Hall effect under hopping conduction conditions

Yu. M. Gal'perin, E. P. German, and V. G. Karpov

A. F. Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad
(Submitted 1 August 1990)
Zh. Eksp. Teor. Fiz. **99**, 343–356 (January 1991)

A theory of the Hall effect under hopping conduction conditions is developed. General expressions are obtained and these are used to derive the Hall mobility μ_H in specific cases such as variable-range-hopping and ε_3 conduction. An allowance is made of the effects of the electron–electron Coulomb correlations, leading to the appearance of a Coulomb gap. It is shown that under hopping conduction conditions the Hall effect has a correlation radius $L_H \gg L$, when L is the correlation radius of an infinite cluster governing the conductivity. In practice, the value of L_H can exceed the dimensions of a sample, which gives rise to size effects. These effects include a possible activated temperature dependence of the mobility $\mu_H(T)$. The results are used to analyze the experimental data on the Hall effect in amorphous semiconductors.

1. INTRODUCTION

Many experimental and theoretical investigations have been made of hopping conduction in disordered systems. This conduction is characterized by an activated temperature dependence

$$\sigma=\sigma_0 \exp[-\xi_c(T)], \qquad (1)$$

where $\xi_c(T)\gg 1$ is governed by the hopping conduction mechanism. For example, when the Mott law is obeyed it is found that $\xi_c(T)=(T_0/T)^{1/4}$ or $\xi_c(T)=(T_0/T)^{1/2}$, if there is a Coulomb gap. At higher temperatures the corresponding relationship is $\xi=\varepsilon_3/T$ (Refs. 1 and 2).

One of the important and still unsolved problems is the mechanism of the Hall effect under hopping conduction conditions. This problem is tackled below. We shall be interested mainly in the activation energy of the Hall effect (Hall field E_H), and its sign. We shall also obtain a rough estimate of the absolute value of the Hall effect and compare it with experimental data.

The data obtained in experiments on the Hall effect under hopping conditions are largely contradictory. A general observation is that the Hall effect is much less than under the usual band conduction conditions. In the hopping case the Hall effect in doped crystalline semiconductors is so small (because of the low resistivity of the samples) that the measurements themselves become very difficult. Nevertheless, some authors (see, for example, Refs. 3–5) report an observation of the Hall effect in doped semiconductors in the hopping conduction range. The Hall effect in amorphous semiconductors exhibits what is known as the double sign anomaly.[6-9] This is manifested as follows: when a sample has n-type conduction (deduced from measurements of the thermoelectric power) the Hall effect is positive, whereas in the case of p-type conduction it is negative.

Moreover, there is no agreement about the theory of the Hall effect under hopping conduction conditions. The first theoretical investigation of this case was made by Holstein.[10] He considered the ac hopping Hall effect. The problem then simplifies because the main contribution to the Hall emf comes from spatially isolated configurations of centers for which the population relaxation time is of the order of the reciprocal of the frequency of the current ($\omega\tau\approx 1$), similar to that encountered in the familiar Pollak–Geballe ac hopping conduction mechanism. These configurations are the sources of an alternating Hall emf and, moreover, in the ac case there is no need to solve the nontrivial problem of how these hole emf sources are connected to the Hall contacts. We shall see later that in the dc case this problem can be solved using percolation theory. Some variants of this approach had been proposed earlier by Bottger and Bryksin,[11] Butcher and Kumar,[12] and by Friedman and Pollak.[13] Nemeth and Muhlschlegel[14] proposed a heuristic approach to calculations of the Hall effect based on the formal similarity of the expressions for the transverse (σ_{xy}) and longitudinal (σ_{xx}) conductivities. Following the optimization procedure employed by Mott to derive Eq. (1), the authors of Ref. 14 demonstrated that the activation energy of the Hall effect in the variable-range hopping (VRH) case is of the order of $0.35T(T_0/T)^{1/4}$, i.e., it is considerably less than the activation energy of conduction. One should mention here that the optimization procedure used by Mott to obtain Eq. (1) is essentially based on the requirement of connectivity of a hopping path. This requirement is not necessary in optimization of the Hall emf. The results reported in Refs. 11–14 do however differ even in respect of the argument of the exponential function describing the Hall effect.

Our aim will be to derive a general expression which can be used to consider the various limiting cases representing amorphous and doped crystalline semiconductors. We shall consider hopping conduction in an impurity band with constant and variable length of the jumps, as well as the role of the Coulomb correlation of the occupancy numbers of the site electron states. We shall ignore the contribution of the band carriers to the Hall effect.

The present paper is organized as follows. We shall first determine the elementary sources of the Hall emf, starting from microscopic parameters (Sec. 2). We shall then discuss the averaging of the emf's contributed by the elementary Hall emf sources in a disordered system of centers (Sec. 3). Next, we shall consider special cases representing various temperature ranges (Sec. 4). We shall conclude with some final comments and a comparison of the theory with selected experimental data (Sec. 5).

