. 6. Power-law localization in a two-dimensional system.	
Il. Multielectron effects in the physics of disordered systems	676
7. Introduction. 8. Interaction between delocalized electrons or only impurity	
scattering? 9. The problem of the Coulomb gap.	
III. The "tail" of the interban light-absorption coefficient in a disordered semiconductor	681
10. Introduction. 11. Phonons and static disorder. 12. The random field and the	
Urbach rule. 13. The Urbach rule in amorphous silicon.	
Appendices	688
I. Corrections to the results of the kinetic equation for $T \rightarrow 0$. II. On the rule of the	
temperature-dependence of the density of states at low temperatures.	
References	692

INTRODUCTION

The physics of disordered systems-amorphous and vitreous semiconductors, amorphous magnetic materials, metallic and nonmetallic alloys, etc .- is now enjoying heightened attention. As often happens in such cases, there are two reasons for this. One of them involves the inner logic of development of science-the current stage of development of the physics of condensed media is characterized by transition from crystalline materials to materials whose atomic structure lacks long-range order. The other reason arises from the demands of modern electronics. Since B. T. Kolomiets and N. A. Goryunova^{1,2} discovered the semiconductor properties of chalcogenide glasses, photoelectronics has had at its disposal a new class of highly interesting materials whose practical value is now generally known. In recent years amorphous silicon has moved to the forefront (especially its alloys with hydrogen or fluorine). Owing to its unique characteristics (convenient width of the forbidden band, high dark resistance, and long lifetime of the minority charge carriers), this material appears highly promising for applications in a number of photoelectric devices, including solar cells. Half of all the papers at the 9th International Conference on Amorphous and Liquid Semiconductors (Grenoble, July 1981) were devoted to different aspects of the physics, technology, and application of this material.

A number of review articles has been published in recent years, devoted to both fundamental and applied aspects of the physics of disordered semiconductors.³⁻²² However, this branch of physics has developed at such a tempo that not a single review has managed to keep pace with it. Consequently new reviews appear; they share the fate of the preceding ones, and it is not yet clear whether this series converges at all.

This review is designed more for experimentalists than for theoreticians. To a certain extent it can serve to supplement the book,²³ which was published in 1981, but was submitted for publication in 1978. In line with the style of this textbook, the list of references here also does not claim completeness; the references to the original studies are given only in limited amount, since they are essentially not of priority-setting type, while the references to the reviews mainly encompass only the studies published no earlier than 1977.¹⁾

We shall study three problems that have attracted great attention recently.

¹Note added in proof (January 1983). As should have been expected, since the time that this review was submitted for publication, a considerable number of studies has appeared on the same topic. Even brief annotations of them would considerably enlarge the scope of the review, which has already grown profusely. Some references have still been made in the appropriate place in the review (Sec. 2).

1) The nature of the electron states in disordered semiconductors of different dimensionalities (the problem of localization).²⁾

2) The role of interelectron interaction in forming the energy spectrum of the electrons (the problem of the Coulomb gap).

3) The frequency-dependence of the light-absorption coefficient in the "tail"—the region of quantum energy smaller than the width of the forbidden band.

The first of these problems is relatively new, the second already has a relatively long life (about 11 years), but is still a topic of discussion, while the third problem has recently become again very current in connection with the experimental studies on amorphous silicon.

I. NATURE OF THE ELECTRON STATES IN DISORDERED SEMICONDUCTORS OF DIFFERENT DIMENSIONALITIES

1. Introduction. Criterion of localization

A condensed system of particles is termed d-dimensional (d=1,2,3) if the electrons, holes, or other elementary excitations under the conditions of interest to us can move freely in d dimensions. Thus, for example, organic crystals based on TCNQ, now so fashionable, are one-dimensional systems, while an inversion layer arising at the surface of a semiconductor becomes a two-dimensional system when the bending of the bands is great enough. This makes it clear that the study of the behavior of elementary excitations (henceforth we shall speak only of electrons and holes) in systems of different dimensionalities is not at all of only abstract theoretical interest. In particular, the problem of how disorder caused, e.g., by radiation affects the characteristics of thin-film devices reduces to this problem.

The energy spectrum of electrons and holes in weakly doped³⁾ three-dimensional crystalline semiconductors has been well studied: we have here two regions of continuous spectrum—the conduction and hole bands—and a forbidden band separating them, in which only individual discrete levels involving impurities are allowed. The regions of the continuous spectrum correspond to delocalized wave functions—the probability of finding an electron (or hole) in any unit cell of the crystal differs from zero and is the same in all cells. The states of the continuous spectrum are current-bearing states: the electrons and holes occupying them can participate in charge and energy transport at arbitrarily low temperatures, even in the absence of illumination. On the other hand, the discrete levels possess wave functions localized around the corresponding impurity atoms (randomly situated). The states of the discrete spectrum in a specimen of arbitrarily large dimensions (macroscopic) are current-free states: the electrons (or holes) occupying them can participate in charge and energy transport throughout the entire macroscopic specimen only in the presence of thermal or optical activation (hopping conductivity). When $T \rightarrow 0$ and in the absence of illumination, the contribution of the localized electrons to the charge and energy fluxes vanishes.

With increasing impurity concentration, the pattern of the energy spectrum in the forbidden band becomes complicated. However, the separation into the regions of continuous and discrete spectrum, which correspond to current-bearing and current-free states, remains in force here also (Ref. 23, Chap. IV).

These results, which stem from a large set of experimental data on the electrical and optical properties of matter, are obtained naturally in the electron theory of crystals. Here one need not at all formulate the latter as a one-electron theory: the statements expressed above have also an exact multielectron meaning (Ref. 24, Chap. XVII). The conclusion that the spectrum has a band structure remains in force also for ideal crystals of lower dimensionality.

However, the situation changes as soon as we go to real crystals containing impurities randomly distributed in space, and all the more so as we go to amorphous, liquid, or vitreous semiconductors. When d = 3, the pattern of the energy spectrum here still resembles to a certain degree that existing in crystals. Actually, the experimental data^{25,23} irrefutably indicate the existence in three-dimensional disordered semiconductors of regions of a continuous spectrum of electrons and holes that correspond to current-bearing states. These regions are analogous to the conduction and hole bands (as they are often called), but in contrast to ideal or almost ideal crystals, here one cannot speak of dispersion laws of electrons and holes: in the absence of spatial periodicity in the arrangement of the atoms, one cannot introduce the components of the quasimomentum as "good" quantum numbers. There is a region between the conduction and valence bands occupied by discrete levels corresponding to currentfree states of electrons and holes. However, in contrast to weakly doped crystals, these levels can lie arbitrarily close to one another in energy (a discrete spectrum everywhere dense).⁴⁾ For this reason, their energy distribution can be conveniently characterized by a continuous function, the density of states $\rho(E)$, which indicates the mean number of these levels per unit volume and per unit energy interval about a given point E (according to Ref. 23 this is the smoothed density of states averaged over all configurations of the random field). In disordered materials (in contrast to ideal or almost ideal crystals), the density of states does not

²⁾An interesting review³ has been devoted to this problem. However, it was designed mainly for theoreticians. Moreover, already more than two years have passed since its publication.

³⁾The term "weakly doped" is understood in the sense indicated in the book of Ref. 24. The word "impurity" can denote any point defect of the structure. We shall not treat problems here involving extended defects in order not to complicate the presentation with details that will not be essential below.

⁴⁾The discrete character of the levels is maintained because states close in energy generally correspond to localization centers situated far apart.

vanish at the boundaries of the conduction and valence bands, but extends rather deeply into the forbidden band to form "tails" (first studied in the theory of heavily doped semiconductors; see Ref. 24, Chap. XIX). In this situation the very term "forbidden band" becomes not entirely fitting; instead one often speaks of the "mobility gap", while the bottom of the conduction band $E = E_c$ and the top of the valence band $E = E_v$ denote the mobility thresholds.

The reason for the appearance of tails of the density of states is well known. The point is that the random nature of the arrangement of atoms (impurity atoms, or in the case of amorphous, liquid, and vitreous semiconductors, the atoms of the main material) leads also to a random character of the potential energy of the electrons. In particular, potential wells of varying depth and shape arise that correspond to local levels of different energies. These levels are called fluctuational levels. A nontrivial feature of the system is that these levels remain discrete, despite their perhaps large concentration. This is called Anderson localization.²⁶

The difference of the density of states from zero at the mobility threshold leads to posing of the question of how the static conductivity σ of the specimen (as T - 0) behaves when the Fermi level F approaches the mobility threshold from the side of the continuous spectrum.

This question is not at all abstract in nature, since one can control the position of the Fermi level at a given temperature by varying, e.g., the potential V_g at the gate of a field transistor (Fig. 1). A metal (e.g., aluminum) is shown in Fig. 1; also some other dielectric can be used instead of silicon dioxide.

In principle there are two possibilities here, schematically drawn in Fig. 2, a-c: the conductivity can either reach some finite value σ_m as $F - E_c + 0$, and then fall to zero in a jump (Fig. 2a), or it can fall to zero continuously, while the derivative $d\sigma/dF$ can either remain bounded as $F - E_c + 0$ (Fig. 2b), or become infinite (Fig. 2c). An especially essential point is the difference between the situations shown in Fig. 2a, on the one hand, and Figs. 2b, c on the other. The quantity σ_m has been called the minimum metallic conductivity. It can be determined theoretically under certain assumptions. It seemed at one time that the experimental problem was almost solved in favor of the conductivity shown in Fig. 2a. However, in the past two years the problem of the



FIG. 1. Standard diagram of an experiment to measure the electrical conductivity of an inversion layer. The voltage V_{ξ} controls the position of the Fermi level in the semiconductor.



FIG. 2. Possible types of behavior of the conductivity at the mobility threshold.

minimum metallic conductivity has been the object of serious discussion: we shall return to it in Sec. 5.

We note that, under conditions of strong degeneracy of the electron gas, σ is effectively determined in the conduction band only by the electrons of energies close to *F*. This enables one to introduce the concept of the conductivity (and mobility) for a given electron energy *E*, by taking it to be represented by the functions $\sigma(F)$ [or $\mu(F)$]. One can keep this concept also at any arbitrary temperature and arbitrary degree of degeneracy (including also a nondegenerate gas) by assuming that by definition

$$\sigma(T) = \int_{E_c}^{\infty} \sigma(E, T) \left(-\frac{\partial n_F}{\partial E}\right) dE.$$
(1.1)

Here $n_{\rm F}(E)$ is the Fermi function, $\sigma(E, T)$ is a function of the temperature and energy, while the upper limit has been replaced by infinity owing to the rapid decline of $n_{\rm F}$ with increasing energy argument E. When T = 0, Eq. (1.1) acquires the form $\sigma(0) = \sigma(E)_{E*F}$. In a oneelectron formulation of the problem, the relationship (1.1) directly stems from the formula of Kubo,²³ with $\sigma(E, T)$ expressed explicitly in terms of the corresponding Green's functions. One can also write an analogous formula for holes. The concept of a conductivity dependent on the energy enables one to formulate a criterion of localization of electron states based only on using quantities directly measured experimentally.27 Namely, one can naturally call states of energy E delocalized (or localized) if, as $T \neq 0$, the static electrical conductivity o(E) differs from zero (or equals zero). We shall henceforth treat the following equation as the criterion of localization:5)

$$\lim_{T \to 0} \sigma(E, T) = 0.$$
 (1.2)

However, we should note that the criterion (1.2), while theoretically irreproachable, appears inconvenient for numerical calculations. The point is that in such calculations one always—of necessity—treats systems of limited dimensions. But, strictly speaking the concept of electrical conductivity as a self-averaging quantity²³ is justified only in the case of a transition to the thermodynamic limit. It is not fully clear to what extent the potentialities of modern computers allow one to approximate the latter operation. For this reason other criteria have been proposed²⁹⁻³² in the hope that they will prove more convenient computationally. Apparently their equivalence to the criterion (1.2) has not been strictly proved, but it seems rather likely.

⁵³Apparently it is equivalent to the criterion proposed in Ref. 28.

The existence of potential wells and humps of random height, width, and shape distributed at random in the specimen leads to especially important consequences in the case of specimens of lower dimensionality. The idea had already been expressed^{33,34} rather long ago that all states of an electron in a one-dimensional disordered lattice are current-free, i.e., there is no mobility threshold. A rigorous proof of this statement was first derived in Refs. 35 and 36. In particular, it yields rather serious consequences in the physics of one-dimensional "crystals", and also in the physics of crystals with one-dimensional defects, e.g., dislocations. It is known³⁸ that edge and screw dislocations having no impurity atmosphere lead to the appearance of one-dimensional dislocation bands. In a crystal ideal in other respects, the charge carriers occupying these bands, being localized near the dislocations, can move freely along them-in this sense they are delocalized. However, since their motion is one-dimensional, an arbitrarily weak random field will lead here to complete Anderson localization. Correspondingly, in real materials with a low concentration of dislocations, the conductivity in the dislocation bands at low temperatures must be of a hopping type.⁶⁾ The same is true for long organic molecules. A detailed theory of hopping conductivity of one-dimensional systems has been constructed in recent years³⁹⁻⁴² (see also the review⁴³).

For a long time the pattern of the energy spectrum in a two-dimensional disordered semiconductor remained not entirely clear. Now the opinion has become widespread that, just as for d=1, all electron states in these materials are current-free (in any case, if the dimensions of the specimen are large enough). This problem will be treated in greater detail in the next section.

2. Scaling theory of localization

The shift from the regime characteristic of the case $F > E_c$ to the regime $F < E_c$ is to a certain extent analogous to a phase transition. Hence it is not surprising that attempts have been made to study the problem of presence or absence of a mobility threshold in disordered systems by using the methods of the modern theory of critical phenomena (Ref. 44).⁷⁾ Apparently this formulation of the problem was first proposed by F. Wegner;⁴⁶⁻⁴⁸ see also Ref. 49. An elegant approach using the renormalization group has been proposed.^{50,51} However, here the equations of the renormalization group obtained essentially by guessing. One can find a systematic derivation of them within the framework of a rather general model in Ref. 52.

The presentation of the material in this section follows to a considerable extent the method of Refs. 50 and 51, although differing in individual points.

The central idea of the scaling theory of localization consists of the following. For a specimen of finite dimensions one defines a certain quantity associated with the conductivity, and generally dependent on the dimensions of the specimen. Then one studies the behavior of this quantity with increasing dimensions of the specimen and ultimately carries out a transition to the thermodynamic limit. Apparently the problem of choosing this quantity is not completely trivial. The first studies^{50,51} adopted as such a quantity the reciprocal resistance (conductance), which can be connected with the mean shift of the energy levels upon replacing the periodic boundary conditions at the "ends" of the specimen with antiperiodic conditions. (Calculation²⁹ shows that this quantity is expressed by the same quantity as the ordinary reciprocal resistance, R^{-1} , which appears in Ohm's law.) We can naturally expect that the energy levels corresponding to the localized states should practically not be altered by this substitution. On the other hand, the eigenvalues of the energy corresponding to the delocalized wave functions must be shifted by this substitution, however large are the dimensions of the specimen. The criterion of localization proposed in Ref. 29 is based on this argument.

< 6

Subsequent studies^{53,54} have proposed the reciprocal localization length as the quantity undergoing the scaling transformation. However, it appears to be more natural to use a quantity directly measurable experimentally. As such one can still use the reciprocal resistance of the specimen as defined by Ohm's law. Apparently the use of this self-averaging quantity in physically realistic systems need not lead to the complications pointed out in Ref. 53.

We should expect that, as the dimensions of the specimen increase without limit, its shape will not play an essential role in the problem of interest to us. Hence we are justified in treating a specimen of any shape, e.g., cubic (or when d=2, square; henceforth for brevity we shall use the term "cube" for any dimensionality). Here the standard formula for the resistance of the specimen has the form

$$R^{-1} = \sigma L^{d-2}.$$
 (2.1)

Here L is the length of the cube edge.

It is convenient to introduce the dimensionless reciprocal resistance g by setting

$$g = R^{-1} \frac{2\hbar}{s^2} = \frac{2\hbar\sigma}{s^2} L^{d-2}.$$
 (2.2)

In employing this formula, we must consider the fact that the length L has a lower bound: it must exceed the other characteristic lengths that appear in the problem under discussion. There are at least three of them. The first of them arises in treating phenomena of transport by free charge carriers, i.e., in the region of the continuous energy spectrum. This is the mean free path with respect to momentum; under the conditions of interest to us it is also the mean free path for elastic collisions. The second characteristic length appears (or its presence is implicitly understood) when one treats transport phenomena in the region of the discrete spectrum. This is the quantum correlation length R_a , which

⁶^WWe note that the hopping nature of the conduction does not necessarily entail a strong temperature-dependence of it (of exponential type). Cases are known²⁰² in which conduction of clearly hopping type depended on the temperature according to a power law.

⁷⁾We can recommend the book of Ref. 45 for an initial acquaintance with this theory.

defines the smallest possible distance between levels of the same energy. The point is that, as $T \rightarrow 0$, charge transport in a given region of energies can occur only by isoenergetic tunneling. However, in view of the effect known from quantum mechanics of "repulsion of levels", levels close in energy directly correspond to wave functions localized near points sufficiently remote from one another (Ref. 23, Chap. III, Sec. 3). If the levels did not possess an intrinsic width δ , then the length R_0 would become infinite—the specimen could not contain two identical discrete levels with a finite distance between the corresponding centers of localization.⁸⁾ However, actually the value of δ always differs from zero, at least for levels lying above the Fermi level. There are two reasons for this. One of them stems in principle from the irremovable interaction of the electrons with the electromagnetic vacuum (as in the atom) and with the heat bath,⁵⁵ the other arises from the current passing through the specimen.⁵⁶⁻⁵⁸ Indeed, the very fact of passage of current implies that the levels, which are stationary in its absence, become nonstationary when it arises. That is, they acquire a finite width. When $\delta \neq 0$, the length R_0 proves to be finite. It is not easy to derive an exact expression for it; However, one can obtain an orientational estimate by using the very simple argument of Mott (Ref. 25, Chap. 2). For levels close to a Fermi level situated in the mobility gap, we have

$$R_0 = \frac{1}{2\alpha} \ln \frac{V}{\delta} ; \qquad (2.3)$$

Here α is the reciprocal localization length (assumed to be the same for both levels being studied, in view of their closeness), and V is the characteristic energy of interaction of an electron localized near one of the centers with the other center.⁹⁾

Evidently, isoenergetic tunneling can occur only if $L \ge R_0$. For an orientational estimate of R_0 , we shall assume that

$$\alpha = 10^{5} \text{ cm}^{-1}, \quad \tau^{-1} \equiv \delta/\hbar = 10^{4} \text{ s}^{-1}, \quad V = 1.4 \text{ eV}.$$

Then $R_o = 10^{-4}$ cm. Apparently this value is somewhat excessive, since the values of α and τ^{-1} that we have used are too low. Yet it gives a certain orientation. We note that the third characteristic length—the correlation length of the random field γ^{-1} —can prove to be of the same order of magnitude (or larger). It defines the distance at which the binary correlation function of the fluctuations of the potential energy of an electron U declines appreciably:

$$\mathbf{I}'(\mathbf{r}) = \langle U(0) | U(\mathbf{r}) \rangle; \qquad (2.4)$$

Here the angle brackets denote averaging over the random field, $\langle U \rangle \approx 0$. In particular, if the random field is created by a set of charged impurities situated at random in the specimen, then the screening radius plays the role of γ^{-1} . In a system like amorphous silicon it can amount to about 10^{-3} cm.

Henceforth we shall be dealing with specimens of dimensions considerably exceeding both R_0 and $\gamma^{-1,10\rangle}$ Here the latter lengths need not enter into the expression for the reciprocal resistance, and the equation for the function g(L) can be written by the same method as in Refs. 50 and 51.

Let us take b^d cubes, each of edge length L, and mentally build from them a new cube of edge length bL. This cube will correspond to a new value of the dimensionless reciprocal resistance g(bL). The fundamental assumption of the scaling theory consists of the idea that, under the conditions being discussed, the function g(bL) can depend only on its former value and on the number b:¹¹⁾

$$g(bL) = f[b, g(L)].$$
 (2.5)

Now let us assume that

$$b = 1 + \varepsilon, \quad \varepsilon \to 0,$$

. .

and expand the function g(bL) in a power series in ε :

$$g(b, L) = g(L) + \varepsilon g(L) - \frac{d \ln g}{d \ln L} + O(\varepsilon^2).$$
(2.6)

After elementary transformations, we obtain:

$$\frac{d \ln g(L')}{d \ln L'} = \beta [g(L')].$$
(2.7)

Here L' is a variable which upon completing the calculation must be identified with the length L of the specimen, and we have

$$\beta[g(L')] = g^{-1}(L') \frac{df(b, g(L'))}{db} \Big|_{b=1}.$$
(2.8)

Equation (2.7), which has been derived in Refs. 50 and 51, is fundamental in what follows. It is autonomous its right-hand side contains the argument L' only implicitly—only via the sought function g(L'). In other words, Eq. (2.7) describes a one-parameter group of transformations—the renormalization group.¹²⁾ By

⁸⁾Strictly speaking, with a number of atoms greater than two, one can choose a configuration of them such that certain levels are not split. However, the probability of realizing such a configuration in a system of many particles is very small.

⁹⁾In the presence of a mobility threshold $E_{\rm cr}$, the energy V agrees in order of magnitude with the difference $E_{\rm c}-F$. Here one obtains from Eq. (2.3) the expression given in the book of Ref. 23. We note that Eq. (2.3) is based on the concept of an exponential decline of the wave functions at a great enough distance from the corresponding localization centers.

¹⁰⁾The problem of short specimens requires special study⁵⁵ that we shall not take up here. We note only that under these conditions the static conductivity can prove to differ from zero for any energy of the electron. It is more convenient to define the localized states here as those whose contribution to the static conductivity declines monotonically with increasing L for a sufficiently large dimension of the specimen.

¹¹⁾The role of the other lengths was not treated in Refs. 50 and 51. If they were comparable with L, then one would have to generalize Eq. (2.5) by also including the quantities LR_0^{-1} and γL among the arguments of the function f.

¹²)^{If} the lengths R_0 and γ^{-1} were comparable with L, then the equation analogous to (2.8) would prove to be nonautonomous, and the discussions would become somewhat complicated. ⁵⁵

studying the behavior of its solutions as $L \rightarrow \infty$, which corresponds to taking the thermodynamic limit, we can determine whether the static electrical conductivity of a large enough specimen is zero or not as $T \rightarrow 0$. That is, we can establish whether the wave functions of the charge carriers are localized or delocalized.¹³⁾

3. Limiting forms of the function $\beta(g)$. Thermal and nonohomic effects

In order to extract any information from Eq. (2.7), one must formulate a concept of the behavior of the function $\beta(g)$, i.e., ultimately of the behavior of the dimensionless reciprocal resistance as L varies.

First of all we note that the function β for finite values of *L* describes the grouping of a finite number of cubes of finite dimensions. For this reason, we can naturally expect that it will behave sufficiently regularly. This expectation is expressed in the following two hypotheses:^{50,51}

- a) The function $\beta(g)$ is continuous and differentiable.
- b) The function $\beta(g)$ is monotonic.

There is as yet no complete, rigorous proof of these hypotheses, but they seem quite plausible. We note also that their consequences for d = 2 have been confirmed also by the results of Ref. 52, which was carried out by another method.

As soon as we have adopted the hypotheses (a) and (b), for a qualitative study of Eq. (2.7) it suffices to determine only how the function $\beta(g)$ behaves in the limiting cases $g \gg 1$ and $g \ll 1$. In the former case we have relatively weak scattering; this can be the situation in a system of free charge carriers. In the limit as $L \to \infty$, this corresponds to the region of the continuous energy spectrum, i.e., to delocalized electrons. The latter case corresponds to a system of localized electrons, i.e., the region of the discrete energy spectrum.

First let us examine the case $g \gg 1$. Here we can use the ordinary Boltzmann kinetic equation for calculating g. Here the dependence of σ on L generally vanishes in the limit of $L - \infty$. However, when we take into account the corrections to the results of the kinetic equation, such a dependence arises. As we shall soon see, it proves substantial in treating a two-dimensional system. The appropriate calculations have been performed^{59,60} under the assumption that the electrons are scattered by impurities having short-range forces. For d=2 the answer is

$$\sigma = \sigma_{\rm B} - \frac{e^3}{2\pi^2\hbar} \ln \frac{L}{l_{\rm p}}, \qquad (3.1)$$

Here $\sigma_{\rm B}$ is the result obtained from the kinetic equation, while $l_{\rm p}$ is the mean free path with respect to momentum. Formula (3.1) holds as long as the first term

on the right-hand side is large in comparison with the second, i.e., as long as $g \gg 1$; a schematic derivation of this is given in Appendix I.¹⁴)

< h

In the general case for delocalized states we obtain (see Appendix I)

$$\beta = d - 2 - \frac{c_d}{\sigma}, \quad g \gg 1. \tag{3.2}$$

As before, *d* denotes here the dimensionality of the system, while the c_d are positive constants; in particular, when only potential scattering exists (without spin reversal), we have $c_2 = \pi^{-2}$.

Now let us turn to the case $g \ll 1$. The probability of an isoenergetic tunneling transition through a distance $R \ge R_0$ is proportional to $\exp(-2\alpha R)$. If $L \gg R_0$, such transitions can occur at distances considerably less than L; charge transport throughout the specimen will occur only as the result of a chain of isoenergetic transitions of this type, while the resulting value of g will be determined by considerations of percolation and will depend only relatively weakly on L. At the same time, tunneling transitions can also occur "throughout the specimen"; their probability is proportional to $\exp(-2\alpha L)$. Thus, when $g \ll 1$, we have⁵⁶

$$g(L) = g_0(L) + g_1(L) \exp(-2\alpha L).$$
 (3.3)

The first term here corresponds to the chain of transitions pointed out above, while the second term corresponds to tunneling percolation of electrons throughout the specimen.

In line with what we have said above, the functions g_0 and g_1 depend on L far more weakly than $\exp(-2\alpha L)$. One can neglect their derivatives with respect to L in the following differentiation. In other words, they can be treated as parameters in the equation of the renormalization group. At the same time, the very fact of dependence of g_0 on L can be important in another respect: if, as we shall soon see, in a two-dimensional system we have g(L) = 0 as $L \to \infty$, then the quantity $g_0(L)$ must vanish here.

According to Eqs. (2.8) and (3.3), we now obtain

$$\beta = -\frac{2\alpha Lg_1 \exp\left(-2\alpha L\right)}{g_0 + g_1 \exp\left(-2\alpha L\right)} = -\frac{g - g_0}{g} \ln \frac{g_1}{g - g_0}.$$
 (3.4)

Since the quantities g_0 and g_1 appear here as parameters, this relationship agrees with the assertion of the autonomousness of Eqs. (2.7).

In the next section we shall employ Eqs. (3.3) and (3.4) for a qualitative study of the function g(L). Here it is appropriate to make some remarks on the temperature- and field-dependence of the electrical conductivity of a two-dimensional system at a low, but finite temperature (one can be dealing, e.g., with an inversion layer arising under certain conditions near the contact of a semiconductor with a dielectric in a metal-

¹³Generalizations were derived in Refs. 208 and 209 of Eq. (2.7) to the case of conduction in an alternating field and to the question of taking into account the interaction between the charge carriers. Naturally the theory here becomes complicated, but the fundamental conclusions presented in the following subsections remain in force.

¹⁴)An analogous result⁵⁶ is also obtained on calculating the real component of the conductivity at a frequency $\omega > 0$: in the two-dimensional problem a correction arises to the result of the kinetic equation that is proportional to $-\ln \omega \tau_r$, where τ_r is the relaxation time for momentum. It is assumed that $\omega \tau_r \ll 1$.

dielectric-semiconductor structure (see Fig. 1). Under these conditions the conductivity of the specimen is known to be finite (though perhaps small), since thermal excitation ensures the possibility of displacement of the electrons throughout the specimen. For an orientation in the possibilities that arise here, let us integrate Eq. (2.7), using here the formula (3.2) for β . When d=2 we obtain the following, taking into account the definition (2.2):

$$\sigma = \sigma_0 - \frac{e^2}{2\hbar} c_d \ln \frac{L}{L_0} , \qquad (3.5)$$

Here σ_0 and L_0 are constants. Of course, for weak scattering this expression is a tautology, while one must take respectively for σ_0 and L_0 the result of the kinetic equation and the mean free path with respect to momentum. However, another point is important. When $T \neq 0$, the lower limit of the integral in (I.11) [i.e., the numerator in the argument of the logarithm on the right-hand side of (3.5)] can be determined not only by the dimensions of the specimen, but also by inelastic scattering of electrons by phonons—as soon as the corresponding mean free path l_i is less than L.^{50,51} Under these conditions we must replace L in the integral in (I.11) with $\sqrt{l_1 l_2/2}$ (Ref. 203).¹⁵⁾

Further, we have the following expression for the relaxation time τ_i for inelastic scattering:⁶¹

 $\tau_1 \sim T^{-q}, \quad q > 0.$

Thus, when $l_i < L$, the second term in Eq. (3.1) is replaced by $cq \ln T$, where c is a positive constant. At the same time, the quantity $\sigma_{\rm g}$ practically does not depend on the temperature, being governed mainly by elastic scattering of charge carriers by impurities.

Such a logarithmic variation of the static conductivity has actually been observed experimentally.^{62,63} However, as we shall see in Sec. II, this fact alone does not yet imply an experimental confirmation of the correctness of the concepts that we have presented above.

We note that,⁵⁰ as soon as we are dealing only with processes of emission of phonons at a low lattice temperature, we can also take the temperature to be the electron temperature T_e , which is defined as a measure of the mean energy of the charge carriers. Thus the problem arises of the dependence of σ on the voltage applied to the specimen (or on the strength of the measuring current).

Another possible reason has been pointed out⁶⁴ for nonohmicity of specimens having localized charge carriers. It involves the "current" broadening of the levels. In fact, when $T \neq 0$ the diffusion coefficient D of the electrons also becomes non-zero. Consequently a localized state of an electron becomes nonstationary. The corresponding energy uncertainty ΔE is connected with the time τ_0 in which an electron diffuses over a distance of the order of l_1 . Evidently we have $\tau_0 \approx l_1^2/D$, and hence, $\Delta E \sim \hbar D/l_1^2$. On the other hand, the energy acquired by an electron in moving over the distance l_1 in a homogeneous electric field of intensity \mathscr{C} is \mathscr{eSl}_1 . This quantity has meaning if it is larger than ΔE , i.e., if

$$l_1 < \left(\frac{\hbar D}{\epsilon g}\right)^{1/3} \equiv l_g. \tag{3.6}$$

In the converse case in which $l_1 > l_s$, the maximum distance that an electron can traverse without changing energy must be identified with l_s . This is the situation at sufficiently low temperatures. If, moreover, we have

$$L_g < L, \tag{3.7}$$

then we must take as the lower limit of the integral in (I.11) the quantity I_s^{-1} . The inequality (3.7) implies that the voltage on the specimen $\mathscr{C}L$ exceeds $\hbar D/eL^2$. For macroscopic specimens this condition is satisfied apparently under all realistic experimental conditions. Thus, if we assume for an estimate that $D = 10 \text{ cm}^2/\text{s}$ and L = 0.1 cm, we find that the voltage $\mathscr{C}L$ must exceed $0.6 \times 10^{-12} \text{ V}$.

Evidently Eq. (3.1) must be rewritten under the conditions being discussed in the form¹⁶⁾

$$\sigma = \sigma_{\rm B} + \delta\sigma, \qquad (3.8)$$

Here we have

$$\delta\sigma = -\frac{e^2}{2\pi^3\hbar} \ln \frac{l_{\mathfrak{C}}}{l_p} = \operatorname{const} \cdot \ln \mathfrak{E}$$
(3.9)

(const>0). The formulas (3.8) and (3.9) remain valid as long as $l_{g} < l_i(T)$ and $l_{g} < L$. It is not yet fully clear to what extent this effect is appreciable against the back-ground of the heating of the electron gas by the electric field.

4. Qualitative study of Eq. (2.7). Localization of states for d = 1,2

According to hypothesis (a) of Sec. 3, the singular points of Eq. (2.7) can be only the zeros of the function $\beta(g)$, while hypothesis (b) implies that there is no more than one of these zeros. Formulas (3.2) and (3.4) give:

a) when $g \gg 1$: $\beta > 0, \ d = 3; \ \beta < 0, \ d = 1, 2.$ (4.1)

b) When $g \ll 1$:

$$b < 0, \quad d = 1, 2, 3.$$
 (4.2)

We note that when d = 1 and d = 3, the inequalities (4.1) arise from purely "geometric" considerations, being determined only by the dimensionality of the system. Here the small corrections to the results of the kinetic equation play no role. On the other hand, when d = 2 it is important in principle to take these corrections into account. In their absence we would obtain $\beta = 0$.

The inequalities (4.1) and (4.2) enable one to construct a qualitative picture of the behavior of the function $\beta(g)$ for different values of g (Fig. 3). We see that,

¹⁵In the presence of a magnetic field, and also when the interaction between the electrons is taken into account, other characteristic lengths arise (e.g., the well known magnetic length $\sqrt{\hbar c/Be}$, where B is the magnetic induction). A certain combination of them can also appear instead of L.

¹⁶Of course, the ideas presented above on the role of the electric field are justified only when l_i $(T_e) > l_p$ and $l_l > l_p$.



FIG. 3. Schematic form of the function $\beta(g)$ in systems of different dimensionalities.

when d = 3, the singular point $g = g_c$ exists, as is known [although Eqs. (3.3) and (3.4) do not enable one to find g_c exactly], while it does not exist for d = 1 or 2. One can understand the physical meaning of the singular point of Eq. (2.7) by employing the concepts of the modern theory of critical phenomena.⁴⁴ In this theory the unstable singular points of equations like (2.7) are critical points. We can easily convince ourselves that in our case something similar is taking place. In fact, let us study the curve for d = 3; let $\beta(g) = 0$ when $g = g_c$, and let the value of g deviate slightly from g_c . Then, according to (2.7), we have

$$\operatorname{sign} \frac{d \ln g}{d \ln L} = \operatorname{sign} \left(g - g_{c} \right). \tag{4.3}$$

This implies that the singular point being studied is unstable: when $g < g_c$ the values of g(L) will decline with increasing L. That is, in the limit of an infinitely large specimen we obtain $g(L) \rightarrow 0$. This is the region of localized states. On the other hand, when $g > g_c$ the function g(L) will increase with increasing L. In the limit of an infinite specimen, we obtain the asymptotic form (2.2) for the conductivity σ , which does not depend on L. This is the region of delocalized states.

Thus the singular point of Eq. (2.7)—the zero of the function $\beta(g)$ —defines nothing other than the mobility threshold. As we see from Fig. 3, Eq. (2.7) combined with the hypotheses adopted above yields the following conclusions:

a) When d = 3 there is a mobility threshold, i.e., the system being studied possesses both localized and delocalized states.

b) When d = 1 and d = 2, a mobility threshold is absent everywhere, i.e., all states are localized.¹⁷⁾ We should note that, when d = 2, this conclusion [cf. the remark following Eq. (4.2)] is essentially based on the existence and sign of the correction term on the right-hand side of (3.2).

However, the sign of this term depends on the nature of the interaction forces between the electrons and the scattering centers. Only potential scattering of charge carriers was treated in Refs. 59 and 60. Taking into account scattering by magnetic impurities⁵⁵ can lead to disappearance of a dependence of the conductivity on the dimensions of the specimen. Actually, another characteristic mean free path, $l_{\rm m} = v_{\rm F}\sqrt{\tau \tau_{\rm m}}$ is manifested here

(see Appendix I). Here $v_{\rm F}$ is the Fermi velocity, and τ and $\tau_{\rm m}$ are respectively the decay times of one-electron states in scattering by nonmagnetic and magnetic impurities (with short-range forces in both cases). If $l_{\rm m}$ < L, it is precisely the quantity $l_{\rm m}^{-1}$ that determines the lower limit of the integral in (1.11). Here the function $\beta(g)$ for d=2 generally ceases to vary upon reaching zero. Since there are no reasons to assume that g_c $\gg 1$, generally the case $g \gg 1$ is not realized here, and the entire approach based on treating an electron gas with weak scattering proves to be incorrect.

4

A zero of $\beta(g)$ for $g \gg 1$ does not arise, even in an external magnetic field, if scattering by magnetic impurities takes place.⁵² In this situation, localization of all the electron states in a two-dimensional system arises apparently only in a sufficiently strong random field. This implies that one should observe negative magnetoresistance in the system being studied: by causing a delocalization of the charge carriers, however partial, a magnetic field thus increases the electrical conductivity of the specimen.

Thus we see that the answer to the question of whether all the states in a two-dimensional system are localized depends on the concrete nature of the forces acting on the electrons and holes, i.e., not only on the geometry, but also on the dynamics.

On the other hand, the conclusion that, when d=1, all the states prove to be current-free in an arbitrarily weak (but not zero) random field does not involve the concrete nature of the latter. In this sense the scaling theory of localization supplements the prior results only to a certain extent.

The physical reason for the difference between the three-dimensional system, on the one hand, and the two- and one-dimensional systems, on the other, apparently involves the character of the random fluctuations in these systems.^{52,66} According to a well-known theorem of Polya,^{67,68} when d=1 and d=2, these fluctuations cause a particle to return again and again arbitrarily close to the initial point. Correspondingly, the scattering at each center unavoidably proves to be multiple, while the interaction of the electron with each individual scattering center proves strong, even at low amplitude of scattering per unit event. On the other hand, in a three-dimensional system such multiple returns can be insubstantial.

5. The problem of the minimal metallic conductivity (d = 3)

An attempt was also made in Refs. 50 and 51 to study the problem of the minimal metallic conductivity on the basis of Eq. (2.7). For this purpose, one must solve this equation in the region of values of g close to g_c , and then associate the scaling parameter L with the quantity

$$\eta = E - E_{\rm c} \, (>0). \tag{5.1}$$

Thereupon Eq. (2.2) will determine $\sigma(\eta)$.

It is convenient to rewrite Eq. (2.7) in the form

¹⁷⁾The same result was obtained also in Refs. 101 and 102.

$$\frac{\mathrm{d}g}{\mathrm{d}L} = \frac{g}{L} \beta(g). \tag{5.2}$$

In line with the hypotheses adopted in Sec. 3, the authors of Refs. 50 and 51 assume that

$$g\beta(g) = \frac{g - g_c}{v} \,. \tag{5.3}$$

Here, as we see from Fig. 3, we have $\nu > 0$ for either sign of the difference $g-g_c$. Henceforth in this section we shall be interested in the case $g > g_c$. Near the point g_c , one can replace the factor g on the left-hand side of (5.3) by g_c . Thus Eq. (5.2) acquires the form

$$L \frac{\mathrm{d}g}{\mathrm{d}L} = \frac{g - g_{\mathrm{c}}}{\mathrm{v}g_{\mathrm{c}}} \,. \tag{5.2'}$$

Let us integrate this equation from some point g_0 = $g(L_0)$ arbitrarily close to g_c (corresponding to the energy E) to an arbitrary point g(L) [within the limits of applicability of the approximation (5.3)]. We obtain

$$\ln \frac{g - g_c}{g_0 - g_c} = v^{-1} \ln \frac{L}{L_0}$$

That is, we have

$$L = L_0 \left[\frac{g(L) - g_c}{g_c - g_c} \right]^{\circ}.$$
(5.4)

In order to obtain from this an idea of the form of the function $\sigma(\eta)$, we must make some assumptions on the dependence of g on the energy E of the electron. In Refs. 50 and 51 the very simple linear relation

$$g - g_c = \left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{g=g_c} (E - E_c), \qquad (5.5)$$

was used. Then, according to (5.4), we have

$$L \sim \eta^{-\gamma}. \tag{5.6}$$

Then Eq. (2.2) yields (for $\eta \ll 1$):

$$\sigma = \frac{e^3}{2\hbar} g_c \left[L(\eta) \right]^{2-d} \sim \eta^{-(2-d)\nu}.$$
(5.7)

Thus, when d = 3, we obtain $\sigma \sim \eta^{\nu} \to 0$ when $E \to E_c$. The concept of a nonzero minimal metallic conductivity is not confirmed. Here, depending on the magnitude of the critical index ν , one of the two cases shown in Figs. 2b, c is realized.

However, this conclusion seems not completely convincing for two reasons. 58,69

First,⁶⁹ even if we adopt the general approach adopted in Refs. 50 and 51, it remains unclear whether the relationships (5.3) and (5.5) are always valid. Actually, if the assumption of the existence of a minimal metallic conductivity is valid, then when $\eta = 0$, the function $\sigma(\eta)$ must have a first-order point of discontinuity.

If it seems natural in the spirit of the scaling theory to assume the function $L(\eta)$ to be continuous, then also the dimensionless reciprocal resistance $g(\eta)$ [or g(L)] must undergo a discontinuity; in essence this is what makes it risky to integrate Eq. (5.2') directly from the point $g = g_c$. Hence, when $g = g_c$, the function $\beta(g)$ must have a singularity, and we must write Eq. (2.7) separately for $g > g_c$ and $g < g_c$. Here we must understand the left-hand side of (2.7) as being the right- or left-hand derivative.

Upon introducing, as before, a point g_0 close to g_e ,

we obtain the following from (2.7) in the region $g > g_c$:

$$L = L_{g} \exp\left[\frac{1}{g_{c}} \int_{g_{0}}^{g} \beta^{-1}(g) dg\right].$$
 (5.8)

However, in contrast to (5.3) and (5.5), now we assume that¹⁸⁾

$$g = g_c + a\eta^m, \quad \beta = \frac{(g - g_c)^n}{vg_c} \,. \tag{5.9}$$

Here we have a > 0, m > 0, $1 \ge n > 0$.

Then we find (when
$$n < 1$$
)

$$L \sim \exp\left[-\frac{v}{1-n} (a\eta^m)^{1-n}\right].$$
 (5.10)

Hence, when d = 3, we have

$$\sigma \sim \exp\left[\frac{\nu}{1-n} (a\eta^m)^{1-n}\right]. \tag{5.11}$$

As $\eta \neq 0$, this expression remains nonzero. Thus we see that the assumption of some particular analytic structure of the function $\beta(g)$ as $g \neq g_c$ is essentially equivalent to the *a priori* assumption of the existence or absence of a finite minimal metallic conductivity. In other words, solution of the problem posed here requires independent information of dynamic character.

Second, the statement was made in Ref. 58 that no characteristic physical length is involved in the quantity $g_{\rm c}$ —the zero of the function β . For this reason, the integration of Eq. (5.2) with account taken of (5.3) yields the following when $g - g_{\rm c} \ll g_{\rm c}$:

$$1 - \frac{g_0}{g} = \left(\frac{L}{L_0}\right)^{1/\nu}.$$
 (5.12)

Here L_0 is nothing other than an integration constant.¹⁹⁾ Hence, according to Ref. 58, one can draw no conclusions on the value of $\sigma(E_c)$. Actually the equality $\sigma(E_c)$ = 0 for d = 3 would mean that g remains finite as $L \rightarrow \infty$. However, it is clearly impossible to take the limit as $L \rightarrow \infty$ in Eq. (5.12). In the opinion of the authors of Ref. 58 (with which the author of this review concurs), the latter can involve the inadequacy of the approximate relationship (5.3) as $L \rightarrow \infty$.

The statement that one cannot find from Eq. (5.2) the value of $\sigma(E_c)$ without defining the constant of integration, i.e., without supplementing (5.2) with some boundary condition, can hardly arouse doubts. However, perhaps the solution of the problem of whether this value is zero or not requires somewhat lesser information. At present the problem seems open.

Another formulation of the problem has been presented in the studies of N. F. $Mott^{70,71}$ and M. Pepper.²⁰⁴ The idea consists of taking into account the large-scale fluctuations of the potential energy of an electron in the random field. Here it proves possible to associate the problem of the minimal metallic conductivity with calculating the localization length characterizing the wave functions of the electrons in the discrete spectrum.

¹⁸) In Ref. 69 the parameter ν was defined in a different way: in the notation adopted here it corresponds to $1 / \nu g_{c}$.

¹⁹In Ref. 58, ν appears instead of $1/\nu$ in the exponent on the right-hand side of the equation corresponding to (5.12). Apparently this is a misprint. However, it does not affect the subsequent discussions.



FIG. 4. Smooth band bending in a semiconductor with a random field. The marks denote discrete levels.

One can describe these fluctuations as a smooth random bending of the bands (Fig. 4; one can find a more detailed treatment in the book of Ref. 23, Sec. II.12). The mean value of this bending can be small or large, depending on the concrete nature of the field, but in principle it always exists. The existence of this bending makes it necessary to distinguish the local boundaries of the conduction and valence bands, E'_{v} and E'_{c} , from the boundaries of the bands taken as the mobility thresholds E_{v} and E_{c} . In a macroscopically homogeneous specimen, the latter do not depend on the coordinates, while the former do. Upon denoting by $\delta U(\mathbf{x})$ the smooth component of the random field, we can write

$$E_{c}'(\mathbf{x}) = E_{c} + \delta U(\mathbf{x}). \tag{5.13}$$

One can understand the role of the smooth bending of the bands in the problem in which we are interested by using the following qualitative arguments, mainly due to Mott.

Initially let $\delta U(\mathbf{x}) = 0$. We should expect that here the wave function of a conduction electron ψ will vary relatively slowly in space as the energy approaches the mobility threshold. The exact meaning of this expression consists of the following: Let us represent ψ as an expansion in the atomic wave functions φ_n , as is done in the well-known method of strongly bound electrons:²⁴

$$\psi = \sum_{n} a_n \varphi_n. \tag{5.14}$$

With a random distribution of atoms in space, the phases of the coefficients a_n can be random. However, in the sense of the delocalization concept, the amplitudes as $E \rightarrow E_c + 0$ must not fluctuate strongly upon going from one volume containing many atoms to another such volume. In other words, there are no factors here that could hinder the free translation of electrons of the stated energy throughout the specimen.