 0038-5646/91/010193-08$03.00 © 1991 American Institute of Physics

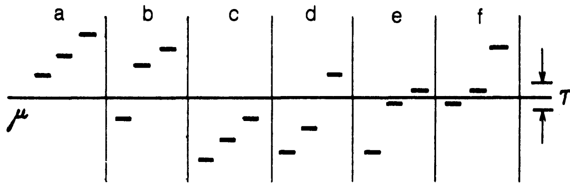

FIG. 1. Configuration of levels corresponding to electron (a, b), hole (c, d), and mixed (e, f) triads. The Fermi level is labeled μ and the thermal energy scale T is shown.

2. ELEMENTARY SOURCES OF THE HALL emf

A self-consistent analysis of hopping conduction is based on the concept of a random network of resistances put forward by Abrahams and Miller.[2] Each resistance connects two sites and is selected to be equal to the coefficient of proportionality between the difference between the electrochemical potentials at this pair of sites, on the one hand, and the hopping current, on the other. However, this approach is insufficient for the description of the Hall effect because a magnetic field **H**, considered in the linear approximation, does not influence the probability of a transition between sites. The minimum number of sites needed for the appearance of a linear (in **H**) change in the probabilities of intersite transitions is three. We shall consider configurations of three sites with random energies ε_i, ε_j, and ε_k (measured from the Fermi level μ) and random intersite distances r_{ij}, r_{jk}, and r_{ki}. We shall call these site triads and we shall show that they are elementary sources of the Hall emf. We shall obtain an expression for the Hall emf developed by such a triad.

We shall base our analysis on the balance equation for transitions from one of the sites in a triad (i) under steady-state conditions:

$$-n_i(1-n_j)P_{ij}{}'-n_i(1-n_k)P_{ik}{}'+n_j(1-n_i)P_{ji}{}'+n_k(1-n_i)P_{ki}{}'=0. \quad (2)$$

Here, n_i is the occupancy number of the ith site and P'_{ij} is the total probability of transitions between the sites i and j. In writing down Eq. (2) we shall assume that the currents flowing into and out of a triad are not affected by the application of a magnetic field. This formulation of the problem corresponds physically to finding the Hall emf under constant-current conditions. It allows us to consider only the transitions within a triad in the balance equation (2).

It must be stressed that the quantities P'_{ij} in Eq. (2) represent the probabilities of single-particle transitions. This does not require any special comment if we are dealing with a pair of sites, so that a transition occurs when just one particle is present at one of the two centers. In the case of a triad the situation is fundamentally different. The transitions in a triad are possible in the presence of one or two particles at the sites. Both cases are relevant, because we are speaking of states close to the Fermi level (this is true in any case in the VRH range). Single-particle states can be used if right from the beginning we can distinguish the electron and hole triads. We shall identify as the electron triads those for which the probability of occupancy by more than one electron is negligible. In the case of the hole triads the probability of occupancy by fewer than two electrons is negligible. The corresponding configurations of the energy levels are shown in Fig. 1. This figure includes also the examples of mixed cases when triads with comparable probabilities can be populated by any number of electrons from zero to three. We shall show later that the contribution of mixed triads to the Hall effect is exponentially small.

This classification of triads makes it possible to use Eq. (2) with single-particle probabilities for which explicit and relatively simple expressions are available. At the end of the present section we shall consider the specific case of an electron triad.

We shall be interested in the Hall emf which is a linear function of the magnetic field **H**. According to Ref. 10, only real transitions involving states of the third site are then possible, so that the effects of correlation of the site occupancy numbers are manifested. The contribution of these effects depends on the intrasite correlation energy U of the particles. We shall assume that this energy is sufficiently high, i.e., that $U \gg T\xi_c$, which is usually true in real situations. Hence, it follows that the presence of the third site affects the probability of a transition between the other two sites only when the third site is unoccupied.

Bearing all these points in mind, we can write down

$$P_{ij}{}'=P_{ij}+P_{ij}{}^H(1-n_k), \quad (3)$$

where P_{ij} is the probability of a transition from a site i to a site j in the absence of a magnetic field and P^H_{ij} is the change in this probability due to the application of a magnetic field. Using a factor $(1-n_k)$, we shall allow for the intrasite correlation at an intermediate site.