However, when $\delta U \neq 0$ the situation can change. Indeed, let us denote by E the energy of any electron, localized or delocalized. With a large enough value of $|\delta U(\mathbf{x})|$ the differences $E - E_c$ and $E - E'_c$ can prove to differ in sign. In this sense one can say that the smooth bending of the bands mixes the "localized" and "delocalized" states.²⁰⁾ Here, evidently, smooth large-scale Now let us examine the wave functions that would belong, for $\delta U = 0$, to the discrete spectrum as $E \rightarrow E_c - 0$. Just as in the case of the continuous spectrum, we can represent them in the form (5.14). However now we have

$$a_n \sim \exp\left(-\alpha r\right),$$
 (5.15)

Here α is the reciprocal localization length, and the coordinate r is measured from the localization center. (Henceforth we shall be interested in values of r considerably exceeding the interatomic distances; hence the exact position of the localization center is not essential.) The functions under consideration are of interest for our problem because, in the presence of smooth bending of the bands, the levels of the discrete spectrum must be measured from the local boundaries of the bands E'_{c} and E'_{r} . Then the radius of localization of an electron with energy $E < E'_{c}$ is determined by the difference $E'_{c}(x) - E$. Near the mobility threshold we can assume that α obeys

$$\alpha \sim (E_{\rm c}' - E)^{\rm s}.\tag{5.16}$$

Here s is a non-negative number. In contrast to the deep levels, s depends here on the parameters characterizing the random field in the specimen; in general we have $s \neq 1/2$.

Now if we assume that $E - E_c$, we obtain from (5.16) and (5.13)

$$\alpha \sim [\delta U(\mathbf{x})]^3. \tag{5.17}$$

We are interested in the case $\delta U > 0$. Indeed, then for $E > E'_c - 0$, the difference $E - E_c = E - E'_c + \delta U$ proves to be positive. That is, in the given region of space the level being studied lies above the mobility threshold. We note that the exponential (non-oscillating) form of the wave function does not mean now that the state with energy E is actually localized: the condition $\delta U(\mathbf{x}) > 0$ is obeyed only within the limits of certain restricted volumes. However, this is just why great and large-scale fluctuations of $|\psi|^2$ can arise in going to other regions of space. Now the problem consists of understanding when such fluctuations actually arise. For this purpose we note that, in the absence of non-point structural defects (of the type of radiation defects), the large-scale fluctuation of the potential energy of an electron covering a volume having a large number of atoms directly arises from the fluctuations of the number of atoms in this volume. Let us denote the characteristic linear dimension of the latter as R, and the mean distance between atoms as \overline{r} . Then the mean number of atoms in the given volume is $N \sim (R/\bar{r})^3$. We are interested in the case $N \gg 1$. Here we have the following expression for the rms fluctuation of the number of particles in the given volume δN :

²⁰⁾To avoid misunderstanding, we stress that we are here dealing with spatial delocalization of electrons and holes within the limits of a region in which the functions $E'_{c}(\mathbf{x})$ and $E'_{\gamma}(\mathbf{x})$ remain approximately constant. Of course, the classification of states into localized and delocalized in the sense of the condition $\sigma = 0$ ($\sigma \neq 0$) at $T \rightarrow \infty$ is determined only by the differences $E_{c} - E$ and $E - E_{\gamma}$.

$$\frac{\delta N}{N} \sim N^{-1/2} \sim \left(\frac{\bar{r}}{R}\right)^{3/2} \ll 1.$$
 (5.18)

We can naturally expect that, under the conditions of (5.18), the rms value of δU will satisfy the relationship

$$\delta U \sim \frac{\delta N}{N}$$

That is, according to (5.17) and (5.18) we have $\alpha R \sim (R/\tilde{r})^{1-(3s/2)} \sim (\delta U)^{s-(2/3)}$. Hence, when $E \sim E_c - 0$, we have

$$\alpha R \sim |E - E_c|^{s-(2/3)}. \tag{5.19}$$

If we wish to estimate how strongly $|\psi|^2$ fluctuates, we must substitute this expression for α_r into Eq. (5.15). Thus we see that, when $s \ge 2/3$, the magnitude of $|\psi|^2$ does not fluctuate too strongly, and does not succeed in decaying within the bounds of the volume that we are studying. In line with what we have said above, we can expect here a nonzero value of the minimal metallic conductivity.

On the other hand, when s < 2/3, the magnitude of αR increases without limit as $E - E_c \neq 0$. Hence $|\psi|^2$ varies rapidly in space, and decays without reaching the bounds of the volume being studied. This is the second of the cases pointed out above; we should expect that $\sigma_m = 0$ here.

Thus, in this approach the problem of the presence or absence of minimal metallic conductivity is reduced to the problem of the relationship between the actual exponent s and its critical value $s_c = 2/3$. According to K. F. Freed⁷² and S. F. Edwards,⁷³ the value of s is 2/3. Other values of s have also been indicated. However, the essential point is that all these results pertain only to a special type of random field for which the binary correlation function of (2.4) has a δ -function form:

$$\Psi(\mathbf{r}) = \Phi_0 \delta(\mathbf{r}).$$

With other types of random field, the situation apparently proves to be more complicated, since the corresponding correlation functions are no longer characterized by a single parameter Φ_0 , but by at least two. The concrete values of the index s here are yet to be obtained. Yet it appears clear that the results will depend on the characteristics of the random field being studied. In other words, we arrive at the previous conclusion: the eventual conclusion concerning a zero or nonzero value of σ_m is not of universal character—it involves the dynamic features of the system.

6. Power-law localization in a two-dimensional system

As we have seen in subsections 4 and 5, complete localization of all states of electrons and holes can occur in a two-dimensional disordered system. Apparently the same is true of other elementary excitations phonons, magnons, etc. However, a set of experimental data (see the review of Ref. 74) compels us to think that in a two-dimensional system, as in a three-dimensional one, there is a certain critical energy E_c that separates states of different types. Electrons occupying states of different types make substantially different contributions to charge transport. Namely, hopping conduction of the usual type occurs at energies $E < E_c$, and depends exponentially on the temperature, whereas the temperature-dependence of the static conductivity due to electrons of energies $E > E_c$ must be considerably weaker.

What we have said implies that, when $E_v < E < E_c$, discrete energy levels of the usual type most likely have the following distributions at a large distance r from the given center of localization, the wave function ψ of the electron has the following asymptotic form, apart from possible (but inessential) power-law factors:

$$\psi \sim \exp\left(-\alpha r\right). \tag{6.1}$$

Here we have $\alpha \sim (E_c - E)^s$ or $\alpha \sim (E - E_s)^s$ (the values of the non-negative number s can vary). This is termed exponential localization. Several years ago it was assumed that, when $E > E_c$ (or when $E < E_s$; hereinafter this possibility will always be understood), the electron states in the system being studied are delocalized, just as when d = 3. Now one must acknowledge this viewpoint to be nonuniversal. The problem naturally arises: what will happen in a two-dimensional system when $E > E_c$? The idea was advanced in Ref. 75 that power-law localization of charge-carriers occurs in this region, rather than exponential: at large distances from the center of localization, the wave functions of the electrons fall off asymptotically according to a power law²¹ (this is sometimes termed weak localization).

In principle the possibility of power-law localization of the solutions of the Schrödinger equation has been known for a very long time;⁷⁶ it has also been treated in the current monograph literature.^{77,78} In connection with the physics of disordered materials, this possibility has been discussed in the articles of Refs. 79, 80 and in the book of Ref. 23 (for d = 3). However, all these studies treat the case of a single potential well responsible for the appearance of wave functions of the type being studied. Evidently, the problem in which we are interested requires generalization of these results to the case of a finite concentration n, of wells, and also to the case of taking into account the total random field (we shall understand the latter term to denote the component of the random potential energy of an electron that does not reduce to a set of wells responsible for the power-law localization).81

Still it is convenient to start with studying the twodimensional problem of an individual well (at low enough values of n_t , this might have a real meaning). Let us set the origin of coordinates (for d = 2) at the center of the given well and denote by r and φ the polar coordinates in the plane, while $V(r, \varphi)$ is the potential energy of the electron in the given well. In line with what we have said above, we shall seek solutions of the Schrödinger equation for the potential energy V that have the following form at sufficiently large distances from the center:

²¹⁾The hypothesis of the possibility of power-function localization has also been advanced in Refs. 103 and 104 on the basis of numerical calculations.

$$\psi = Nr^{-n}f(r, \varphi). \tag{6.2}$$

Here N is a normalizing factor, and $f(r, \varphi)$ is a dimensionless function that is periodic in φ with the period 2π . As $r \rightarrow \infty$, it either has a finite bound, $f_{\infty}(\varphi) \neq 0$, or it oscillates (with an amplitude independent of r). The character of the approach of f to the limit can vary; we note only that, since the point " $r \rightarrow \infty$ " is not singled out physically in any way, the derivatives $\partial f/\partial r$ and $\partial^2 f/\partial r^2$ must be bounded as $r \rightarrow \infty$. It will also be evident from the material below that the condition that the states being studied be current-free reduces to the condition that f should be real.

Let us denote by a the distance from the center at which the function fits the asymptotic form (6.2) to an acceptable accuracy.

For simplicity we shall assume that this is the sole characteristic length in the problem (with a single well). Then considerations of dimensionality imply that we have $N \sim a^{\pi^{-1}}$ when n > 1. When n = 1 (this case has been treated in Ref. 75), the function (6.1) is normalizable only in a restricted volume, and we have $N \sim -\ln(an_t^{1/2})$.

Evidently, the treatment of solutions of the type of (6.2) can have a meaning only when

$$a \ll n_t^{-1/2}$$
. (6.3)

This inequality defines the exact meaning of the words "sufficiently low concentration of wells".

Upon substituting the function (6.2) into the corresponding Schrödinger equation, we obtain

$$-\frac{\partial^2 f}{\partial r^2} + \frac{2n-1}{r} \frac{\partial f}{\partial r} - \frac{1}{r^3} \frac{\partial^2 f}{\partial \varphi^3} - \frac{n^3}{r^2} f = \frac{2m}{\hbar^3} (E-V) f.$$
(6.4)

When $r \gg a$, the function $V(r, \varphi)$ must approach some constant value V_0 . In the contrary case the concept of a single potential well would make no sense in the system being studied.

What we have said earlier on the asymptotic behavior of the derivatives of f implies that a nontrivial solution of the type of (6.2) exists only for the single eigenvalue of the energy $E = V_0$. Finally, by choosing the reference origin of the energy just as in the case of (6.1), we see that $V_0 = E_c$, a discrete level belonging to the solution (6.2), coincides with E_c . This is a well known result.^{76-78, 80, 23} Now we should treat Eq. (6.4) as the definition of the potential energy $V(r, \varphi)$ for an a priori imposed function $f(r, \varphi)$. Evidently a function $V(r, \varphi)$ defined in this way actually falls off toward infinity. Owing to the great arbitrariness of choice of fand n, there can be quite a number of such wells. However, we should note that such a situation is unstable: even small perturbations of the potential energy can lead to a shift in the energy and correspondingly to disappearance of solutions of the type of (6.2). These perturbations can be associated with either the presence of other potential wells of the same type as in (6.4), or with a total random field in the disordered material. First let us examine the role of the first of the stated factors. The main role here is played by the known effect of "repulsion" of levels; under the conditions of (6.2) it evidently should be even more noticeable than in the case of (6.1).

Conveniently we can start with treating two wells lying at the points r = 0 and r = R ($r = \{r, \varphi\}$). This approach is justified if $R \gg a$, but on the other hand, the rest of the wells are even more remote from the two that are being explicitly studied. This is possible under the conditions of (6.3) as long as $R < n_t^{-1/2}$. The essence of the matter can be understood most simply on the example of identical wells, in which V(r) = V(r + R). Let ψ_1 and ψ_2 be the wave functions of the single-well problem, respectively lying at the points r = 0 and r = R. By definition, these functions are identical in form, but centered at different points. As always in such problems, we can seek the wave function of an electron in the system of two wells in the form

$$\psi = c_1 \psi_1 + c_2 \psi_2. \tag{6.5}$$

Here c_1 and c_2 are constant coefficients. Upon substituting this expression into the Schrödinger equation with the potential energy $V(\mathbf{r}) + V(\mathbf{r} + \mathbf{R})$, we easily obtain a system of equations for the coefficients c_1 and c_2 . By solving it, we can find the eigenvalues of the energy of the electron moving in the field of the two wells.

Let us assume that

a k

$$A = \int \psi_1^2(\mathbf{r}) V(\mathbf{r} + \mathbf{R}) d\mathbf{r}, \quad S = \int \psi_1(\mathbf{r}) \psi_2(\mathbf{r} + \mathbf{R}) d\mathbf{r},$$

$$B = \int \psi_1(\mathbf{r}) V(\mathbf{r} + \mathbf{R}) \psi_2(\mathbf{r} + \mathbf{R}) d\mathbf{r}.$$
(6.6)

We note that it is convenient to choose an orthogonalized combination of the type of (6.5) as the basis system of functions when calculating the electrical conductivity. As the temperature T = 0, the static electrical conductivity σ is determined by isoenergetic transitions. Hence on the given basis only the diagonal elements of the current density figure in the formula for σ . For this reason, the condition adopted above that f should be real implies a current-free character of the electron states being studied.

We can easily derive in the standard way the two eigenvalues of the energy of the electron:

$$E_{\pm} = \frac{A \pm B}{A \pm S} \,. \tag{6.7}$$

Here we have

$$\left(\frac{c_{1}}{c_{1}}\right)_{\pm} = \mp 1, |c_{1}|_{\pm} = |c_{2}|_{\pm} = \frac{1}{(1 \mp S)\sqrt{2}}.$$
 (6.8)

(The coefficient $\sqrt{2}$ in the denominator is for normalization.)

Thus the former level $E = V_0$ is split into two separated by the amount

$$\Delta(R) = E_{+} - E_{-} = \frac{2(B - AS)}{1 - S^{2}} \cdot$$
(6.9)

(Here ΔR can have either sign.)

Upon employing the expression pointed out above for the normalizing factor N (for n > 1), we can easily derive the asymptotic estimates

$$|B| \sim R^{-n} |V(R)|, \quad |A| \sim |V(R)|, \quad S \sim \left(\frac{a}{R}\right)^n. \tag{6.10}$$

This implies that $|\Delta(R)|$ declines more rapidly with increasing R than $\psi_1(R)$, which essentially justifies this whole calculation.

Formally we have $\Delta(R) \neq 0$ for any R. However, as we have already noted in Sec. 4, "discrete" levels lying above the Fermi level always have a finite intrinsic width. Let the total intrinsic width of the levels being studied be δ . Just as in Sec. 4, we shall define the correlation length R_0 by the equation

$$|\Delta(R_0)| = \delta. \tag{6.11}$$

Then, when $R > R_o$, the levels under consideration must be considered isoenergetic. We see that, when

$$a \ll R_0 < n_t^{-1/2} \tag{6.12}$$

there is an interval of values of R in which

a) the splitting of levels is negligibly small,b) one can single out two asymptotic regions:

$$a < r \ll \frac{R}{2}, \quad \frac{R}{2} < r < R - a \tag{6.13}$$

and

c) the effect of the other wells is insignificant.

We can easily generalize these conclusions to the case of several wells. Thus, under the conditions (6.12) and with neglect of the effect of the total random field, the concept of power-law localization has a literal meaning, but it can hardly explain the character of the transport phenomena observed⁷⁴ when the Fermi level lies in the region $E > E_c$.

Actually, under the conditions being treated, the states with power-law localization occupy only a very narrow band of energies around $E_{\rm e}$.

However, the situation changes when we reject the conditions (6.12) and/or when we take into account the component of the random field U_1 that does not reduce to the summation

$$V = \sum V_i (\mathbf{r} + \mathbf{R}_i). \tag{6.14}$$

Here the subscript *i* numbers the centers of power-law localization situated at the points R_i . Indeed, let us examine first what happens if, while still neglecting the term U_1 in the potential energy of the electron, we turn to the case of a rather large concentration of wells of the form $V(\mathbf{r} + \mathbf{R}_i)$, in which $n_i^{-1/2} \leq a$ (and *a fortiori*, $n_t^{-1/2} \leq R_0$). Here we must treat the problem of an electron in the field of (6.14) and correspondingly seek a wave function in the form

$$\psi = \sum_{i} c_i \psi_i. \tag{6.15}$$

Here the summation is performed over all the wells, rather than two, of the type being studied. The function ψ_i has the same meaning as in the case of (6.4).

The system of equations for the coefficients c_i can prove no simpler than the initial Schrödinger equation. However, the result in which we are interested is clear without calculations: since none of the asymptotic regions of (6.13) exist under the conditions being considered now, the arguments that led us to the conclusion of a single eigenvalue $E = E_c$ become unsound. In particular, one must not neglect the effect of repulsion of levels. In combination with the usual considerations of the small probability of having wells similar in depth and form in close-lying regions of space, it must lead to spread of the levels in a finite (at least!) region of energy values.

These conclusions are even more strengthened if we bear in mind the term U_1 in the potential energy of the electron. In order to have the concept of an individual potential retain its meaning, we require here not only the inequality (6.3), but also the condition that the variation of the random field in space is sufficiently slow on the average: the characteristic length over which the function $U_1(r)$ varies appreciably must be large in comparison with R and $n_{\star}^{-1/2}$. In the converse case the asymptotic regions of (6.13) cannot be distinguished for any values of $an_{+}^{-1/2}$ and aR^{-1} . Apparently we should expect here the appearance of an ordinary²³ everywheredense energy spectrum, while we can retain the concept of power-law localization only in a certain conventional sense: we are speaking of states having wave functions that would have the asymptotic behavior of (6.2) if some region existed in which such behavior is possible. However, actually one obtains here something like "tissues" of wave functions occupying the entire specimen.

Yet the fundamental idea of Kaveh and Mott⁷⁵ remains in force: there is a region of energies of an electron that corresponds to current-free states having wave functions of nonexponential type. It is not yet fully clear what will be the temperature-dependence of the static electrical conductivity associated with these states. However, evidently we should not expect here the strong dependence characteristic of hopping conduction of the usual type.

Another treatment of the electron states with energies above E, has been proposed by I. P. Zvyagin.²⁰⁵ He proposed a simple interpolation formula for the function $\beta(g)$ that covers both limiting cases pointed out in Sec. 3. The solution of Eq. (2.7) with this function on the right-hand side led to the conclusion of exponential localization of all electron states for d = 2. However, there is a critical energy E_c , above which the localization length increases very rapidly with the energy. When this length becomes comparable with the dimensions of the specimen, the corresponding states become practically delocalized. Apparently, in the sense of answering the question posed at the beginning of this section, this idea is no worse than the concept of powerlaw localization. At the same time, the comments made above on the actual lack of asymptotic regions in a material with a not too low concentration of localization centers and/or a random field of general type also pertain to it.

II. MULTIELECTRON EFFECTS IN THE PHYSICS OF DISORDERED SYSTEMS

7. Introduction

The multielectron problem in the physics of disordered systems has two aspects. First, the question involves taking into account the interaction between delocalized charge carriers. It had already been noted^{82,83} rather long ago that, in the presence of a random field, the role of the interelectron interaction must increase in comparison with what happens in an ideal crystal. Actually, in the latter case the joint action of the laws of conservation of energy and quasimomentum appreciably restrict the probability of electron-electron scattering-so much so that it becomes arbitrarily small near the Fermi surface (and upon this, strictly speaking, the very concept of a Fermi surface is grounded). However, in a nonperiodic force field, the components of the quasimomentum can no longer be treated as "good" quantum numbers. That is, the law of conservation of quasimomentum does not hold, and hence, the restrictions cited above are lifted.

Correspondingly it becomes topical to study the effect of interelectron interaction on the kinetic and other characteristics of a system of delocalized charge carriers. This problem will be treated briefly in Sec. 8.

Second, the role of interaction between the localized charge carriers may prove to be very substantial. This situation is well known in the physics of crystalline, weakly doped semiconductors.²⁴ Coulomb repulsion can render the localization of two electrons near the same impurity atom energetically unfavorable, provided only that here we do not gain the energy of a chemical bond or the energy of deformation (in particular, polarization) of the atomic matrix. The latter reservation is essential: the effects that we have just pointed out can radically alter the situation, by allowing, e.g., localization of up to three extra electrons on gold and copper atoms in crystalline germanium. Moreover, the resultant potential energy of interaction of the electrons can prove to be negative, i.e., the electrons in a solid can undergo mutual attraction rather than repulsion²²⁾.⁸⁴⁻⁸⁷ Yet another effect has attracted serious attention in recent years that is due to interaction between electrons. It consists of a substantial change in the density of states near the Fermi level whenever the latter is situated in a region of energies occupied by discrete levels ("Coulomb gap"). We shall turn to the problems associated with this effect in Sec. 9.

8. Interaction between delocalized electrons or only impurity scattering

A quantitative study of the role played by interaction between delocalized electrons in a not fully ordered system was apparently first carried out with the necessary degree of completeness in Refs. 88-90 (for the case d=3). There the behavior was studied of a degenerate gas of electrons interacting with one another and undergoing elastic scattering at point structural defects (impurities). The interaction forces between the electrons and the impurities are assumed to be of shortrange.²³⁾ The scattering by an impurity was assumed to be weak in the sense of the usual condition

- 6

$$k_F l_p \gg 1. \tag{8.1}$$

Here I_p is the corresponding mean free path with respect to momentum, and k_F is the Fermi wave number. At the same time, the values of I_p and τ —the corresponding time of free flight—were considered still to be not too large, such that the electrons could undergo multiple scattering by the impurity within the characteristic time of interelectron interaction. This means that, in addition to (8.1), the following inequalities must be satisfied:

$$ql \ll 1, \quad q_0 \tau \ll 1. \tag{8.2}$$

Here $\hbar q$ and $\hbar q_0$ are the momentum and energy transferred in a single event of inelastic scattering of an electron by an electron.²⁴⁾

Under the conditions (8.2), interference of the two scattering mechanisms occurs. Consequently corrections arise for the density of states, the electronic heat capacity, the resistivity, and other electronic characteristics of the material. These corrections depend on the temperature (for $T \ll \hbar/\tau$) and, for the resistivity, on the frequency of the external field ω (for $\omega\tau \ll 1$).

The results of Al'tshuler and Aronov have been generalized⁹¹ to the case d = 2 (under the assumption that it makes sense to speak of delocalized electrons). It turned out that the temperature correction to the conductivity is given by the same formula (3.1) with a different numerical coefficient in front of the logarithm: the following expression appears instead of $e^2/2\pi^2\bar{n}$:

$$\frac{e^2}{2\pi^2\hbar} [1 - F(n)]. \tag{8.3}$$

Here n is the concentration of delocalized electrons, while we have

$$F = \frac{1}{2\pi} \int_{0}^{2\pi} d\phi \left[1 + 2k_F \kappa^{-1} \sin \frac{\phi}{2} \right]^{-1}.$$

Here h_F is the same as in Eq. (8.1), and κ^{-1} is the screening radius. Since $k_F \sim n^{1/2}$ for a degenerate gas, while $\kappa \sim n^{1/4}$, ¹²⁹ we find $F \rightarrow 0$ for a sufficiently high electron concentration. In the absence of scattering with spin reversal, even the coefficients prove to be the same.²⁵⁾ Thus the question arises: just what do the experimental data confirm?^{62,63} A detailed calculation of the numerical coefficient in Eq. (3.1) (and the analogous equation) unavoidably involves using some approximations. Hence it is more convenient to turn to studying other effects—magnetoresistance and the Hall effect.^{93,95} In the presence of a magnetic field with the

²²⁾This effect can also occur for free electrons. It is well known in the theory of superconductivity. However, in the case of localized electrons it can prove to be far stronger, being manifested even in the region of relatively high temperatures.

²³Apparently the last restriction plays no fundamental role.³⁰ It only simplifies the calculations by enabling one to identify the transport relaxation time (time of free flight) with the time for departure of the electrons from a state having a given momentum.

²⁴⁾Formally this means that the vertex component describing the interelectron interaction must be renormalized with the aim of taking impurity scattering into account. Here it becomes dependent on q and q_0 .

²⁵)One can find a simple derivation of this result in Ref. 92.

induction \mathscr{D} , another characteristic length appears in the problem, $L_{\mathscr{D}} = (\hbar c/\mathscr{D} e)^{1/2}$. Consequently the formula for $\delta\sigma$ is modified. The general expression for $\delta\sigma$ proves to be somewhat unwieldy. However, it is simplified if the length $L_{\mathscr{D}}$ is comparable with the length Lof the specimen or smaller.⁹⁵ Then (with $l_t > L$), one obtains the previous formula (3.1) with a replacement of the length of the specimen by

$$L_D = (L^{-2} + L_{-2})^{-1/2}. \tag{8.3'}$$

Evidently this implies the appearance of a negative magnetoresistance in any magnetic field, even a weak one.²⁶⁾ An essential point is that, in a theory based on the concept of localization of electrons, the role of the magnetic field consists of its effect on the orbital motion of the charge carriers.⁹⁶ This implies that we should understand \mathscr{D} in the expression for $L_{\mathscr{D}}$ to be the component of the magnetic induction normal to the plane of the specimen.

The formula of the type of (3.1) also becomes modified in a theory based on taking into account the interaction between the electrons: the factor 1 - F is replaced by 1 - F/2 (according to Ref. 97 this result is due to Lee and Ramakrishnan). However, here the effect involves the influence of the magnetic field on the spin of the electron, and is appreciable only when $g\beta_{\mathscr{B}} \gg T$ (g is the gyromagnetic ratio, and β is the Bohr magneton). Moreover, this effect does not depend on the orientation of the vector $\overline{\mathscr{B}}$. Thus, one can find out by studying the magnetoresistance of a specimen which of the two mechanisms being discussed plays the major role.

The same possibility is offered by studying the Hall effect. According to Ref. 98, in a theory based on a picture of localized electrons, a correction to the standard formula of the Hall constant R does not arise, while the Hall mobility declines logarithmically with decreasing temperature. On the other hand, the interaction between the electrons leads⁹¹ to appearance of the correction

$$\frac{\delta R}{R} = 2 \frac{\delta r}{r}$$

Here r is the resistance of the specimen, and the Hall mobility increases logarithmically with decreasing temperature.

In both cases the role of the temperature can also be played by the electron temperature. This enables one to employ a study of the dependence of the kinetic coefficients on both the magnetic and electric fields^{$\theta\theta$} to discriminate between the two mechanisms being studied.

Experimental data^{97, 99} obtained in experiments with inversion layers at a temperature $T \approx 85$ mK and with a magnetic induction in the range from zero to ≈ 1 tesla

FIG. 5. Exciton instability in a crystal. If the binding energy of an electron and a hole $W_B > E_{g}$, then spontaneous formation of electron-hole pairs bound into excitons becomes energetically favorable.

apparently indicate the existence of both mechanisms that give rise to a logarithmic variation of the type of (3.1). In relatively weak magnetic fields (up to $\mathscr{B} \approx 0.1$ tesla), the localization effect dominates. With further increase in the magnetic induction, the well-known delocalization of electrons takes place, and the interaction between them begins to play the main role.

9. The problem of the Coulomb gap

The problem of the influence of the interelectronic interaction on the energy spectrum, the thermodynamics, and the kinetics of localized charge carriers in disordered semiconductors began to attract attention at the beginning of the seventies.¹⁰⁵⁻¹⁰⁷ It has been studied subsequently in the aspect in which we are interested in Refs. 108-127. The essence of the matter consists of the fact that (see Ref. 23, Sec. II.16), in the presence of localized states of acceptor and donor types, the Coulomb attraction between the localized electrons and holes can lead to a phenomenon analogous to a certain extent to ordinary exciton instability (Figs. 5 and 6). Evidently, instability with respect to spontaneous creation of "bound electron-bound hole" pairs (see Fig. 6) must lead to rearrangement of the energy spectrum of the system: owing to the Coulomb interaction, the density of states p(E) changes in such a way that the aforementioned instability cannot arise. Namely, the value of $\rho(E)$ must decrease sharply near the energy E = F_0 —the Fermi level at absolute zero temperature.²⁷⁾



FIG. 6. Instability leading to formation of a Coulomb gap. As $T \rightarrow 0$, the levels of acceptor (donor) type lying above (below) the Fermi level are filled with holes (electrons); these states are electrically neutral. One of the types of excitation of the system consists in transfer of an electron from a donor (with $E = E_1 < F$) to an acceptor (with $E = E_2 > F$). If the energy of mutual attraction that arises here between the "extra" electron in the acceptor and the hole in the donor is greater than the difference of "single-electron" energies $E_2 - E_1$, then spontaneous formation of such pairs becomes energetically favorable.

²⁶)An additional effect can also appear in a strong magnetic field.¹⁰⁰ Indeed, let the magnetic length $L_{\mathfrak{B}}$ be considerably smaller than the characteristic length over which the random field varies appreciably on the average. Then a force exerted by the random field will practically not act on an electron magnetically localized in a region with linear dimensions of the order of $L_{\mathfrak{B}}$.

 $^{2^{\}overline{(1)}}$ Here and henceforth we shall use the term "density of states" in the sense of its exact multielectron definition (see Ref. 23 and Appendix II). Here the energy argument *E* need not coincide with any "single-electron" energy.

This is commonly referred to as the appearance of a Coulomb gap—a region with a sharply decreased density of states.¹⁰⁸

According to Refs. 108, 111, 112, 117, 120, 121, 113, 115, and 116, at T = 0 the density of states, calculated with account taken of the interaction between electrons, vanishes at $E = F_0$. However, at values of Earbitrarily close to F_0 , the function $\rho(E)$ already differs from zero ("soft gap"¹¹²). On the other hand, in Refs. 106, 109, and 119, the value of $\rho(E)$ was still found to be finite, though diminished in comparison with the value obtained upon neglecting the interaction between electrons. Apparently, the difference in results involves the differing treatment of the random nature of both the "single-electron" energies and the distance between the localization centers; complete agreement here has not yet been attained.

In a three-dimensional system (in the sense of Ref. 23), the density of states $\rho(E)$ as $E - E_0$ is approximated by the following expression¹¹² ¹¹³:²⁸⁾

$$\rho(E) = \frac{1}{2} \rho''(E) (E - F_0)^2 + \dots$$
(9.1)

Apparently this is confirmed also by numerical calculation.¹²⁶

However, we note that this result is most likely of a modeling nature. On the other hand, the conclusion that $\rho(F_0)$ is zero has been derived^{113,115} without any assumptions of a modeling nature. The only essential point was that in a condensed medium the electric field is unavoidably screened, e.g., by the spatial redistribution of the delocalized or localized charge carriers. We should note in this regard that one can approach the question of taking account of screening in two ways.

First, one can formally sum all the diagrams for which the ordinary loops play the role of skeletal loops, e.g., by calculating the potential of a point charge. This approach, which was adopted in Refs. 113 and 115 and in the book of Ref. 23, is apparently convenient for proving general theorems. However, it can prove ineffective in concrete calculations if one does not restrict the treatment to some particular form of perturbation theory.

Second, one can explicitly solve the ordinary selfconsistent Poisson equation, the right-hand side of which involves the bulk charge density, the latter depending on the sought potential. Probably this approach is useful for numerical calculations. However, being essentially approximate, it can lead to extra complications in discussions of a general nature. Which of the two cited approaches one should choose in any particular problem is a question of convenience. Naturally, the appearance of a Coulomb gap should affect the thermodynamic and kinetic characteristics of the system of electrons being discussed. First let us turn to the thermodynamics.

As is well known, one can find the thermodynamic characteristics of a material by specifying the number of particles as a function of the temperature, the magnetic induction, and other quantities of this type. On the other hand, the following general formula $holds^{23}$:

$$\frac{N}{V} = \int \rho(E) n_F(E) dE.$$
(9.2)

Here V is the volume of the system, while

$$n_F = \left[\exp\left(\frac{E-F}{T}\right) + 1 \right]^{-1}$$

- **h**

is the Fermi function. The smoothed density of states $\rho(E)$ is given by the formula

$$\rho(E) = \frac{2}{V} \langle \operatorname{Sp} \operatorname{Im} G_{\mathbf{r}}(E) \rangle \equiv \rho(E, T, F).$$
(9.3)

Here G_r denotes the Fourier transform of the retarded single-particle Green's function, while the angle brackets indicate the need to average over the random field.

The formulas (9.2) and (9.3) are exact. They hold for any system of interacting particles, and do not involve any assumptions whatever of modeling type nor classification of elementary excitations in the system, etc. However, using them in concrete problems can face serious difficulties, since the density of states defined by Eq. (9.3) generally depends itself on the temperature and on the Fermi level. The situation is eased only for low enough temperatures when-under certain conditions-a theorem holds that was proved in Ref. 128. For convenience we have presented its proof in Appendix II together with certain supplements. This theorem states that one can take as $\rho(E)$ the value of the density of states calculated for T = 0 with an accuracy no worse than quantities of the order of T^4 . Here the temperature-dependence of the number of particles (using the thermodynamic variables T and F) is determined only by the argument T, which explicitly enters into the Fermi function. The conditions for applicability of the theorem are pointed out in Appendix II. These conditions are fulfilled in the problem of the thermodynamics of a system with a Coulomb gap. This implies, in combination with Eq. (9.1), that the electronic heat capacity of the (three-dimensional) system under consideration (at constant volume) is a cubic function of the temperature^{23,115,116}:

$$c_{\rm v} \sim T^3. \tag{9.4}$$

We note that this conclusion is based solely on a single assumption of modeling type, which is Eq. (9.1).

On the other hand, a different relationship was derived in Ref. 125 for the heat capacity at low temperatures:

$$c_{\rm v} \sim T |\ln T|^{-1/2}$$
 (9.5)

Here the following formula was used:

$$\rho(E) \sim \left(\ln \frac{\text{const}}{E - F_0} \right)^{-1}. \tag{9.6}$$

The latter was derived by the authors by using their

²⁸Of course, the notation of the density of states in this form does not at all imply its analyticity as a function of $E - F_0$, even if we add to the right-hand side of (9.1) another constant¹¹⁹ or several of the next terms of the expansion.¹¹⁵ We assume only the possibility of such an approximation of $\rho(E)$ in the region $E \approx F_0$, but not at all the possibility of expanding the density of states in an absolutely convergent series in powers of $(E - F_0)$.

self-consistent equation for the density of states²⁹⁾ (our notation is not the same as the one used in Ref. 125). Apparently it would be premature to enter here into a detailed discussion of this equation (the authors themselves¹²⁵ gave a preliminary evaluation in the cited article). However, we should note that the derivation of Eq. (9.6) seems to be not entirely free of objections, even within the framework of this equation. One of these consists of the following. In Ref. 125 the probability $F(\mathbf{r}, \omega)$ of formation of a "pair" excitation is first introduced, i.e., the probability that there is a node occupied by an electron in the interval dr_1 about the point r_1 , and an empty node in the interval dr_2 about the point r_2 , with the energy ω needed to transfer the electron from the latter to the former node assumed to be fixed. Then the interaction of these pairs ("dipole excitations" in the terminology of Ref. 125) is examined. To do this one writes the self-consistent equation for the function $F(\mathbf{r}, \omega)$, where $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Here the latter is implicitly identified with the conditional probability of creation of a pair of excitations with the parameters \mathbf{r}' and ω' at the distance s from the coordinate origin when the pair indicated above is situated near the origin. However, as is known, this conditional probability differs from $F[\mathbf{r}, \omega]$ by the factor of the correlation function. Thus, the latter is considered constant in Ref. 125. Is the singularity of (9.6) [and hence of (9.5)] conserved when one rejects this assumption, and also takes into account the unavoidable screening of the interaction between the dipoles?

The experimental situation is as yet not fully clear. On the one hand, perhaps the cubic relationship (9.4) is difficult to distinguish against the background of the heat capacity of the atomic matrix. On the other hand, as yet (December 1981) no experimental indications have appeared of a singularity of the type of (9.5), insofar as the author of this review knows.

On proceeding to the role of the Coulomb gap in kinetics, we note first of all that it is essentially assumed in the standard theory of hopping conductivity with a variable hop length that the density of states at the Fermi level differs from zero. Here the temperaturedependence of the hopping conductivity σ_h is given by the expression

$$\sigma_{\rm h} \sim \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right]. \tag{9.7}$$

Here the "Mott temperature" varies as $T_o \sim \rho^{-1}$ (see Ref. 23, Sec. IV.10; at a low enough temperature we can here replace F with F_o). Under the conditions of (9.1), instead of (9.7), the standard percolation procedure leads^{107,111,126,131} to the relationship

$$\sigma_{\rm h} \sim \exp\left(-\left(\frac{T_0'}{T}\right)^{1/2}\right]. \tag{9.8}$$

Here we have $T'_o \neq T_o$.

680 Sov. Phys. Usp. 26(8), Aug. 1983

However, it has been noted^{110,119,123,127} that another factor can play a role in addition to the change in density of states: simultaneous correlated displacement of many electrons. One can distinguish two types of such hops.

First, when an electron hops between two centers of localization (atomic sites), a redistribution of the other electrons in space also occurs near the starting and end points. Usually this is referred to as the polaron effect, with the idea that "polaron" hops are occurring, rather than "pure electron" hops.³⁰ According to Mott,¹²² one again obtains here the formula (9.7) with the former value of T_0 , but with a diminished (and temperature-dependent) coefficient in front of the exponential.

Second, a cascade process^{123,127} can occur, which amounts to a set of simultaneously occurring "singlepolaron" hops. According to Pollak,¹²⁷ here the temperature-dependence of the hopping conductivity proves to follow a power law:

 $\sigma_{\rm h} \sim T^{\rm p} \tag{9.9}$

with an exponent p difficult to determine.

However, we note that both Mott and Pollak did not employ the density of states in the form (9.1), while apparently assuming that, when $E = F_0$, the density of states has only a dip, but is not zero. Further, as was noted in Ref. 123 itself, the result (9.9) was obtained by a simple maximization of the probability of hopping without taking percolation considerations into account. All this renders the theoretical situation not entirely clear at present. The experimental situation is also not very clear. Apparently the relationship (9.8) was discovered in Ref. 132, and the power-law relationship in Refs. 133 and 134. Nevertheless, for a final assurance that these effects involve precisely the existence of a Coulomb gap, one would wish to have experimental data on other characteristics of the specimen-the electronic heat capacity and the magnetic susceptibility, as well as the temperature- and frequency-dependences of the hopping conductivity in an alternating field, $\sigma_h(\omega)$. According to Ref. 116, one should expect here the following relationships (at low temperatures and high enough frequencies, at which one can restrict the treatment to phononless transitions):

$$\operatorname{Re} \sigma_{h} (\omega) \sim \sigma_{0} (\omega) \begin{pmatrix} T^{4}, & T \gg \hbar \omega, \\ \omega^{4}, & T \ll \hbar \omega. \end{cases}$$

$$(9.10)$$

We have here denoted by $\sigma_0(\omega)$ the real component of the hopping conductivity at the frequency ω that would have been obtained in the absence of the Coulomb gap. The explicit expressions for $\sigma_0(\omega)$ are well known.^{25,135}

²⁸⁾An analogous result was obtained ¹³⁶ by a numerical method. Unfortunately, owing to the unavoidable approximations adopted in this study, the degree of reliability of the extrapolation made there is not fully clear. Also the relationship of the density of states studied there to that appearing in Eq. (9.2) is not fully clear.

³⁰⁾Apparently, if we employ the term "polaron" here, we should add to it the word "electronic," since an ordinary polaron, as we know, amounts to an elementary excitation caused by the interaction of charge carriers with the lag polarization of the atomic matrix.

III. THE "TAIL" OF THE INTERBAND LIGHT-ABSORPTION COEFFICIENT IN A DISORDERED SEMICONDUCTOR

10. Introduction

One refers to a "tail" of the absorption coefficient ("optical tail") when photons are absorbed with energies smaller than the optical width of the forbidden band E_g . In the physics of weakly doped crystalline semiconductors one usually understands here that the concept of the "tail" does not include light-induced electron transitions between discrete levels and the conduction or hole bands. However, in a disordered semiconductor, including a heavily doped one, strictly speaking, the spectrum of discrete levels in the mobility gap becomes everywhere dense, and this reservation can lose sense.

The light-absorption coefficient α in the region of the tail is very often described by a formula that has become called the Urbach rule:

$$\alpha = \alpha_{\mathbf{m}} \exp\left(-\frac{E_g - \hbar \omega}{\overline{w}}\right). \tag{10.1}$$

Here \overline{W} is a parameter independent of the frequency, while the pre-exponential coefficient α_m depends on the frequency relatively weakly.

Apparently this relationship was first established by an empirical method in Ref. 137, in which light absorption was studied in alkali-halide crystals. Subsequently a relation of the form of (10.1) has been repeatedly observed in the most varied materials and under different experimental conditions.³¹⁾

According to Ref. 137, in ionic crystals the parameter W is approximately proportional to the temperature T. This conclusion has been confirmed also in a number of other studies; it has proved to be valid also as applied to crystals of a different type if the temperature is not too low. Thus, in strongly doped (and not too well compensated) p-type gallium arsenide,¹³⁹⁻¹⁴¹ a linear temperature-dependence of \overline{W} was observed down to ~100 K, becoming considerably weaker with a further decrease in temperature. In strongly doped and compensated semiconductors and also in amorphous materials¹⁴³⁻¹⁵⁶ and glasses,¹⁵⁷⁻¹⁷⁰ the temperature-dependence of \overline{W} is also not very substantial, although sometimes appreciable. 32 On the other hand, a dependence of \overline{W} on the concentration of impurity is clearly visible.¹⁴² As the latter declines, \overline{W} also declines, and the effect becomes very small. The dependence of \overline{W} on the conditions of preparation of the specimen is also very appreciable^{143,144} (in the cited studies the absorption of light in amorphous silicon films was studied). This suggests the conclusion that \overline{W} declines as the concentration of structural defects in the material declines.³²⁾

A substantial role played by structural defects in the formation of the optical tail is also evident from experiments on light absorption in radiation-damaged materials.^{171,172}

Recently the problem of the light-absorption coefficient in amorphous materials, especially in amorphous silicon, has become rather pressing in connection with the prospects of applying them in various photoelectric devices, including solar cells.¹⁷³ For this reason it seems expedient to summarize briefly here the fundamental theoretical results obtained up to now.

11. Phonons and static disorder

Apparently the first convincing theory that allows one to explain the Urbach rule in ionic crystals was proposed in Ref. 174. The idea is very simple: one assumes that the photon supplies only a part of the energy needed for transition of the electron from the valence band to the conduction band, while the energy deficit $E_s - \hbar \omega$ is covered by simultaneous absorption of phonons. Consistent application of the theory of many-phonon transitions indeed allows one to derive an expression for $\alpha(\omega)$ that satisfactorily approximates Eq. $(10.1)^{174}$:

$$\alpha \sim \exp\left[-\frac{E_g - \hbar\omega}{c_1 T}\right] \ln \frac{E_g - \hbar\omega}{c_s T} \Big].$$
(11.1)

Here c_1 and c_2 are constants characterizing the energy of interaction of the electrons with phonons and the form of the phonon spectrum. As $T \rightarrow 0$, naturally, the effect vanishes owing to the disappearance of phonons that might be absorbed. Subsequently the ideas of Ref. 174 were transferred¹⁷⁵ also to the case of optical transitions between the conduction (or valence) band and local levels lying in the forbidden band.³³⁾

On the other hand, one can hardly explain the lowtemperature component of the absorption coefficient in strongly doped (compensated) amorphous and vitreous semiconductors, nor its dependence on the conditions of specimen preparation, without taking into account the effect of the static disorder in the arrangement of atoms on the energy spectrum of electrons and holes.³⁴⁾ In other words, one must study the absorption of light in a random field. Here, as always happens in such problems,²³ the formula for the experimentally observed absorption coefficient can contain only quantities

³¹One can find a rather complete review of the experimental data accumulated up to 1971 in Ref. 138.

³²In line with what we have said in the Introduction, we are restricting ourselves to the minimum of literature references, and the author of this review hopes that those interested will not anathematize him for it: just a list of the last names would occupy more than a page here.

³³Important studies^{206,207} have recently appeared that are devoted to a more detailed treatment of the Urbach rule on the basis of the concept of multiphonon transitions. However, we note that the following remarks apply to them as well.

³⁴⁾An attempt was made in Ref. 158 to explain the tail of the absorption coefficient in vitreous arsenic sulfide by the influence of the zero-point vibrations of the atomic matrix. In the range of phonon energies of about 0.3 eV, the experimentally observed $\alpha(\omega)$ was satisfactorily matched. However, this study dealt with absorption involving transitions between the conduction band and a local level lying near the middle of the mobility gap. The problem of the dependence of α on the conditions of specimen preparation is blurred here to a considerable extent.

averaged over all configurations of the random field.³⁵⁾ They are all expressed in terms of correlation functions of the form:³⁶⁾

$$\Psi_n (\mathbf{r}_1, \ldots, \mathbf{r}_n) = \langle U (\mathbf{r}_1) \ldots U (\mathbf{r}_n) \rangle, \qquad (11.2)$$

Here the angle brackets denote the averaging noted above, while as before, $U(\mathbf{r})$ is the fluctuation in the potential energy of an electron at the point r: U = V $-\langle V \rangle$, where V(r) is the potential energy of the electron for some fixed configuration of the atoms. We note that the "effective" potential energy remaining after removing the periodic component V_{per} describing some auxiliary ordered system is used most often for V(r). For a strongly doped crystalline semiconductor, V_{per} is simply the potential energy of an electron in an ideal crystal (as derived, e.g., in some particular variant of the single-electron theory). The choice of V_{per} for an amorphous or vitreous semiconductor is less unequivocal.²³ In essence, by subtracting V_{per} we are using the same method as in the theory of polarons¹⁷⁶; here the final expression for the experimentally observed quantities contains only quantities of the type of the effective masses corresponding to the problem with the potential energy V_{per} . One must treat them as parameters of the theory to be determined experimentally (for more details, see Ref. 23, Sec. II.8).

An essential point below is that the effective potential energy can depend on the band index. Correspondingly, one sometimes writes $U_l(\mathbf{r})$ instead of $U(\mathbf{r})$, where *l* is the band index. Similar indices appear for the function Ψ_n . For example, we face a situation of this type when the random field stems from elastic stresses produced in the atomic matrix when structural defects of some particular type arise in it: generally the deformation potentials for the conduction and valence bands differ and can even have opposite signs.

In the macroscopically homogeneous system treated in this section and the next two sections, the function Ψ_n can depend on only n-1 differences $\mathbf{r}_1 - \mathbf{r}_n$, etc. We shall be especially interested in the binary correlation function of the random field $\Psi_2(\mathbf{r}_1, \mathbf{r}) \equiv \Psi(\mathbf{r}_1 - \mathbf{r}_2)$, since only quantities associated with it enter into the results obtained up to now. Its explicit form (just like the form of all the rest of the functions Ψ_n) depends on the nature of the forces acting on the electrons, or as one says, on the nature of the random field. General considerations imply only that the function $\Psi(\mathbf{r})$ must be bounded and must fall off to zero as $r - \infty$.

Three fundamental types of static random fields have been treated up to now in the theory of disordered systems.

a) Coulomb field.³⁷⁾ This is commonly taken to mean a screened field created by a set of point charges situated at random throughout the specimen (with or without some correlation between the coordinates of the charged centers). For example, such a field is created by the ions of a foreign dopant whenever no role is played by relatively small distances between them such that the ions cannot be treated as point charges. Here the screening arises either as a result of a spatial redistribution of the free or (in a compensated specimen) bound charge carriers, or because of spatial redistribution of the donors and acceptors themselves that arises in the process of specimen preparation. The expressions for the corresponding screening radius, r_0 $\equiv \pi^{-1}$ are well known in a number of cases.^{24,129,177} Below we shall treat × simply as a phenomenological parameter. Evidently in such a field the function V, and hence also Ψ , does not depend on the number of the band. Let us restrict the treatment to the very simple case in which the impurity ions are uncorrelated in space, while the potential energy V_{a} of an electron in the field of an individual ion is given by the ordinary Debye expression

$$V_a(r) = \frac{Z_a e^2}{e_r} \exp\left(-\varkappa r\right) , \qquad (11.3)$$

Here the subscript *a* specifies the type of impurity, *e* is the charge of an electron, and Z_a is the charge of the ion in units of |e|, while ε is the static dielectric permittivity of the material. The overall potential energy V of the electron is given here by the summation of the expressions (11.3) over all the ions, while we obtain the following expression for the binary correlation function:

$$\Psi(r) = \frac{2\pi n_{\xi}^{*} e^{4r_{0}}}{e^{4}} \exp\left(-\frac{r}{r_{0}}\right), \qquad (11.4)$$

Here

$$\boldsymbol{n}_{\boldsymbol{b}}^{*} = \sum_{\alpha} n_{\alpha} Z_{\alpha}^{*} \tag{11.5}$$

is the effective concentration of the charged impurity.

Also the intrinsic (in the sense indicated in Ref. 23, Sec. II.8) random field in disordered semiconductors proves to be of Coulomb nature when the chemical bonding has even partially ionic character and only shortrange order exists. In particular, such materials include the vitreous chalcogenide semiconductors-arsenic selenide and sulfide, etc. In general one cannot use Eq. (11.4) in this case, since correlation unquestionably exists in the arrangement of the intrinsic (nonimpurity) atoms of the material. However, in contrast to crystals, the corresponding correlation length r_{e} is finite in noncrystalline materials. If it is small in comparison with the screening radius, then if we neglect quantities of the order of $r_c r_0^{-1}$, we again obtain Eqs. (11.4) and (11.5). Here, in contrast to doped crystalline semiconductors, the subscript a specifies the chemically different types of atoms of the main material, and we must take n_a to mean the concentration of

³⁵⁾Further, to a considerable extent this section repeats the content of one of the sections of the book of Ref. 23, so that we have omitted the detailed conclusions and literature references given in the cited book. At the same time, for the reader's convenience, it seemed reasonable to repeat here the fundamental physical arguments and results.

³⁶)Sometimes these functions are not introduced explicitly by employing directly the distribution function of any of the quantities of physical interest (e.g., the value of the intensity of the random field). Of course, these two approaches are completely equivalent.

³⁷⁾To avoid misunderstanding, we stress that the term "Coulomb" is used here only in the narrow sense pointed out below.

atoms of the given type, while Z_a is the corresponding fractional ionic bond character.

b) Smooth field. This is what a field from any source is called if on the average it varies sufficiently smoothly in space. Namely, let us introduce the quantities

$$\psi_{1} = \langle U^{2} (\mathbf{r}) \rangle,$$

$$\psi_{2} = \nabla_{\mathbf{r}} \nabla_{\mathbf{r}'} \Psi | (\mathbf{r} - \mathbf{r}') |_{\mathbf{r}' = \mathbf{r}} \equiv \langle (\nabla U)^{2} \rangle.$$
(11.6)

Their physical meaning is obvious; we note that they do not depend on \mathbf{r} in the macroscopically homogeneous system that we are treating. Further, let us denote by m the smallest of the effective masses corresponding to the auxiliary problem pointed out above. Then the condition for "sufficient smoothness" is reduced to the inequality

$$\frac{\hbar^3 \psi_a}{4m \psi_i^{3/2}} \ll 1.$$
 (11.7)

The explicit form of $\Psi(\mathbf{r})$ plays no role here. The experimentally measurable quantities are expressed only in terms of the parameters ψ_1 and ψ_2 .

A smooth field arises, for example, under conditions in which one can replace a system of point charges with a continuous charge distribution, i.e., shift to a formulation of the problem characteristic of Lorentz electrodynamics.³⁸⁾ In particular, this is the situation when the random field is created by a system of three-dimensional extended defects of technological origin chaotically arranged in the specimen or by defects of the type of clusters arising in the specimen (including crystalline defects) due to neutron irradiation.¹⁷⁸⁻¹⁸¹

Defects of this type also arise in crystals.^{213,214} For this reason many specimens that had previously been considered pure and "defect-free" actually may not be so. We note that, precisely because of its smoothness, the random field being discussed may not—in a certain temperature region—affect too strongly the mobility of the charge carriers (the appropriate formulas were derived in Refs. 215–217). At the same time, the influence of this field can affect the more subtle effects, in particular the magnitude of the optical tail. In this sense perhaps one must include even many crystalline specimens—and not at all only those strongly doped among the disordered materials.

Thus the study of effects involving the existence of a smooth random field (e.g., the discovery of a tail of the absorption coefficient in a weakly doped crystal) can serve for non-destructive quality control of specimens. As we shall see below (Sec. 13), the same is true of "pure" amorphous silicon, i.e., not doped with hydrogen, fluorine, etc.

Finally, a smooth field of any origin can appear in the presence of intermediate-range order in the arrangement of the atoms.182 39)

We note that the Coulomb field is not smooth, whatever the screening radius r_0 is. Actually, we can easily convince ourselves that in the case of (11.4) the parameter ψ_2 diverges [owing to the singularity of the potential in (11.3) at small distances]. Formally this involves the singularity of the second derivatives of Ψ with respect to the components of the vector **r**. The binary correlation function of a smooth isotropic random field can contain only integral powers of \mathbf{r}^2 .

c) Field of short-range forces. The meaning of this term in this case needs no explanation. A random field of this type arises, e.g., as a result of a not fully ordered-including random disorder-arrangement of atoms in a homopolar material. Upon considering the potential energy (or the pseudopotential) of an electron in the random field as the sum of its energy V_{a} in the field of the individual atoms, we can easily convince ourselves that the bindary function $\Psi(r)$ differs appreciably from zero only if the distance r does not exceed the range of the corresponding forces. If the latter is considerably smaller than the other characteristic lengths of the problem (e.g., the wavelength of the light, if one is referring to light absorption, the localization radius of an electron, etc.), then one can approximate the function Ψ with the expression

$$\Psi(\mathbf{r}) = \Phi_0 \delta(\mathbf{r}) \tag{11.8}$$

(the "white-noise" case). Under certain conditions the same formula also characterizes the field of elastic deformations arising from fluctuations of composition in an inhomogeneous solid solution. Here Φ_0 depends on the band indices c and v. One can find explicit expressions for the parameter Φ_0 for the two stated cases in the book of Ref. 23.

The function (11.8) is the "most singular" one of those that we have treated. For it, one can determine neither the parameter ψ_2 nor ψ_1 . The role of the latter is played by the quantity

 $\Phi_0 = \int \Psi(\mathbf{r}) \, \mathrm{d}\mathbf{r}.$

We note that the correlation functions of Coulomb and smooth fields are characterized by at least two independent parameters, of which one determines the mean length of the fluctuations, and the other determines the correlation radius, i.e., the distance at which the function $\Psi(r)$ falls off rather rapidly. In a smooth field these are $\psi_1^{1/2}$ and $(\psi_1/\psi_2)^{1/2}$, and in a Coulomb field they are $\psi_1^{1/2} = (2\pi n_i^* e^4 r_0 \varepsilon^{-2})^{1/2}$ and r_0 . On the other hand, only one parameter enters into the right-hand side of (11.8): Φ_0 . The correlation radius was formally assumed to be zero. The latter circumstance determines the special simplicity of the function in (11.8),

³⁸Non-random force fields of this type are very well known in semiconductor physics. They arise in nearly every n-pjunction if the Debye length is great enough in comparison with the de Broglie wavelength (in this case one must simply insert the potential energy of an electron and the modulus of its gradient into (11.7) instead of $\psi_1^{1/2}$ and $\psi_2^{1/2}$).

³³The somewhat vague term "intermediate-range order" that we have just used has been defined up to now in a more or less intuitive way. It means only that we are dealing with a distribution of atoms that are ordered, e. g., within a region with linear dimensions of tens of mean atomic distances, rather than within one or two coordination spheres or throughout the specimen.

which has for this reason enjoyed great popularity among theoreticians. There are problems in which the approximation (11.8) is admissible and quite sufficient. However, in studying the density of states or the lightabsorption coefficient, the form of the correlation function, and especially its analytical structure, affects the result.

12. The random field and the Urbach rule

The frequency-dependence of the light-absorption coefficient of semiconductors having a random field has been calculated in a number of studies. Table I summarizes the results obtained²³ under the assumption that $E_{a} \gg \overline{W}$, neglecting the interaction of the electrons and holes with phonons, and without taking into account the creation of excitons. The latter effect hardly plays a role in semiconductors with a Coulomb random field, since at large impurity concentrations the screening radius proves to be small. If it is smaller than the Bohr radius of an exciton (as happens under conditions of strong doping), then the exciton does not exist at all as a stationary state. In a material with a smooth field that does not depend on the band number, excitons can exist if the value of \widetilde{W} is smaller than the energy of ionization E_{J} of an exciton. In the converse case the forces exerted on the electron and hole by the random field "tear the exciton apart" (under the stated conditions they are oppositely directed). As soon as the inequality $\overline{W} < E_1$ holds, one must replace the quantity \overline{W} with

$$\overline{\widetilde{W}} = \frac{\hbar^2 \psi_a}{27 m_* W_a^2}.$$
(12.1)

Here we have

 $W_{\rm B}=\frac{m_{\rm I}e^4}{2e^2\hbar^3}.$

The "tail" corresponds here to the region of energies lying below the unperturbed exciton level:

TABLE I. Frequency-dependence of the light-absorption coefficient in a disordered semiconductor for $\hbar \omega < E_g$.

Random field	α	w
Coulomb (withour cor- relation in the distri- bution of point charges)	Eq. (10.1)	2,2 $W_{\rm B} (n_{\rm t}^{\bullet} a_{\rm B}^2)^{2/5}$ $W_{\rm B} = \frac{e^2}{2ea_{\rm B}}, a_{\rm B} = \frac{e\hbar^2}{m_{\rm R}e^2}$
Smooth	Eg. (10.1)	$\left(\frac{\hbar^2 \psi_2}{36m_r}\right)^{2/3}, m_r^{-1} = m_n^{-1} + m_p^{-1}$
Short-range forces	$\sim \exp\left(-\sqrt{\frac{E_{g}-\hbar\omega}{\overline{W}}}\right)$	$\left(\frac{(2m)^{3/2}\Phi_{0,c,v}}{16\pi\hbar^3}\right)^2, m=\sqrt{m_nm_p}$
		$\Phi_{0,c,v} = (\Phi_{0,c} \Phi_{0,v})^{1/2}$

1. The symbols $c_r m_n$, and m_p respectively denote the dielectric permittivity of the material and the effective masses of the electrons and holes corresponding to the auxiliary problem with a periodic field. For similicity we have assumed isotropic dispersion laws; the generalization to the anisotropic case¹⁴¹ reduces to replacing m_n and m_p by some combination of effective masses. 2. In strongly doped semiconductors a tail of the absorption coefficient can be observed only with strong enough compensation. In the opposite case one will observe the well known Burshtein Moss shift instead of a tail (Ref. 24, Sec. XVIII, §7). 3. The formula for W in a Coulomb field was derived under the assumption that $m_n < m_p$. With the conversion size of the ingenuity one must tendece $m_n \to m_p$.

S. The composite sign of the inequality one must replace m_n by m_p . 4. In a Coulomb random field the absorption coefficient is described by Eq. (10.1) only in the

region $E_g - \hbar\omega \leq 2W$. Upon further decrease in the photon energy the $\alpha(\omega)$ relationship be sharper. When $E_g - \hbar\omega > W$, one obtains the following expression instead of (10.1): nship proves to

 $\alpha \sim \exp[-((E_g - \hbar\omega)/W) \cdot 1.1]^{5/3}$.

The result obtained for a field of short-range forces holds asymptotically for $W > E_g - \hbar\omega$. S. The result for a smooth field has been derived under the assumption that it is Gaussian.²³ The generalization to the non-Gaussian case²³ reduces to replacing ψ_2 with another parameter of analogous type.

$$\hbar\omega < E_{\rm g} - W_{\rm B}.$$

The problem of the role of excitons in light absorption in a disordered semiconductor with a random field of short-range forces has not yet been studied with the requisite completeness, insofar as we know.

In connection with the results presented in Table I, we should make two remarks.

First, in Coulomb or smooth random fields, the frequency-dependence of the absorption coefficient for $\hbar \omega < E_{\mu}$ does not at all match the density of states $\rho(E)$ as a function of the energy E in the tail formed below the bottom of the conduction band or above the ceiling of the valence band, nor the combined density of states (for more details, see Ref. 23, Chap. V). Only in a material having short-range forces do the functions $\rho(E)$ and $\alpha(\omega)$ prove to be asymptotically of the same form, while differing only in the coefficients of $\sqrt{E_c - E}$ and $\sqrt{E_{\mu} - \hbar \omega}$.²⁰¹ This makes it obvious that the attempts to employ the data on the absorption coefficient directly for studying the density of states in the mobility gap that unfortunately are occasionally being undertaken even now seems extremely risky.

Second, in recent studies^{184,185} using the model of a semiconductor with a random field of short-range forces, an expression of the type of (10.1) has been derived (numerically), both for the absorption coefficient and for the density of states in the tails near the conduction and valence bands (with replacement of $E_{\mu} - \hbar \omega$ by the absolute value of the energy at a given point of the mobility gap as measured respectively from the bottom of the conduction band or the top of the valence band). Since the formula indicated in Table I, just like the analogous expressions for the density of states, amounts to an exact asymptotic form of the explicit analytic expression, we should suppose that the results of Refs. 184 and 185 are valid only in a certain restricted energy region. Of course, this does not deprive them of interest; on the other hand, it is hardly possible to accept these results as universal, and apparently it is not justified to compare their calculations with Coulomb or smooth random fields: the point is that the systems are physically essentially different. We also note that the cited studies still have no *a priori* estimate of the parameter \overline{W} (in the notation of Refs. 184 and 185, the parameter Γ).

It has proved necessary to generalize the results in order to take into account the interaction of electrons and holes, not only with the random field, but also with phonons, first of all in the problem of the optical tail in crystalline semiconductors.²⁸⁶ We are dealing with materials in which indirect transitions would dominate in the absence of the random field. As the above-cited studies of B. Esser have shown, the formulas indicated in Table I for the Coulomb and smooth fields remain in force in this case as well; only the pre-exponential factor α_m is changed in Eq. (10.1). In particular, this factor becomes temperature-dependent-owing to the number of phonons absorbed in indirect transitions. On the other hand, a certain temperature-dependence also of the parameter W is observed in chalcogenide glasses.^{161,167,168} This has required generalization of the theory in order to take into account jointly both the effect of the random field and of the possibility of electron transitions with simultaneous absorption of phonons. Such a generalization has been proposed in two series of studies (Refs. 186-188 and 189, 190). The former of these studied transitions between deep local levels lying in the mobility gap, and between these levels and regions of the continuous spectrum (both light absorption and luminescence). The nature of the random field was not specified explicitlyinstead of this assumptions were made on the spectrum of the discrete levels. They either could be connected with some type of inherent defects or with fluctuational defects. In the former case the role of the random field consisted, by hypothesis, simply of a Gaussian "smearing" of the peaks of the density of states. In the latter case an exponential form of the density of states in the tails was postulated.⁴⁰⁾ Upon combining these ideas with the standard theory of many-phonon transitions in solids,¹⁹² the authors constructed graphs for the temperature-dependence of the light-absorption coefficient in the tail and for the frequency-dependence of the luminescence intensity of disordered semiconductors for transitions of different types. Apparently these results still await comparison with experiment.

The second series of studies dealt with light absorption in smooth and Coulomb random fields with simultaneous emission of an arbitrary number of phonons. The influence of the random field was treated within the framework of the same approximations as in Ref. 23 (both without and with taking exciton effects into account). Here it proved possible in the Condon approximation to employ the ordinary theory of many-phonon transitions. As a result, the expression (10.1) was again obtained for the absorption coefficient (under the same conditions as in the absence of many-phonon transitions), but with a function \overline{W} that depended on the heatrelease parameter and on the temperature. The calculation was carried out to completion in the case of dispersion-free optical (or local) phonons characterized by the frequency ω_0 . As should have been expected, the temperature-dependence of \overline{W} becomes weaker as the dimensionless (in units of $\overline{\pi}\omega_0$) heat-release parameter decreases and as the following parameter increases:

5		Ŵ
		ħω

At low enough temperatures, $T \ll \hbar \omega_0$, this dependence almost vanishes, as should have been expected—hardly any phonons of the type being discussed remain. At higher temperatures the curve $\overline{W}(T)$ proves to be almost linear. Thus, apparently one can¹⁹⁰ quantitatively explain the experimentally observed^{167,168} temperaturedependence of the characteristic energy \overline{W} in vitreous arsenic selenide and sulfide (in line with what we have said in Sec. 11, the random field in these materials was assumed to be a Coulomb field).

13. The Urbach rule in amorphous silicon

Upon examining the experimental data on amorphous silicon, we should note first of all that the problem of determining experimentally the optical width of the forbidden band in a disordered semiconductor is more complex than in a crystal. Actually, the optical width of the forbidden band is taken to mean (see Ref. 23, Secs. II.12 and V.1) the energy spacing between the local boundaries of the conduction and valence bands. However, outside these boundaries the density of states does not at all vanish-a dense spectrum of discrete levels lies everywhere there (the tail of the density of states). Ultimately this gives rise to the appearance of the tail of the absorption coefficient ("optical tail") in the absence of phonon effects. The absence of a sharp red absorption edge compels us to define E_g more or less arbitrarily. Sometimes the frequency region in which the light-absorption coefficient is less than 10⁴ cm⁻¹ is assigned to the optical tail. In other studies the following formula is used for defining E_{a} :

$$\alpha \sim (\hbar \omega - E_{\rm g})^2.$$

This is often justified when $\hbar\omega > E_g^{210}$ by extrapolating it to $\hbar\omega - E_g$. However, in both cases it remains not entirely clear whether one is attributing here a fraction of the states to the conduction band (or the valence band or both) that actually belong to the discrete spectrum.

Further, the very term "amorphous silicon" still does not characterize the specimen fully. The question is not only of the natural dependence of many characteristics of the material on the composition and concentration of doping impurities and on the method of preparation, but also of the fact that amorphous alloys of sili-

⁴⁰⁾Such a semiphenomenological way of taking the random field into account might cause a certain degree of discontent. However, it is not superfluous to note that, according to model calculations,^{184, 191} a relationship of this type actually can approximate the course of the density of states $\rho(E)$ over a rather broad interval of energies *E*.

con with hydrogen or fluorine (or chlorine) have already been of special interest for several years. Recent mention has been made of alloys of silicon with other elements—tin or germanium. Here the concentration, e.g., of hydrogen can amount to 20% or more. The properties of these alloys and of "pure" amorphous silicon prove to differ substantially.

It was shown even in the early studies¹⁴⁴ ¹⁴⁵ that amorphous silicon (1971 specimen, i.e., not specially doped), manifests an absorption tail or not, depending on the conditions of preparation. Here the former possibility was realized, as a rule, in specimens annealed insufficiently long (for several hours), but it disappeared on longer annealing. This naturally led to the conclusion that the absorption of photons with frequency $\omega \leq E_{g}/\hbar$ is not an "intrinsic" property of amorphous silicon, but a property of a specimen with some type of structural defects. In later studies (see, e.g., Refs. 146-156) on the alloy of amorphous silicon with hydrogen, light absorption in this frequency region was regularly found even under different conditions of preparation. In specimens containing hydrogen but not especially doped with boron or phosphorus, the logarithm of the absorption coefficient α proved to be an almost linear function of the frequency [i.e., Eq. (10.1) is confirmed in the interval of values of α from $\leq 10^4$ cm⁻¹ to $\approx 1 \text{ cm}^{-1}$ (this corresponds to photon energies from E_{g} $\approx 1.7 \text{ eV to} \leq 1.3 \text{ eV}$). At the same time the value of \overline{W} varies in the range from ≈ 0.05 eV to ≈ 0.09 eV-depending on the conditions of preparation of the film (in particular, on the temperature of the substrate maintained while depositing the film from a glow discharge and on the temperature of the subsequent annealing). At the same time, the most frequently employed value \overline{W} $\approx 0.06 \text{ eV}^{41}$ is obtained also for films produced by sputtering. Thus it is not yet entirely clear precisely on what conditions of preparation does \overline{W} depend.⁴²⁾ In this connection we should note that this characteristic energy depends only weakly on the concentration of hydrogen. The idea naturally arises that the fundamental responsibility for the formation of the optical tail is borne by the random field, which is not associated with hydrogen per se, but with structural defects of technological origin, perhaps including also non-point defects, that constantly and stably arise in the process of formation of the film. Correspondingly, one can hardly be amazed at the relationship established experimentally in the cited studies between the quantities E_{e} and \overline{W} :

the former declines linearly as the latter increases. One of the possible reasons for this has been pointed out at the beginning of this section: as the density of states increases in the tail, the danger also increases of assigning them, e.g., to the conduction band. The essential point here is that the characteristic energy that determines the density of states in the tail differs from W only by a factor that either does not depend on the effective concentration of defects (the case of short-range forces) or depends on it only very weakly. Thus, in a smooth field, and also in a Coulomb field with the correlation function (11.4), the corresponding characteristic energy is $\psi_1^{1/2} = [\Psi(0)]^{1/2}$. Upon using the expression for \overline{W} from Table I, we can easily convince ourselves that we have the following expression in the case of (11.4):

$$\psi_1^{1/2}/\overline{W} \sim (r_0 a_{\rm B}^{-1})^{1/2} (n_{\rm t}^* a_{\rm B}^3)^{1/10} \frac{\sqrt{2\pi}}{14}.$$

The screening radius r_0 can also depend on n_t^* . Here, depending on the screening mechanism, one obtains different powers of n_t^* on the right-hand side of the latter relationship, but none of them exceeds 3/20 in modulus. An analogous situation arises also in a smooth field.

With all the attractiveness of this viewpoint, apparently it still is not the only possible one. The point is that amorphous-silicon films containing hydrogen prove to be spatially inhomogeneous on scales that cannot be taken as microscopic.¹⁹³⁻¹⁹⁸ They can not only contain internal cavities and/or inhomogeneities resembling pebbles in shape (with linear dimensions about 1000 Å), but also possess "columnar" structure: a film obtained by the glow-discharge method is penetrated throughout by columns with transverse linear dimensions of the order of 10 µm. A more complicated structure of inhomogeneities is also possible, but only one point is essential in principle: amorphous silicon containing hydrogen amounts to an essentially two-phase system. Here the spatial distribution of either phase is to a certain extent random. The hydrogen concentration $n_{\rm H}$ in these two phases proves to be different,¹⁹⁶ although it is not yet fully clear to what extent. However, let the value of $n_{\rm H}$ in one of the phases be rather large—so much so that, despite the small ionic character of the Si-H bond, the Coulomb random field becomes substantial, arising from the disordered distribution of these bonds in space. The expression for \overline{W} given in Table I can prove in this case to be rather inaccurate, since the possibility of neglecting the correlation in the distribution of Si-H bonds is at best not evident. For orientation, let us try to use it nevertheless. In the SiH₄ molecule the effective charge of an H atom amounts to -0.05|e|, while correspondingly, that of the silicon atom is 0.20 |e|.¹⁹⁹ These values can change somewhat in a condensed medium, but hardly so much that it would make sense to take this into account in such a rough orientational estimate. As the Bohr radius a_B we shall take the "crystal" value $a_{\rm B} = 1.5 \times 10^{-7}$ cm. We note that $a_{\rm B}^{1/5}$ enters into the formula for \widetilde{W} . That is, the numerical coefficient here plays no special role. Then we can easily convince ourselves that one obtains the value W = 0.06 eV if the concentration of bound hy-

⁴¹⁾The value $\overline{W} = 0.06 \pm 0.01$ eV was given in Ref. 154b (the film was prepared by deposition from a glow discharge). ⁴² With a relatively low temperature (100 °C) of the substrate in the specimen-preparation conditions, the Urbach rule holds in a narrower energy region—from 1.7 eV to ~1.3— 1.4 eV (depending on the annealing temperature); the latter energy corresponds to $\alpha \approx 10^2$ cm⁻³. Beyond this a smoother dependence sets in—there is a broad absorption band sometimes extending (depending on the annealing temperature) to 0.9 eV. According to Ref. 154a, the light-absorption coefficient in this band involves the concentration of unpaired spins as determined in experiments on electron paramagnetic resonance (it can attain values of 10¹⁸ cm⁻³). Apparently this means that, under the given conditions of preparation, hydrogen does not block all the "dangling bonds."

drogen amounts to about 3×10^{22} , 2×10^{22} and 1.5×10^{22} cm⁻³ if one is speaking respectively of monohydride (Si-H), dihydride (Si-H₂) and trihydride (Si-H₃) groups. Apparently the first of these values corresponds to the alloy Si: H with 20 atomic percent hydrogen. We note that, if one putatively increases the hydrogen concentration by a factor of five, then \overline{W} increases by a factor of only 1.9. This makes it clear that the value of \overline{W} found above is not so far from the maximum attainable in the material that we are studying. If (as seems probable) the formation of such bonds is energetically favorable, then the concentration of hydrogen in the given phase can prove to depend only weakly on the mean experimentally measured concentration of hydrogen in the given specimen.

Apparently both the study of the nature and the role of structural defects obtained in a stable manner and careful examination of the two-phase system merit more serious attention on the part of theoreticians than has been paid up to now. At the same time we should point out another circumstance that arises in connection with the optics of a spatially inhomogeneous material. The point is that both the expressions presented in Table I and the standard phenomenological formulas of electrodynamics employed in processing the experimental data on light absorption have been derived within the framework of the problem of propagation of light in a macroscopically homogeneous medium. What we have said above makes it clear that this assumption is not fully justified as applied to films of amorphous silicon. Apparently one can still use the formulas of Table I within the limits of each phase, but the electrodynamic part of the problem must be reexamined. Such an attempt has recently been undertaken^{55,200} in connection with the problem of light absorption in films of a-Si: H having a columnar structure. The heart of the matter is evident from Fig. 7, where such a film is schematically shown in plan view. The hatched circles are the ends of the columns, assumed cylindrical for simplicity. According to Ref. 196 the columns consist of "real" amorphous silicon containing hydrogen, while a medium of not entirely clear composition ("fabric") lies between them. We see that we are essentially dealing with a system of waveguides randomly distributed in space, the columns playing this role. From the standpoint of ordinary radiotechnology these are rather strange waveguides, since they are filled with and surrounded by matter that absorbs the electromagnetic radiation, but still they are waveguides with their inherent features.

Let the thickness of the film be L. Then the intensity S of a wave that has passed through the film in a direc-



FIG. 7. An *a*-Si:H film as a system of randomly distributed waveguides.

tion perpendicular to the plane of Fig. 7 (i.e., along the columns) is not given by a single exponential function $e^{-\alpha L}$, as would happen in a homogeneous medium, but by an infinite series

•

$$S = \sum_{n \ge 1} c_n e^{-\alpha_n L}. \tag{13.1}$$

Here the c_n are coefficients that depend on the boundary conditions at the illuminated surface, while the quantities α_n are expressed in terms of the eigenvalues of the boundary value problem of the propagation of waves in the system of waveguides being studied (we are selecting the numbers α_n in increasing order). The essential point is that they depend also on the light-absorption coefficients, both in the column (α_n) and in the tissue (α_m) calculated as though we were dealing with the corresponding homogeneous medium:

$$\alpha_{k,m} = \frac{4\pi\sigma_{k,m}}{c\,V\,\overline{\epsilon_{k,m}}}.$$
(13.2)

Here $\sigma_{\mathbf{k}}(\sigma_{\mathbf{m}})$ and $\varepsilon_{\mathbf{k}}(\varepsilon_{\mathbf{m}})$ are the real components of the electric conductivity and the dielectric permittivity of the material of the columns (or the medium) at the frequency of the light being absorbed. If we have

$$\alpha_h L \gg 1, \quad \alpha_m L \gg 1,$$
 (13.3)

then only one term, $c_1 e^{-\alpha_1 L}$, remains in the series (13.1). Here the quantity α_1 is proportional to the lesser of the coefficients α_k and α_m . Here we formally return to the problem of a homogeneous medium. However, actually α_1 can depend on the radius of the columns. Thus, if all the columns are identical, $\alpha_k \ll \alpha_m$, and the mean distance between the columns is large in comparison with α_m^{-1} , then calculation yields²⁰⁰

$$\alpha_{t} = \alpha_{k} \left[1 - 0.086 \left(\frac{\lambda}{R \sqrt{\varepsilon_{k}}} \right)^{2} \right]^{-1/2}.$$
(13.4)

Here *R* is the radius of the columns, while λ is the wavelength of the light in vacuo. We assume that the expression in brackets is large in comparison with $2\pi\sigma_k/\omega\varepsilon_k$ and that $0.086\lambda^2 R^{-2}\varepsilon_k^{-1} < 1$.

We see that, even in this relatively simple case, an attempt to interpret the results obtained in transmission experiments by using the ordinary formulas of electrodynamics entails errors in determining the frequencydependence of the light-absorption coefficient. In fact, under conditions in which the Urbach rule holds, this can be not very essential, since then α_{\star} depends exponentially on the frequency, i.e., far more strongly than the second factor in Eq. (13.4). However, the situation can become very serious if even one of the inequalities of (13.3) is not satisfied. Then the right-hand side of (13.1) does not reduce to a single term, and the description of the experimental data in the ordinary language of a homogeneous medium can yield completely erroneous results. Thus, if we assume, "by definition", that

$$S \sim \exp(-\alpha L)$$
,

we would draw the "conclusion" that the light-absorption coefficient depends on the thickness of the film in the absence of any real cause of it.

We note in closing that, as was noted in Ref. 200, Eq. (13.1) in essence is not necessarily associated with the columnar structure model discussed above. Apparently the only essential point is the presence of internal phase boundaries. It is this circumstance that gives rise to the boundary value problem, which generally admits more than one eigenvalue.

The author is very grateful to I. P. Zvyagin, A. G. Mironov, and R. A. Suris for a critical reading of the manuscript and useful remarks, and also to Sh. Abe and Yu. Toyazawa, F. Wegner, A. A. Gogolin, N. F. Mott, M. Olivier and F. Boucher, M. Pepper, J. L. Pichard, and J. Sarma, M. Pollak, and T. Tanaka for kindly sending preprints.

APPENDIX I. CORRECTIONS TO THE RESULTS OF THE KINETIC EQUATION FOR $T \rightarrow 0$

A dependence of the static electrical conductivity σ on the dimensions L of the specimen can arise also within the framework of the kinetic equation. This can happen if, e.g., the charge carriers are scattered by a charged impurity and the Debye screening radius is comparable with L (see Sec. I.5). However, we shall be interested here in the dependence of σ on L that arises when we take into account the corrections to the kinetic equation. It will suffice below to treat the case of weak scattering $(g \gg 1)$ by randomly distributed point centers that interact with the electrons via short-range forces. 50-52, 59, 60, 65 Let us denote the concentration of the centers, the concentration of electrons, and the amplitude of elastic scattering of an electron by an individual center respectively by n_i , n, and v. It is convenient to use the graph method; here the dotted lines in the diagrams will correspond to the factors v, and the solid lines to the averaged single-electron Green's functions calculated with account taken of elastic scattering; each circle corresponds to a factor n_i . The static electrical conductivity (at T = 0) is given by the expression

$$\sigma = \sum_{\mathbf{k}} \kappa^2 \left[\underbrace{\mathbf{k}}_{\mathbf{k}} \right] + \sum_{\mathbf{k}, \mathbf{k}'} (\mathbf{k}, \mathbf{k}') \left[\underbrace{\mathbf{W}(\mathbf{k}, \mathbf{k}')}_{\mathbf{k}, \mathbf{k}'} \right]. \tag{I.1}$$

Here W(k,k') denotes the two-particle Green's function:

$$w(\mathbf{k},\mathbf{k}) \equiv \underbrace{\mathbf{k}}_{\mathbf{k}} \underbrace{\mathbf{k}}_{\mathbf{k}} = \underbrace{\mathbf{0}}_{\mathbf{k}} + \underbrace{\mathbf{0}}_{\mathbf{k}} \underbrace{\mathbf{0}}_{\mathbf{k}} + \underbrace{\mathbf{0}}_{\mathbf{k}} \underbrace{\mathbf{0}}_{\mathbf{k}} + \cdots$$
(I.2)

Here, since we are dealing with elastic scattering, we have $k^2 = k'^2$ under conditions of weak interaction. As one can show,^{51, 56, 65} in the higher orders in *n* with d = 2, the essential contribution to $W(\mathbf{k}, \mathbf{k}')$ proves to arise from the graphs in which all the dotted lines intersect. This contribution has the form

$$\widetilde{W} = \underbrace{\circ} \cdot \underbrace{\circ} + \underbrace{\circ} \underbrace{\circ} \underbrace{\circ} + \cdots$$
(I.3)

The contribution from the graph having s dotted lines is given by the expression

$$W^{(s)} = (nv^{3})^{s} \left[\frac{2\pi \rho(F)}{\sqrt{4p^{2} + v_{p}^{2}(\mathbf{k} + \mathbf{k}^{2})^{3}}} \right]^{s-1}.$$
 (I.4)

Here $\rho(F)$ is the density of states at the Fermi level,

 v_F is the Fermi velocity, and Γ is the imaginary component of the mass operator associated with the singleelectron (averaged) Green's function.⁴³⁾ Within the framework of the model that we have adopted, the expression for Γ is well known:

$$\Gamma = \pi \rho (F) n v^2. \tag{I.5}$$

The summation of all such graphs (over s) forms a geometric progression. Upon summing it, we obtain

$$\widetilde{W}(\mathbf{k}, \mathbf{k}') = \sum_{s \ge 2} \widetilde{W}^{s}(\mathbf{k}, \mathbf{k}') = nv^{aJ}(\mathbf{k}, \mathbf{k}') [1 - nv^{aJ}(\mathbf{k}, \mathbf{k}')]^{-1}.$$
(I.6)

Here we have

$$J_{k}(\mathbf{k}, \mathbf{k}') = \frac{2\pi\rho(F)}{\sqrt{4\Gamma^{2} + v_{F}^{2}(\mathbf{k} + \mathbf{k}')^{2}}}.$$
 (I.7)

Evidently the contribution being discussed is especially large when $(\mathbf{k} + \mathbf{k'}) \rightarrow 0$, i.e., in back-scattering. Actually, if we assume

 $4\Gamma^2 \gg v_F^2 (\mathbf{k} + \mathbf{k}')^2,$ (I.8)

we obtain

$$\widetilde{W} = nv^{2} \left[-1 + \frac{\Gamma \left(1 + \frac{1}{2} l^{2} \left(k + k' \right)^{3} \right)}{\pi \rho \left(l^{2} \right) nv^{2}} \right]^{-1}.$$
(I.6')

Here $l = (1.2) v_{\rm F} \Gamma^{-1}$ is the mean free path for momentum. In view of (I.5), this gives

$$\widetilde{W} = \frac{2nv^3}{l^2 \left(\mathbf{k} + \mathbf{k}'\right)^2} \,. \tag{I.9}$$

Correspondingly, the second term on the right-hand side of (I.1) is proportional to

$$\int d\mathbf{k} \int d\mathbf{k}' \, \frac{(\mathbf{k}, \, \mathbf{k}')^2}{l^2 \, (\mathbf{k} + \mathbf{k}')^2} \approx - \int d\mathbf{k} \int d\mathbf{k}' \, \frac{\mathbf{k}^2}{l^2 \, (\mathbf{k} + \mathbf{k}')^2}. \tag{I.10}$$

We can conveniently perform a substitution of variables by taking $\mathbf{k} + \mathbf{k'} = \mathbf{p}$. Then the following integral appears on the right-hand side of (I.10):

$$J = -\int \frac{\mathrm{d}\mathbf{p}}{p^{\mathbf{a}}}.$$
 (I.11a)

The upper limit here is not essential for our purposes: it determines the finite (and independent of L) contribution to σ and g involving the breakdown of the approximate equation $\mathbf{k} + \mathbf{k}' = 0$. The lower limit as T - 0 is determined by the condition—apart from the factor 2π —that the wave number must not be smaller than the reciprocal length of the specimen.⁴⁴⁾ Thus we obtain (writing out only the value of the integral at the lower limit and denoting by c_1 , c_2 , c_3 and σ_B some positive constants and the value of the conductivity obtained from the kinetic Boltzmann equation):

⁴³⁾ We have omitted the real part of the mass operator on account of its smallness. The validity of neglecting it is not always evident, but it does not affect the appearance of the singularity treated below.

⁴⁴⁾When $T \neq 0$, another lower bound is imposed on the minimum value of the wave number. In fact, inelastic scattering becomes possible here. Consequently the treatment adopted above is valid only as long as the length of the specimen is smaller than the characteristic distance L_{eff} to which an electron can diffuse in the field of an impurity without undergoing inelastic scattering. When $L > L_{eff}$, the lower limit of the integral in (I. 11) must be replaced with L_{eff}^{-1} . According to Ref. 203, we have $L_{eff} = \sqrt{ll_1}/2$, where l_1 is the mean free path for inelastic scattering.

a)
$$d = 1$$
:

$$J = -L, \quad g = L^{-1}\sigma_{\rm B} - c_{\rm L}. \tag{I.12a}$$

Hence we have

$$\beta = -1 - \frac{c_1}{g}, \quad \frac{d\beta}{dg} = \frac{c_1}{g^2}. \quad (I.13a)$$

b) *d* = 2:

$$J = -\ln L, \quad g = \sigma_{\rm B} - c_{\rm a} \ln L. \tag{I.12b}$$

Hence we have:

$$\beta = -\frac{c_2}{g}, \quad \frac{d\beta}{dg} = \frac{c_2}{g^2}. \quad (I.13b)$$

c)
$$d = 3$$
:

 $J = L^{-1}, g = Lo_{\rm B} + c_{\rm s}.$ (I.12c)

Hence we have

$$\beta = \mathbf{i} - \frac{c_s}{g}, \quad \frac{d\beta}{dg} = \frac{c_s}{g^{\mathbf{k}}}. \tag{I.13c}$$

We see that in the three-dimensional case the dependence of σ on L is quite insignificant; it vanishes in the limit as $L \rightarrow \infty$. On the other hand, it is important for d = 2. This dependence would prove to be even more important in a one-dimensional system. However, it is already clear there^{35,36} that as $T \rightarrow 0$ the case $g \gg 1$ is not realized, and the entire approach employed in this Appendix is inapplicable. The appearance of the logarithm of a dimensioned quantity in Eq. (I.12b) need not cause confusion: we were interested only in the singularity of the function g as $L \rightarrow \infty$, without paying attention to the upper limit of integration. It turns out in an exact calculation that $c_2 = \pi^{-2}$, while $\ln L$ in (I.12b) is replaced by $\ln L/l_p$, where l_p is the mean free path for momentum [here one can calculate it by taking into account only the first term in (I.1).

APPENDIX II. ON THE ROLE OF THE TEMPERATURE-DEPENDENCE OF THE DENSITY OF STATES AT LOW TEMPERATURES

The calculation of the thermodynamic characteristics of a system of many particles at low temperatures is considerably facilitated if one employs the identities that arise under certain conditions from the fact of equivalence of two definitions of the number of particles $N.^{128}$ On the one hand, the latter is given by the thermodynamic relationship¹²⁹

$$N = -\left(\frac{\partial \Omega}{\partial F}\right)_T$$
(II.1)

Here Ω is the thermodynamic potential in the Gibbs grand ensemble, T is the temperature in energy units, and F is the chemical potential (Fermi level). On the other hand we have

$$N = V \int \rho(E) n_{\mathbf{F}}(E) dE.$$
 (II.2)

Here we have $n_{\rm F}(E) = [\exp\beta(E-F) + 1]^{-1}$, $\beta = T^{-1}$, and $\rho(E)$ is the density of states as defined by the equation

$$\rho(E) = \frac{2}{V} \operatorname{Sp} \operatorname{Im} G_{r}(E), \qquad (II.3)$$

Here V and G_r denote respectively the volume of the system and the retarded one-particle Green's function.

689 Sov. Phys. Usp. 26(8), Aug. 1983

The integral over the energy variable E on the righthand side of (II.2) is taken over the entire energy range in which $\rho(E) \neq 0$. The formula (II.2) is exact; it holds when applied to any system of interacting particles. The effects of interaction are described by the mass operator M_r that enters into the expression for the Green's function. Namely, in an arbitrary representation given by the quantum numbers λ , we have¹³⁰

$$G_{\mathbf{r}}^{-1}(\lambda, \lambda'; E) = 2\pi \left(\lambda \mid E - \mathscr{H}_{\mathbf{0}} - M_{\mathbf{r}} \mid \lambda'\right).$$
(II.4)

Here \mathscr{H}_0 is the additive component of the complete Hamiltonian \mathscr{H} of the system.⁴⁵⁾ Naturally, the mass operator generally depends on any intensive variables that characterize the given system of particles, including the concentration of particles and the temperature.

Let us denote by \mathscr{H}' the Hamiltonian of the system, in which we have included the term -FN:

$$\mathscr{H}' = \mathscr{H} - FN. \tag{II.5}$$

We obtain from the condition of normalization of the total density matrix:

$$\Omega = -T \ln \operatorname{Sp} \exp \left(-\beta \mathscr{H}'\right). \tag{II.6}$$

Let us differentiate this equation with respect to T for F = const, and then with respect to F with T = const, using the thermodynamic relationships (II.1) and

$$S = -\left(\frac{\partial\Omega}{\partial T}\right)_{F}, \quad \left(\frac{\partial N}{\partial T}\right)_{F} = \left(\frac{\partial S}{\partial F}\right)_{T}. \tag{II.7}$$

Here S is the entropy. We obtain the following identities:

$$T\left(\frac{\partial\Omega}{\partial T}\right)_{F} = \Omega - \langle \partial\mathcal{C}' \rangle - FT\left(\frac{\partial N}{\partial T}\right)_{F}$$
(II.8)

and

$$\left(\frac{\partial\left(\partial \mathcal{E}'\right)}{\partial F}\right)_{T} + N = -TF\left(\frac{\partial^{2}S}{\partial F^{2}}\right)_{T}$$
(II.9)

or as is equivalent,

$$\left(\frac{\partial\left(\partial\mathcal{C}\right)}{\partial F}\right)_{T} - F\left(\frac{\partial N}{\partial F}\right)_{T} = -TF\left(\frac{\partial^{2}S}{\partial F^{2}}\right)_{T}.$$
(II.9')

Here and below, the angle brackets denote averaging both over the Gibbs grand ensemble and over the random field, if such exists.

We note that, as $T \rightarrow 0$, the right-hand sides of Eqs. (II.9) and (II.9') approach zero faster than linearly.⁴⁶

One can write the expression for $\langle \mathscr{H} \rangle$ in a system of particles having a binary interaction in the form¹³⁰

 $^{^{45}}$ In contrast to Ref. 130, where the variable *E* was measured from the Fermi level, we shall not specify here the reference origin of the energy.

⁴⁶⁾Here we use the Nernst theorem. This requires clarification when studying disordered materials, since the configuration of atoms in a glass and in an amorphous material is metastable, rather than being in full equilibrium. However, we shall be interested here only in electronic processes. Since the relaxation times of the atomic structure are large in comparison with the characteristic electronic times, we are justified in treating the states of the electronic subsystem in a given external field as determined by some particular configuration of atoms.

$$\langle \mathscr{H}' \rangle = \frac{i}{2} \lim_{\substack{j \ t \to -0}} \int_{-\infty}^{+\infty} d\omega \, e^{-i\omega t} \sum_{\lambda'} \\ \times \left\{ EG_{\mathbf{c}} \left(\lambda', \, \lambda'; \, \omega \right) + \sum_{\lambda} \left(\lambda' \mid \mathscr{H}_{\mathbf{0}} - F \mid \lambda \right) G_{\mathbf{c}} \left(\lambda, \, \lambda'; \, \omega \right) \right\}.$$

$$(II.10)$$

Here G_c is the causal Green's function, we have $\omega = E - F$, while the integration over ω extends over all the region where the integrand differs from zero. The apparent lack here of a second Green's function need not cause surprise. It has been eliminated by using the equations of motion for G_c ; actually this function determines the mass operator that appears in Eq. (II.4).

It is convenient to choose the system of eigenfunctions of the operator \mathscr{H}_0 as the basis system; we shall denote its eigenvalues by W_{λ} . Then, upon expressing the function G_c in terms of G_r , we can convert Eq. (II.10) into the more convenient form:¹²⁸

$$(\mathscr{B}') = (\mathscr{B}) - NF = \frac{1}{2} \sum_{\lambda} \int_{-\infty}^{+\infty} dE \operatorname{Im} G_{\Gamma} \langle \lambda, \lambda; E \rangle a_{\Gamma} \langle E \rangle \langle E + W_{\lambda} \rangle - NF.$$
 (II.11)

We must substitute this expression into the identities (II.8)-(II.9'). Since we envision using them to study the temperature-dependence of the number of particles and of the density of states, we note that the chemical potential and the temperature enter into Eq. (II.11) both explicitly—via the Fermi function, and implicitly—via the mass operator.⁴⁷⁾ Consequently we have

$$\left(\frac{\partial \langle \mathscr{B} \varepsilon' \rangle}{\partial F}\right)_{T} = \left(\frac{\partial \langle \mathscr{B} \varepsilon' \rangle}{\partial F}\right)_{T. M_{T}} + \frac{\delta \langle \mathscr{B} \varepsilon' \rangle}{\delta M_{T}} \left(\frac{\partial M_{T}}{\partial F}\right)_{T}.$$
 (II.12)

The second term on the right-hand side of (II.12) amounts to a contracted notation of the rather unwieldy expression

$$\int_{-\infty}^{+\infty} \mathrm{d}E \sum_{\boldsymbol{\lambda},\,\boldsymbol{\lambda}'} \frac{\delta\left(\mathcal{F}'\right)}{\delta M_{T}\left(\boldsymbol{\lambda},\,\boldsymbol{\lambda}';\,E\right)} \left(\frac{\partial M_{T}(\boldsymbol{\lambda},\,\boldsymbol{\lambda}';E)}{\partial F}\right)_{T} \cdot$$

According to (II.9), as $T \rightarrow 0$ the first term on the righthand side of (II.12) is -N. Hence we can rewrite the identity (II.9) in the form

$$\frac{\delta \langle \mathcal{G}\mathcal{E}' \rangle}{\delta M_{\rm r}} \left(\frac{\partial M_{\rm r}}{\partial F} \right)_{\rm T} = -TF \left(\frac{\partial^2 S}{\partial F^2} \right)_{\rm T}.$$
 (II.13)

Since generally we have $(\partial M_r/\partial F)_r \neq 0$, this implies that

$$\frac{\delta \langle \mathscr{B} \varepsilon' \rangle}{\delta M_{r}} \xrightarrow[T \to 0]{} 0. \tag{II.14}$$

In view of (II.7) and (II.8), this implies that

$$\frac{\delta\Omega}{\delta M_r} \underset{T\to 0}{\longrightarrow} 0, \quad \left(\frac{\delta N}{\delta M_r}\right)_F \underset{T\to 0}{\longrightarrow} 0. \tag{II.15}$$

This means that, as $T \rightarrow 0$, in calculating N by Eq. (II.2) we must take into account only the explicit dependence on F contained in the Fermi function.

Now let us turn to the corrections that arise in the function N(F,T) at a low, but finite temperature owing to the interaction between the particles. Let us assume that

$$M_r = M_0 + \Delta M$$
, $\langle \mathcal{F} \rangle = \langle \mathcal{F} \rangle_0 + \Delta \langle \mathcal{F} \rangle$, $N = N_0 + \Delta N$. (II.16)

Here we understand M_0 , N_0 , and $\langle \mathcal{H} \rangle_0$ to be the corre-

sponding quantities at T = 0, while ΔM , ΔN , and $\Delta \langle \mathscr{H} \rangle$ are the temperature corrections. Below we shall assume that $\Delta M \ll M_0$ (for brevity we shall not write the absolute-value symbol here and below, albeit understanding it wherever necessary). However, we note that this assumption may not be fulfilled. Actually, upon using the formula for finite increments, we can write

$$\Delta M = \left(\frac{\partial M_{\rm r}}{\partial T}\right)_{T=T^*} T.$$

Here we have $0 \le T^* \le T$. This expression is actually small in comparison with M_0 if the mass operator has no poles ("normal" system) or if the positions of these poles do not depend appreciably on the temperature. However, if such a dependence exists, then the "correction" ΔM can prove to be arbitrarily large in comparison with M_0 . Here the representation of the mass operator in the form (II.16) loses meaning, and the conclusions that follow below generally prove unwarranted. We shall discuss the problem of when does such a situation arise at the end of this Appendix, restricting the treatment for now to the case $\Delta M \ll M_0$. Then the temperature correction to the number of particles associated with ΔM is

$$\Delta N \approx \left(\frac{\delta N}{\delta M_{\rm r}}\right)_{F, T=0} \Delta M + O\left((\Delta M)^2\right). \tag{II.17}$$

According to Eq. (II.15) the first term here equals zero. To estimate the second term we note that, according to the Nernst theorem, the temperature-derivative of $\langle \mathscr{H} \rangle$ must vanish as $T \rightarrow 0$. In fact, in the variables T and F we have¹²⁹

$$\mathfrak{a} \ (\langle \mathscr{H} \rangle - TS - FN) = -S \, \mathrm{d} T - N \, \mathrm{d} F.$$

That is, we have

$$\left(\frac{\partial \left(\partial \mathcal{E}\right)}{\partial T}\right)_{F} = T \left(\frac{\partial S}{\partial T}\right)_{F} + F \left(\frac{\partial S}{\partial F}\right)_{T} \xrightarrow[\Gamma \to 0]{} 0.$$
(II.18)

As $T \rightarrow 0$, the right-hand side of (II.18) vanishes no more weakly than linearly. Now, upon using Eq. (II.11) and the relationship

$$\left(\frac{\partial n_{\mathbf{F}}}{\partial T}\right)_{\mathbf{T} \to 0} \longrightarrow -\pi^{2}T\delta' (E-F),$$

we can easily convince ourselves that, as T = 0, the derivative $(\partial M_r / \partial T)_F$ also declines no more weakly than linearly. Hence we have

$$\Delta M = O(T^2), \quad \Delta N = O(T^4). \tag{II.19}$$

We note that this is a "minimal" estimate: in various special cases ΔM declines even more strongly with the temperature. However, for our purposes the estimate (II.19) suffices.

Thus, in treating the low-temperature thermodynamics of a degenerate system of Fermi particles, we are correct, apart from quantities of the order of T^4 , in ignoring a possible temperature-dependence of the density of states—if the "single-electron" levels W_{λ} themselves do not depend on the temperature and the poles of the mass operator do not shift appreciably with the temperature.

To illustrate under what conditions such a dependence might arise, it is convenient to treat a simple model

⁴⁷⁾In principle the single-particle levels W_{λ} themselves might depend on the temperature. Naturally the correction associated with this must be taken into account in calculating N.

example-a problem with a very simple Hamiltonian of the Hubbard type describing the short-range attraction and repulsion forces. Let us have a system of electrons moving in some atomic matrix, not necessarily ordered. Let us number the nodes of this matrix with the subscript g and denote by σ the spin quantum number of the electron. Let us choose as the basis system one of orthogonalized wave functions describing the electrons localized at the different nodes (in an ideal crystal these would be Wannier functions). We shall denote the corresponding single-electron energies by W; they can depend both on the number of the node and on σ ; one value of g can, generally speaking, correspond to different values of W, corresponding to the ground and excited single-electron states. Then the role of λ is played by the set of numbers $\{g, W, \sigma\}$; we have

$$\mathscr{H}_0 = \sum_{\lambda} W_{\lambda} a_{\lambda}^{\bullet} a_{\lambda},$$

and the Hamiltonian in which we are interested has the form $% \left({{{\mathbf{x}}_{i}}} \right)$

$$\mathscr{H} = \mathscr{H}_{0} + \frac{1}{2} \sum_{\substack{\lambda_{1}, \lambda_{2} \\ \lambda_{1}, \lambda_{2}}} (\lambda_{1}, \lambda_{2} + V + \lambda_{2}', \lambda_{1}') a_{\lambda_{1}}^{\bullet} a_{\lambda_{2}}^{*} a_{\lambda_{3}}^{*} a_{\lambda_{1}'}^{*}.$$
(II.20)

Here the symbol \sum_{λ} can denote either the sum or the integral, and we have

$$(\lambda_1, \lambda_2 | V | \lambda'_2, \lambda'_1) = \frac{1}{2} V_0 \delta_{g_1 g_1, \delta'_1 g'_2} \delta_{g'_2 g'_2} (\delta_{\sigma_1 \sigma'_1} \delta_{\sigma_1 \sigma'_2} - \delta_{\sigma_1 \sigma'_2} \delta_{\sigma_2 \sigma'_1} \delta_{\sigma_1, -\sigma_{11}}$$
(II.21)

Here $\delta_{\ell_1 \ell_2}$, etc., are the Kronecker symbols, while V_0 is a positive or negative constant.

Let us write in the usual way the equations of motion for the Green's functions⁴⁸⁾

 $G(\lambda, \ \lambda; \ E) = \langle \langle a_{\lambda} \ | \ a_{\lambda}^{*} \rangle \rangle_{E}, \quad G_{2}(\alpha, \ \alpha; \ \lambda, \ \lambda; \ E) = \langle \langle a_{\alpha}^{*} a_{\alpha} a_{\lambda} \ | \ a_{\lambda}^{*} \rangle \rangle_{E},$

where $\alpha = \{g, W, -\sigma\}$. We obtain

$$(E-W) G(\lambda, \lambda; E) = -\frac{1}{2\pi} + V_0 G_2(\alpha, \alpha; \lambda, \lambda; E), \qquad (II.22)$$

$$(E - W - V_0) G_2(\alpha, \alpha; \lambda, \lambda; E) = -\frac{n_\alpha}{2\pi}, \qquad (II.23)$$

where

$$n_{\alpha} = \langle a_{\alpha}^{*} a_{\alpha} \rangle. \tag{II.24}$$

We note that a third Green's function does not appear in Eq. (II.23). That is, instead of an infinite chain of equations, we obtain a simple finite system. This involves the choice of the highly radical approximation (II.21).

For simplicity, let the single-electron energies W_{λ} be independent of the spin orientation. Then we have

$$n_{\alpha} = n_{\alpha} = n \quad (W). \tag{II.25}$$

Upon solving the system of equations (II.22) and (II.23) and taking into account the equation

$$n(W) = 2 \int dE \operatorname{Im} G_{\mathbf{r}}(\lambda, \lambda; E) n_{\mathbf{F}}(E), \qquad (II.26)$$

we obtain

$$G(\lambda, \lambda; E) = -\frac{1}{2\pi} \frac{E - W - (1 - n) V_0}{(E - W - V_0) (E - W)},$$
 (II.27)

$$G_{2}(\alpha, \alpha; \lambda, \lambda; E) = -\frac{n}{2\pi} \cdot \frac{1}{E - W - V_{2}}, \qquad (\text{II.28})$$

$$n(W) = \frac{n_F(W)}{1 + n_F(W) - n_F(W + V_0)}.$$
 (II.29)

Let us denote by $\rho_0(W)$ the density of single-electron states of energy W. Upon using the formulas (II.3) and (II.27), we find the density of states of the many-particle system under discussion:

 $\rho(E) = \rho_0(E) [1 - n(E)] + \rho_0(E - V_0) n(E - V_0).$ (II.30)

The right-hand side of (II.30) depends on the temperature. However, a simple investigation shows that, if the Hamiltonian of (II.20) describes attractive forces, then at low enough temperatures this affects only the magnitude of the right-hand side of (II.2), but not its temperature-dependence. On the other hand, for repulsive forces the situation can change. This fact has an entirely obvious origin. Indeed, when $T \ll |V_0|$ the repulsive forces rule out the possibility of two electrons occupying the same node. This is the same situation as occurs in the problem of the distribution of electrons over impurity atoms in semiconductors. As is known,24 in the case of "simple" centers possessing only one level each and capable of capturing only one electron, the mean occupancy number of the impurity level f_t has the form

$$f_t = \left[\frac{g_0}{g_1} \exp\left(\frac{E_t - F}{T}\right) + 1\right]^{-1} \cdot$$
(II.31)

Here g_0 and g_1 are the statistical weights of the states of an empty center and a center occupied by an electron, while E_t is the energy of an electron in an impurity level. Evidently, we can rewrite the expression (II.31) in the form of a Fermi function by introducing the "effective energy" $E_t^* = E_t + T \ln g_0/g_1$. The corresponding density of states is

 $\rho(E) = N_{\mathbf{t}}\delta(E - E_{\mathbf{t}}^{\bullet}).$

Here N_{\star} is the concentration of centers.

We see that essentially this is the very same case of temperature-dependence of energy levels that we have discussed above (see footnote 47).⁴⁹⁾

Analogous results also prove to be valid in the problem with anomalous pairings, which are essential, e.g., in the problem of the Coulomb gap.¹¹³ The cited study employed the concept of an effective attraction between electrons and holes ($V_0 > 0$). Here, in the approximation (II.21), the fundamental equations of Ref. 113 take on the form

$$\begin{array}{l} (E - W) \ G \ (\mathbf{R}, \ \mathbf{R}; \ E) = - \ V_0 n' G' \ (\mathbf{R}, \ \mathbf{R}; \ E) = - \ 1/2 \pi, \\ (E + W) \ G' \ (\mathbf{R}, \ \mathbf{R}; \ E) = - \ V_0 n' G \ (\mathbf{R}, \ \mathbf{R}; \ E) = 0 \ . \end{array}$$
 (II.32)

Here we have

$$n'(W) = \langle a_{\lambda} a_{\alpha} \rangle = \langle a_{\alpha}^{*} a_{\lambda}^{*} \rangle, \ G'(\mathbf{R}, \, \mathbf{R}; E) = \langle a_{\alpha}^{*} | a_{\lambda}^{*} \rangle. \tag{II.33}$$

This system of equations (apart from the notation) is well known in the theory of superconductivity. Evidently we have

$$\operatorname{Im} G(E) = \frac{E+W}{4(W^2 + n'^2V_0^2)^{1/2}} \left[\delta\left(E - \sqrt{W^2 + n'^2V_0^2}\right) - \delta\left(E + \sqrt{W^2 + n'^2V_0^2}\right) \right].$$
(II.34)

⁴⁸⁾I am grateful to I. P. Zvyagin, who called my attention to this situation.

⁴⁸ The symbols and the normalization of the Green's functions are the same as in the book of Ref. 130.

For determining n', we obtain the standard equation

$$n' = \frac{V_0 n'}{2(W^3 + n'^3 V_0^3)^{1/2}} \tanh\left(\frac{\beta \sqrt{W^3 + n'^3 V_0^3}}{2}\right).$$
(II.35)

As we can easily convince ourselves, when $\beta V_0 \ll 1$, there is no nontrivial solution for *n'*. On the other hand, at low temperatures at which $\beta V_0 \gg 1$, we obtain

$$n' = \frac{1}{2} \left(1 + 4\eta - 4W^2 V_{-0}^2 \right)^{1/2}$$
(II.36)

Here we have

$$\eta = -2 \exp\left(-\frac{V_0}{2T}\right). \tag{II.37}$$

A solution of (II.36) exists as long as

$$\frac{W}{V_0} < \frac{1+2\eta}{2}.$$
 (II.38)

As we should have expected, we see that the temperature-dependence of n' proves to be insignificant as soon as $\beta V_0 \gg 1$. For the mass operator defined as before by Eq. (II.4), we now obtain

$$M = \frac{V_0^* (1+4\eta) - W^2}{E+W}.$$
 (II.39)

We see that its pole, E = -W, does not depend at all on the temperature.

Up to now we have considered the quantity V_0 as being fixed. In averaging over the configurations of the random field, we run through all the values of V_0 . However, under the conditions being studied, a Coulomb gap arises only from the configurations in which $V_0 > 0$ and $\beta V_0 > 1$.

- ¹B. T. Kolomiets and N. A. Goryunova, Zh. Tekh. Fiz. 25, 2069 (1955).
- ²B. T. Kolomiets, Phys. Status Solidi 7, 359 (1964).
- ³M. V. Sadovskil, Usp. Fiz. Nauk **133**, 223 (1981) [Sov. Phys. Usp. **24**, 96 (1981)].
- ⁴G. A. Petrakovskil, Usp. Fiz. Nauk 134, 305 (1981) [Sov. Phys. Usp. 24, 511 (1981)].
- ⁵N. F. Mott, J. Phys. C13, 5433 (1980).
- ⁶R. A. Abram, G. J. Rees, and B. L. H. Wilson, Adv. Phys. 7, 799 (1978).
- ⁷M. H. Brodsky and M. Cardona, J. Non-Cryst. Solids 31, 81 (1978).
- ⁸G. A. N. Connell and G. Lucovsky, *ibid.*, p. 123.
- ⁹T. D. Moustakas, J. Electron. Res. 8, 391 (1979).
- ¹⁰M. Fritzsche, C. C. Tsai, and P. Persans, Solid State Techn., p. 55 (January 1978).
- ¹¹Y. Hamakawa, Surf. Sci. 86, 444 (1979).
- ¹²P. G. Le Comber, Sci. Progr. Oxf. 66, 105 (1979).
- ¹³H. Böttger, R. Keiper, and R. Schuchardt, Wissenschaft und Fortschritt **30**, 209 (1980).
- ¹⁴P. S. Peercy, Nucl. Instrum. Methods 182-183, 337 (1981).
 ¹⁵Fundamental Physics of Amorphous Semiconductors: Proc. of the Kyoto Summer Institute, Kyoto, Japan, Sep. 8-11, 1980, ed. F. Yonezawa, Springer-Verlag, Berlin, Heidelberg, New York, 1981. (This book contains the texts of lectures by G. Fritzsche, D. Adler, I. Solomon, W. E. Spear, P. G. Le Comber, M. H. Brodsky, W. Paul, G. Lucovsky, K. Tanaka, F. Yonezawa and M. H. Cohen, V. L. Bonch-Bruevich, and D. Weaire on various aspects of the physics of amorphous and vitreous materials.)
- ¹⁶I. M. Lifshits, L. A. Pastur, and S. A. Gredeskul, Fiz. Nizk. Temp. 2, 1093 (1976) [Sov. J. Low Temp. Phys. 2, 533 (1976)].
- ¹⁷S. A. Gredeskul and L. A. Pastur, *ibid.* 1, 277 (1975) [Sov.

- J. Low Temp. Phys. 1, 135 (1975)].
- ¹⁸N. F. Mott, Contemp. Phys. 18, 225 (1977).
- ¹⁹B. T. Kolomiets and V. M. Lyubin, Mater. Res. Bull. 13, 1343 (1978).
- ²⁰V. M. Lyubin, in: Uspekhi nauchnol fotografii (Advances of Scientific Photography), Nauka, M., 1980, p. 66.
- ²¹B. T. Kolomiets, in: Amorphous Semiconductors-78, Pardubice, Sep. 1978, Institute of Solid State Physics, Czechoslovak Academy of Sciences, Prague; University
- of Chemical Technology, Pardubice, 1978, p. 3.

- ²³V. L. Bonch-Bruevich, I. P. Zvyagin, R. Kaiper, A. G. Mironov, R. Énderlain, and B. Ésser, Élektronnaya teoriya neuporyadochennykh poluprovodnikov (Electronic Theory of Disordered Semiconductors), Nauka, M., 1981.
- ²⁴V. L. Bonch-Bruevich and S. G. Kalashnikov, Fizika poluprovodnikov (Physics of Semiconductors), Nauka, M., 1977.
- ²⁵N. F. Mott and E. A. Davis, Electronic Processes in Non-Crystalline Materials, Clarendon Press, Oxford, 1971 (Russ. Transl., Mir, M., 1974).
- ²⁶P. W. Anderson, Phys. Rev. **102**, 1492 (1958).
- ²⁷N. F. Mott, Philos, Mag. 65, 1259 (1969).
- ²⁸V. L. Berezinskii and L. P. Gor'kov, Zh. Eksp. Teor. Fiz. 77, 2498 (1979) [Sov. Phys. JETP 50, 1209 (1979)].
- ²⁹J. T. Edwards and D. J. Thouless, J. Phys. C 5, 807 (1972).
- ³⁰D. Weaire and B. Kramer, J. Non-Cryst. Solids 35-36, 9 (1980).
- ³¹ P. Dean, Proc. Phys. Soc. **73**, 413 (1959); Proc. R. Soc. London Ser. A **254**, 507 (1960).
- ³²F. Yonezawa, J. Non-Cryst. Solids **35-36**, 29 (1980).
- ³³N. F. Mott and W. D. Twose, Adv. Phys. **10**, 107 (1961) [Russ. Transl., Usp. Fiz. Nauk **79**, 691 (1963)].
- ³⁴R. E. Borland, Proc. R. Soc. London Ser. A 274, 926 (1963); R. E. Borland and N. E. Bird, Proc. Phys. Soc. London 83, 23 (1964).
- ³⁵V. L. Berezinskii, Zh. Eksp. Teor. Fiz. 65, 1251 (1973) [Sov. Phys. JETP 38, 620 (1974)].
- ³⁶N. Ya. Gol'dsheldt, S. A. Molchanov, and L. A. Pastur, Funkts, analiz i ego pril. 11, 1 (1977).
- ³⁷N. F. Mott, Metal-Insulator Transitions, Taylor & Francis, London, 1974 (Russ. Transl., Nauka, M., 1979).
- ³⁸V. L. Bonch-Bruevich and V. B. Glasko, Fiz. Tverd. Tela (Leningrad) **3**, 36 (1961) [Sov. Phys. Solid State **3**, 26 (1961)]; V. L. Bonch-Bruevich, *ibid.*, p. 47 [Sov. Phys. Solid State **3**, 34 (1961)].
- ³⁹A. A. Gogolin, V. I. Mel'nikov, and E. I. Rashba, Zh. Eksp. Teor. Fiz. **69**, 327 (1975) [Sov. Phys. JETP **42**, 168 (1975)].
- ⁴⁰A. A. Gogolin and V. I. Mel'nikov, Phys. Status Solidi B 88, 377 (1978).
- ⁴¹A. A. Gogolin, V. I. Mel'nikov, and E. I. Rashba, Zh. Eksp. Teor. Fiz. **72**, 629 (1977) [Sov. Phys. JETP **45**, 330 (1977)].
- ⁴²A. A. Gogolin, *ibid.* 71, 1912 (1976) [Sov. Phys. JETP 44, 1003 (1976)].
- ⁴³A. A. Gogolin, Preprint in Theoretical Physics HU-TFT, Research Institute for Theoretical Physics, University of Helsinki, 1981, p. 1.
- ⁴⁴K. Wilson and J. Kogut, The Renormalization Group and the ε-Expansion, (Russ. Transl., Mir, M., 1975).
- ⁴⁵R. Balescu, Equilibrium and Non-Equilibrium Statistical Mechanics, Wiley, New York, 1975, Vol. 1, Chap. 10 (Russ. Transl., Mir, M., 1978).
- ⁴⁶F. Wegner, Z. Phys. B 25, 327 (1976).
- ⁴⁷F. Wegner, *ibid.* 35, 207 (1979).
- ⁴⁸R. Oppermann and F. Wegner, *ibid.* 34, 327 (1979).
- ⁴⁹M. V. Sadovskii, Zh. Eksp. Teor. Phys. 70, 1936 (1976) [Sov. Phys. JETP 43, 1008 (1976)].
- ⁵⁰E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).

²²H. Böttger, *ibid.*, p. 13.

- ⁵¹E. Abrahams and T. V. Ramakrishnan, J. Non-Cryst. Solids 35-36, 15 (1980).
- ⁵²K. B. Efetov, A. I. Larkin, and D. E. Khmel'nitskii, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys. JETP **52**, 568 (1980)].
- ⁵³P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, Phys. Rev. B 22, 3519 (1980).
- ⁵⁴P. W. Anderson, *ibid.* 23, 4829 (1981).
- ⁵⁵V. L. Bonch-Bruevich and A. G. Mironov, J. Phys. (Paris) 42, Suppl. C-4, 33 (1981).
- ⁵⁶L. P. Gor'kov, A. I. Larkin, and D. E. Khmel'nitskii, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 248 (1979) (JETP Lett. **30**, 228 (1979)].
- ⁵⁷D. E. Khmel'nitskiĭ, *ibid.* **32**, 248 (1980) [JETP Lett. **32**, 229 (1980)].
- ⁵⁸N. F. Mott and M. Kaveh, Preprint, 1981.
- ⁵⁹J. S. Langer and T. Neal, Phys. Rev. Lett. 16, 984 (1966).
- ⁶⁰T. Neal, Phys. Rev. 169, 508 (1968).
- ⁶¹D. J. Thouless, Phys. Rev. Lett. 39, 1168 (1977).
- ⁶²D. J. Dolan and D. D. Osheroff, *ibid.* 43, 721 (1979).
- ⁶³D. J. Bishop, C. B. Tsui, and R. C. Dynes, *ibid.* 44, 1153 (1980).
- ⁶⁴M. Kaveh and N. F. Mott, J. Phys. C 14, L177 (1981).
- ⁶⁵P. A. Lee, Phys. Rev. Lett. 42, 21 (1979); J. Non-Cryst. Solids 35-36, 21 (1980).
- ⁶⁶C. H. Hodges, J. Phys. (Paris) 42, Suppl. C-4, 67 (1981).
- ⁶⁷G. Polya, Math. Ann. 84, 149 (1921).
- ⁶⁸W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, New York, 1957 (Russ. Transl., IL, M., 1962).
- ⁶⁹V. L. Bonch-Bruevich, Pis'ma Zh. Eksp. Teor. Fiz. 32, 222 (1980) [JETP Lett. 32, 204 (1980)].
- ⁷⁰N. F. Mott, Commun. Phys. 1, 203 (1976).
- ⁷¹N. F. Mott, Philos. Mag. 44, 265 (1981).
- ⁷²K. F. Freed, Phys. Rev. B 5, 4802 (1972).
- ⁷³S. F. Edwards, J. Non-Cryst. Solids 32, 113 (1979).
- ⁷⁴N. F. Mott, M. Pepper, S. Pollitt, R. H. Wallis, and
- C. Adkins, Proc. R. Soc. London Ser. A **315**, 169 (1975). ⁷⁵M. Kaveh and N. F. Mott, J. Phys. C **14**, L177 (1981).
- (Same as Ref. 64.)
- ⁷⁶E. J. Von Neumann and E. P. Wigner, Phys. Z. **30**, 465 (1929).
- ⁷⁷A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, Rasseyanie, reaktsii i raspady v nerelyativistskoi kvantovoi mekhanike (Scattering, Reactions, and Decays in Nonrelativistic Quantum Mechanics), 2nd edn., Nauka, M., 1971, Appendix A. [Engl. Transl. of 1st edn. Israel Program for Scientific Translations, Jerusalem; Wiley, New York, 1969.]
- ⁷⁸R. G. Newton, Scattering Theory of Waves and Particles, McGraw-Hill, New York, 1966 (Russ. Transl., Mir, M., 1971).
- $^{79}\mathrm{B.}$ J. Last and D. Thouless, J. Phys. C 7, 699 (1974).
- ⁸⁰V. L. Bonch-Bruevich, in: Proc. 6th International Conference on Amorphous and Liquid Semiconductors, Nauka, L., 1976, p. 76; Dokl. Akad. Nauk SSSR 226, 298 (1976) [Sov. Phys. Dokl. 21, 28 (1976)].
- ⁸¹V. L. Bonch-Bruevich, Preprint of the Physics Faculty of the Moscow State University No. 3, Moscow, 1982; Phys. Status Solidi B 113, 493 (1982).
- ⁸²V. L. Bonch-Bruevich, Fiz. Tekh. Poluprovodn. 2, 363 (1968) [Sov. Phys. Semicond. 2, 299 (1968)].
- ⁸³V. L. Bonch-Bruevich, in: Theory of Condensed Matter: Lectures presented at an Intern. Course, Trieste, 3 October 16 December 1967, Intern. Atomic Energy Agency, Vienna, 1968, p. 989.
- ⁸⁴P. W. Anderson, Phys. Rev. Lett. 109, 1492 (1975).
- ⁸⁵N. F. Mott, E. A. Davis, and R. A. Street, Philos. Mag. 3, 961 (1975).
- ⁸⁶R. A. Street and N. F. Mott, Phys. Rev. Lett. 35, 1293 (1975).
- ⁸⁷R. A. Street, see Ref. 80, p. 232.

- ⁸⁸B. L. Al'tshuler and A. G. Aronov, Pis'ma Zh. Eksp. Teor. Fiz. 27, 700 (1978) [JETP Lett. 27, 662 (1978)].
- ⁸⁹B. L. Altshuller and A. G. Aronov, Solid State Commun. 30, 115 (1979).
- ⁹⁰B. L. Al'tshuler and A. G. Aronov, Zh. Eksp. Teor. Fiz. 77, 2028 (1979) [Sov. Phys. JETP 50, 968 (1979)].
- ³¹B. L. Altshuler, A. G. Aronov, and P. A. Lee, Phys. Rev. Lett. 44, 1288 (1980).
- ⁹²M. Kaveh and N. F. Mott, J. Phys. (Paris) C 14, L183 (1981).
- ⁹³B. L. Altshuler, D. Khmelnitzkii, A. I. Larkin, and P. A. Lee, Phys. Rev. B 22, 5142 (1980).
- ⁹⁴S. Hikami, A. I. Larkin, and Y. Nagaoka, Prog. Theor. Phys. **63**, 707 (1980).
- ⁹⁵ M. Kaveh, M. J. Uren, R. A. Davies, and M. Pepper, J. Phys. (Paris) C 14, 413 (1981).
- ⁹⁶Y. Kawaguchi, Y. Kitahara, and S. Kawaji, Solid State Commun. 14, 413 (1978).
- ⁹⁷ M. Pepper, J. Phys. (Paris) 42, Suppl. C-4, 17 (1981).
- 98H. Fukuyama, J. Phys. Soc. Jpn. 48, 2169 (1980).
- ⁹⁹M. J. Uren, R. A. Davies, and M. Pepper, J. Phys. (Paris) C **13**, L985 (1980).
- ¹⁰⁰M. Pepper, Philos. Mag. Ser. B 37, 83 (1978).
- ¹⁰¹A. Houghton, A. Jevcki, R. D. Kenway, and A. M. A. Pruisken, Phys. Rev. Lett. 45, 394 (1980).
- ¹⁰²R. Haydock, Philos. Mag. 43, 203 (1981).
- ¹⁰³J. L. Pichard and G. Sarma, J. Phys. (Paris), 42, Suppl. C-4, 37 (1981).
- ¹⁰⁴J. L. Pichard and G. Sarma, Preprint, 1981.
- ¹⁰⁵M. Pollak, Discuss. Faraday Soc. 50, 13 (1970).
- ¹⁰⁶G. Srinivasan, Phys. Rev. B 4, 2581 (1971).
- ¹⁰⁷M. L. Knotek and M. Pollak, J. Non-Cryst. Solids 8-10, 505 (1972); Phys. Rev. B 9, 664 (1974); Philos. Mag. 35, 117 (1977).
- ¹⁰⁸A. L. Efros and B. I. Shklovskii, J. Phys. (Paris) C 8, L49 (1975).
- ¹⁰⁹T. Kurosawa and H. Sugimoto, Prog. Theor. Phys. Suppl. 8, 234 (1975).
- ¹¹⁰N. F. Mott, Philos. Mag. 34, 643 (1976); J. Phys. (Paris) C 8, L234 (1975).
- ¹¹¹A. L. Éfros, *ibid*. 9, 2021 (1976).
- ¹¹²A. L. Éfros, see Ref. 80, p. 126.
- ¹¹³V. L. Bonch-Bruevich, Vestn. Mosk. Univ. Fiz. Astron. 18, No. 3, 80 (1977).
- ¹¹⁴V. L. Bonch-Bruevich, *ibid.*, No. 4, 51.
- ¹¹⁵V. L. Bonch-Bruevich, Phys. Status Solidi B 84, 99 (1977).
- ¹¹⁶V. L. Bonch-Bruevich, in: Proc. 7th Intern. Conference on Amorphous and Liquid Semiconductors, ed. W. E. Spear, Edinburgh, 1977, p. 239.
- ¹¹⁷B. L. Gel'mont and A. A. Éfros, Pis'ma Zh. Eksp. Teor. Fiz. 25, 77 (1977) [JETP Lett. 25, 67 (1977)].
- ¹¹⁸S. Kirkpatrick and C. M. Varma, Solid State Commun. 21, 183 (1977); J. Non-Cryst. Solids **32**, 141 (1979).
- ¹²⁰S. D. Baranovskii, A. L. Efros, S. L. Gelmont, and B. I.
 Shklovskii, J. Phys. (Paris) C 12, 1023 (1979); Solid State Commun. 27, 1 (1978).
- ¹²¹A. L. Efros, Nguyen Von Lien, and B. I. Shklovskii, J. Phys. (Paris) C 12, 1869 (1979); Solid State Commun. 32, 851 (1979).
- ¹²²N. F. Mott, J. Non-Cryst. Sol. 35-36, 79 (1979).
- ¹²³M. Pollak, *ibid.*, p. 83.
- ¹²⁴N. F. Mott, Philos. Mag. 34, 663 (1976).
- ¹²⁵S. D. BaranovskiĬ, B. I. ShklovskiĬ and A. L. Éfros, Zh. Eksp. Teor. Fiz. **78**, 395 (1980) [Sov. Phys. JETP **51**, 199 (1980)].
- ¹²⁶B. I. Shklovskil and A. L. Éfros, Fiz. Tekh. Poluprovodn.
 14, 825 (1980) [Sov. Phys. Semicond. 14, 487 (1980)].
- ¹²⁷M. Pollak, Preprint, 1981.
- ¹²⁸V. L. Bonch-Bruevich, Acta Universitatis Wratislaviensis, No. 80, 29 (1968).
- ¹²⁹L. D. Landau and E. M. Lifshits, Statisticheskaya fizika

(Statistical Physics), Vol. 1, 3rd edn., with supplement by E. M. Lifshits and L. P. Pitaevskii, Nauka, M., 1976.

[Engl. Transl. Pergamon Press, Oxford (1980)].

- ¹³⁰V. L. Bonch-Bruevich and S. V. Tyablikov, Metod funktsil Grina v statisticheskoi mekhanike (The Green's Function Method in Statistical Mechanics), Fizmatgiz, M., 1961. [Engl. Transl., North Holland, Amsterdam (1962)].
- ¹³¹E. M. Hamilton, Philos. Mag. 26, 1043 (1972).
- ¹³²I. S. Shlimak and E. I. Nikulin, Pis'ma Zh. Eksp. Teor. Fiz. 15, 30 (1972) [JETP Lett. 15, 20 (1972)].
- ¹³³S. Kobayashi, Y. Monden, and W. Sasaki, Solid State Commun. 30, 161 (1979); 33, 793 (1980).
- ¹³⁴Y. Ootuka, P. Komori, Y. Monden, S. Kobayashi, and W. Sasaki, Solid State Commun. 33, 793 (1980); T. G. Castner and G. A. Thomas, Commun. Solid State Phys. 9, 235 (1980).
- ¹³⁵V. L. Bonch-Bruevich, A. G. Mironov, and I. P. Zvyagin, Riv. Nuovo Cimento 3, 321 (1973).
- ¹³⁶S. D. Baranovskii and A. L. Éfros, Fiz. Tekh. Poluprovodn. 14, 2233 (1980) [Sov. Phys. Semicond. 14, 1323 (1980)].
- ¹³⁷F. Urbach, Phys. Rev. 92, 1324 (1953).
- ¹³⁸M. V. Kurik, Phys. Status Solidi A 8, 9 (1971).
- ¹³⁹J. J. Pankove, Phys. Rev. A 140, 2059 (1965).
- ¹⁴⁰D. Redfield and M. A. Afromowitz, Appl. Phys. Lett. 11, 138 (1967).
- ¹⁴¹M. A. Aframovich and D. Redfild, in: Proc. 9th Intern. Conference on Physics of Semiconductors, Nauka, L., 1968, p. 103.
- ¹⁴²H. R. Wittman, Phys. Status Solidi 35, 863 (1969).
- ¹⁴³E. A. Davis, J. Non-Cryst. Solids 4, 107 (1970).
- ¹⁴⁴M. L. Theye, Opt. Commun. 2, 329 (1970); Mater. Res. Bull. 6, 103 (1971).
- ¹⁴⁵D. T. Pierce and W. E. Spicer, Phys. Rev. B 5, 3017 (1972).
- 146C. C. Tsai, ibid. 19, 2041 (1979).
- ¹⁴⁷R. S. Crandall, Phys. Rev. Lett. 44, 749 (1979); J. Non-Cryst. Solids 35-36, 381 (1980).
- ¹⁴⁸G. Moddel, D. A. Anderson, and W. Paul, Phys. Rev. B 22, 1918 (1980).
- ¹⁴⁹T. D. Moustakas, Solid State Commun. 35, 745 (1980).
- ¹⁵⁰B. Abeles, C. R. Wronski, T. Tiedje, and G. D. Cody, *ibid*. 36, 537 (1980).
- ¹⁵¹H. Tokumoto, M. Tokumoto, and T. Ishiguro, J. Phys. Soc. Jpn. 50, 602 (1981).
- ¹⁵²P. J. Zanzucchi, C. R. Wronski, and D. E. Carlson, J. Appl. Phys. 48, 5227 (1977).
- ¹⁵³C. C. Tsai and H. Fritzsche, Solar Energy Mater. 1, 29 (1979).
- ¹⁵⁴a) S. Yamasaki, T. Hata, T. Yoshida, H. Oheda, A. Matsuda, H. Okushi, and K. Tanaka, J. Phys. (Paris) 42, Suppl. C-4. 297 (1981); b) S. Yamasaki, K. Nakagawa, H. Yamamoto, A. Matsuda, H. Okushi, and K. Tanaka, Preprint, 1981.
- ¹⁵⁵M. Olivier and P. Bouchut, J. Phys. (Paris) 42, Suppl. C-4. 305 (1981).
- ¹⁵⁶G. D. Gody, B. Abeles, C. R. Wronski, B. Brooks, and W. A. Lanford, J. Non-Cryst. Solids 35-36, 463 (1980).
- ¹⁵⁷J. Tauc, A. Menthe, and D. L. Wood, Phys. Rev. Lett. 25, 749 (1970).
- ¹⁵⁸R. A. Street, T. M. Searle, J. G. Austin, and R. S. Sussman, J. Phys. (Paris) C 7, 1582 (1974).
- ¹⁵⁹A. M. Andriesh, V. S. Gerasimenko, R. Yu. Lyalikova, M. S. Iovu, V. S. Minaev, S. I. Radautsan, D. I. Tsiulyanu, and S. D. Shutov, see Ref. 80, p. 246.
- ¹⁶⁰R. S. Sussmann, J. G. Austin, and T. M. Searle, *ibid.*, p. 256. ¹⁶¹T. F. Mazets and S. I. Pavlov, *ibid.*, p. 256.
- ¹⁶¹T. F. Mazets and S. I. Pavlov, *ibid.*, p. 260.
- ¹⁶²W. Henrion and H. Friedrich, *ibid.*, p. 269.
- ¹⁶³Yu. R. Zakis and A. V. Moskal'onov, *ibid.*, p. 327.
- ¹⁶⁴B. T. Kolomiets, T. F. Mazets, and Sh. I. Efendiev, Fiz. Tekh. Poluprovodn. 4, 1103 (1970) [Sov. Phys. Semicond. 4, 934 (1970)].
- ¹⁶⁵B. T. Kolomiets, T. F. Mazets, Sh. Sh. Sarsembinov, and
- 694 Sov. Phys. Usp. 26(8), Aug. 1983

- Sh. K. Efendiev, J. Non-Cryst. Solids 8-10, 1010 (1972).
- ¹⁶⁵B. T. Kolomiets, T. F. Mazets, S. I. Pavlov, and Sh. Sh. Sarsembinov, ibid. 9, 710 (1975) [Sov. Phys. Semicond. 9, 465 (1975)].
- ¹⁶⁶B. T. Kolomiets, T. F. Mazets, Sh. Sh. Sarsembinov, and Sh. K. Efendiev, J. Non-Cryst. Solids 8-10, 1010 (1972).
- ¹⁶⁷Sh. Sh. Sarsembinov, M. G. Esenbaev, and G. K. Kozhakhmetova, Izv. Akad. Nauk Kaz. SSR Ser. Fiz. Mat., No. 4, 50 (1976).
- ¹⁶⁸Sh. Sh. Sarsembinov, M. I. Korsunskii, and M. G. Esenbaev, ibid., No. 2, 48 (1977).
- ¹⁶⁹J. T. Edmond and V. Patel, see Ref. 116, p. 551.
- ¹⁷⁰G. R. Moridi and C. A. Hogarth, *ibid.*, p. 688.
- ¹⁷¹N. V. Baramidze and N. I. Kurdiani, Fiz. Tekh. Poluprovodn. 5, 631 (1971) [Sov. Phys. Semicond. 5, 555 (1971)].
- ¹⁷²N. V. Baramidze and N. I. Kurdiani, in: Radiatsionnaya fizika nemetallicheskikh kristallov (Radiation Physics of Nonmetallic Crystals), Naukova Dumka, Kiev, 1971, Vol. 3, Part 2, p. 32.
- ¹⁷³R. W. Griffith, see Ref. 116, p. 457.
- ¹⁷⁴A. S. Davydov, Phys. Status Solidi 27, 51 (1968).
- ¹⁷⁵A. K. Lubchenko and I. M. Zushman, *ibid.* 32, 703 (1969). ¹⁷⁶S. I. Pekar, Issledovaniya po élektronnoi teorii ionnykh
- kristallov (Studies in the Electronic Theory of Ionic Crystals), Gostekhizdat, M., 1952.
- ¹⁷⁷B. I. Shklovskil and A. L. Éfros, Élektronnye svolstva legirovannykh poluprovodnikov (Electronic Properties of Doped Semiconductors), Nauka, M., 1979.
- ¹⁷⁸V. L. Bonch-Bruevich and P. S. Sul'zhenko, Vestn. Mosk. Univ. Fiz. Astron. 23, 37 (1982).
- ¹⁷⁹P. S. Sul'zhenko, *ibid.*, p. 82. Dep. VINITI No. 4858, Oct. 22, 1981.
- ¹⁸⁰N. V. Baramidze, V. L. Bonch-Bruevich, M. P. Giorgadze, and N. I. Kurdiani, Phys. Status Solidi B 110, 33 (1982).
- ¹⁸¹N. V. Baramidze, V. L. Bonch-Bruevich, and N. I. Kurdiani, Soobshch, Akad. Nauk Gruz, SSR, 1983.
- ¹⁸²V. D. Iskra, Zh. Eksp. Teor. Fiz. 82, 646 (1982) [Sov. Phys. JETP 55, 385 (1982)]; Izv. Vyssh. Uchebn. Zaved. Fiz. 25, No. 11, 32 (1982).
- ¹⁸³B. Esser and P. Kleinert, Phys. Status Solidi B 72, 535 (1975).
- ¹⁸⁴S. Abe, J. Phys. Soc. Jpn. 49, Suppl. A, 1179 (1980).
- ¹⁸⁵S. Abe and Y. Toyazawa, Techn. Rept. I.C.S.P. Ser. A, No. 1124 (1981).
- ¹⁸⁶B. Esser, Phys. Status Solidi B 55, 503 (1973); 58, 149 (1973).
- ¹⁸⁷B. Esser and P. Kleinert, in: Proc. of the Conference on Amorphous Semiconductors '76, Budapest, 1976, p. 83; B. Esser and P. Kleinert, see Ref. 116, p. 244.
- ¹⁸⁸B. Esser, R. Keiper, and D. Sprenger, in: Proc. of the Conference on Amorphous Semiconductors '78, Prague, 1978, p. 517.
- ¹⁸⁹V. D. Iskra, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 12, 37 (1979); No. 3, 39 (1980).
- ¹⁹⁰V. D. Iskra, Phys. Status Solidi B **96**, 843 (1979); **106**, 627 (1981).
- ¹⁹¹H. Van Cong, J. Phys. Chem. Solids 36, 1237 (1975).
- ¹⁹²Yu. E. Perlin, Usp. Fiz. Nauk 80, 553 (1963) [Sov. Phys. Usp. 6, 542 (1964)].
- ¹⁹³J. Knights, J. Non-Cryst. Solids 35-36, 159 (1980).
- ¹⁹⁴J. A. Reimer, R. W. Vaughan, and J. Knights, Phys. Rev. Lett. 44, 193 (1980).

¹⁹⁵H. Fritzsche, in: Fundamental Physics of Amorphous Semiconductors: Proc. of the Kyoto Summer Institute, Kyoto, Japan, Sep. 8-11, 1980, ed. F. Yonezawa, Springer-Verlag, Berlin, Heidelberg, New York, 1981, p. 1.

- ¹⁹⁶W. Paul, *ibid.*, p. 72.
- ¹⁹⁷M. H. Brodsky, *ibid*, p. 56; Solid State Commun. 36, 55 (1980).
- ¹⁹⁸P. G. Le Comber and W. E. Spear, Topics in Appl. Phys. 36, 251 (1979).

- ¹⁹⁹G. Lucovsky, see Ref. 195, p. 87.
- ²⁰⁰V. L. Bonch-Bruevich and A. G. Mironov, Opt. Spektrosk., 1983.
- ²⁰¹A. G. Mironov, Phys. Status Solidi B 101, 613 (1980).
- ²⁰²E. M. Gershenzon, V. A. Il'in, I. N. Kurilenko, L. B. Litvak-Gorskaya, R. I. Rabinovich, and S. R. Filonovich, in: Proc. 6th International Conference on Amorphous and Liquid Semiconductors: Electronic States in Noncrystalline Semiconductors, Nauka, L., 1976, p. 338.
- ²⁰³D. J. Thouless, Phys. Rev. Lett. 39, 1167 (1977).
- ²⁰⁴M, Pepper, Proc. R. Soc. London Ser. A 353, 225 (1971).
- ²⁰⁵I. P. Zvyagin, Philos. Mag., 1983.
- ²⁰⁶A. S. Isoelevich, Zh. Eksp. Teor. Fiz. 81, 1508 (1982) [Sov. Phys. JETP 54, 800 (1982)].
- ²⁰⁷V. I. Perel' and I. N. Yassievich, Fiz. Tverd. Tela (Leningrad), 1983.
- ²⁰⁸W. L. McMillan, Phys. Rev. B 24, 2739 (1981).
- ²⁰⁹B. Shapiro and E. Abrahams, *ibid.*, p. 4889.
- ²¹⁰G. Tauc, R. Grigorovici, and A. Vancu, Phys. Status Solidi 15, 627 (1966).

- ²¹¹G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. 47, 1480 (1981).
- ²¹²G. D. Cody, T. Tiedje, B. Abeles, T. D. Moustakas, B. Brooks, and Y. Goldstein, J. Phys. (Paris), 42, Suppl. C-4, 301 (1981).
- ²¹³I. L. Robert, B. Pistoulet, A. Raymond, R. L. Aulombard, C. Bernard, and C. Bousquet, Rev. Phys. Appl. 13, 246 (1978).
- ²¹⁴V. V. Voronkov, G. A. Voronkova, B. V. Zubov, V. V. Kalinushkin, B. B. Kryletskii, T. M. Murina, and A. M. Prokhorov, Fiz. Tverd. Tel a (Leningrad) **19**, 1784 (1977) [Sov. Phys. Solid State **19**, 1040 (1977)].
- ²¹⁵P. G. Zhumatil, see Ref. 202, p. 111.
- ²¹⁶P.G. Zhumatii, Vestn. Mosk. Univ. Fiz. Astron. 17, 436 (1979).
- ²¹⁷P. G. Zhumatii, Phys. Status Solidi B 75, 61 (1975).

Translated by M. V. King