Our task is to determine the influence of an external magnetic field on the difference between the electrochemical potentials at two sites. We shall do this by substituting in Eq. (2) the expression

$$n_i=n_i{}^0+\delta n_i, \quad (4)$$

where δn_i is the change in the population due to the external magnetic field and n_i^0 is the population of the ith level when $H=0$ (in the presence of an electric field). We can write down

$$n_i{}^0=\left[1+\frac{1}{2}\exp\left(\frac{\varepsilon_i-\mu_i}{T}\right)\right]^{-1}. \quad (5)$$

The quasi-Fermi level μ_i differs from the Fermi level μ because of the presence of an electric field. The quantity δn_i may be related to $\delta U_i \equiv \delta(\varepsilon_i - \mu_i)$, i.e., it may be related to the change in the electrochemical potentials of the sites in a magnetic field:

$$\delta n_i=-n_i{}^0(1-n_i{}^0)\delta U_i/2T. \quad (6)$$

Substituting in Eq. (2) the relationships (3), (5), and (6) and retaining only the corrections linear in P^H and δU, we can describe the quantities $\delta U_i, \delta U_j, \delta U_k$ in terms of changes of the probabilities of the intersite transitions on application of an external magnetic field. In the process of obtaining these expressions it is useful to apply the symmetry considerations and represent the probabilities in Eqs. (2) and (3) in the form[10]

$$P_{ij}=L_{ij}\exp(\varepsilon_i/T),\; P_{ji}=L_{ji}\exp(\varepsilon_j/T),\; L_{ij}=L_{ji},$$
$$P_{ij}^H=\mathscr{L}_{jki}\exp(\varepsilon_i/T),\; P_{ji}^H=\mathscr{L}_{ikj}\exp(\varepsilon_j/T),$$
$$P_{ik}^H=\mathscr{L}_{kji}\exp(\varepsilon_i/T),\; P_{ki}^H=\mathscr{L}_{ijk}\exp(\varepsilon_k/T), \tag{7}$$
$$\mathscr{L}_{ikj}=-\mathscr{L}_{jki}=\mathscr{L}_{kji}=-\mathscr{L}_{ijk}$$

(explicit expressions for P_{ij} and P_{ij}^H are given above). Moreover, simplifications can be made using familiar expressions[2] for intersite currents and resistances, which in the notation adopted here become

$$I_{jk}^{(e)}=-e[P_{jk}n_j^{(e)}(1-n_k^{(e)})-P_{kj}n_k^{(e)}(1-n_j^{(e)})] \tag{8}$$

(the expression for a hole triad differs by the substitutions $e\to -e$, $n_s^{(e)}\to n_s^{(h)}$, $eH\to -eH$) and

$$R_{ik}^{-1}=\frac{e^2}{2T}L_{ik}n_i^0n_k^0\exp\left(\frac{\varepsilon_i+\varepsilon_k}{T}\right). \tag{9}$$

Subject to these simplifications, Eq. (2) can be transformed to

$$\frac{\delta U_{ij}^{(e)}}{R_{ij}}-\frac{\delta U_{ki}^{(e)}}{R_{ki}}=e(1-n_i^0)I_{jk}^{(e)}\mathscr{E}_{jk}^{(e)}, \tag{10}$$

where

$$\delta U_{ij}=\delta U_i-\delta U_j.$$

The quantity $\mathscr{E}_{jk}$ is defined by the relationship

$$\mathscr{E}_{jk}=\frac{P_{jk}^H}{P_{jk}}=\frac{\mathscr{L}_{kij}}{L_{jk}}, \tag{11}$$

which shows that it is proportional to the magnetic field H.

We shall be interested in the difference φ_{ik} between electrical potentials at the sites in a triad. We can go over from the difference between the energies U_{ik} to the difference between the electrical potentials φ_{ik} if we divide U_{ik} by the absolute value of the electron charge e. We then obtain from Eq. (10)

$$\frac{\delta\varphi_{ij}}{R_{ij}}-\frac{\delta\varphi_{ki}}{R_{ki}}=(1-n_i^0)I_{jk}^{(e)}\mathscr{E}_{jk}^{(e)}. \tag{12}$$

The physical meaning of Eq. (12) requires some comments. At first sight, the left-hand side of this expression represents the difference between the magnetically induced currents in a triad: those flowing into a site i and flowing out of this site. Under the conditions of constancy of all the other (external to a triad) currents, such an interpretation implies that the left-hand side of Eq. (12) vanishes, which contradicts the fact that the right-hand side is nonzero. This contradiction is removed if we assume that the application of an external magnetic field gives rise to effective sources of the voltage in the branches of the triad. Physically the appearance of such sources is responsible for the Hall emf developed by a triad.

We can use Eq. (12) to write down the expression for the Hall emf by an appropriate selection of the algorithm to be used. We shall assume that in general we are dealing with an object ("black box") through which an external current passes in the presence of a magnetic field. What should be the procedure in finding the Hall emf? The generally accepted answer to this question is that the Hall emf is the distance between the potentials at two points which in the absence of a magnetic field are equipotential.

The application of this algorithm to a triad will be made by replacing it with a star configuration (Fig. 2). We shall consider only those triads which are connected to an external current-carrying circuit mainly by two sites (j and k in Fig. 2), i.e., the external current flowing through the third site i is relatively small. This connection of a triad to a current-carrying circuit is most typical. The probability that all three sites of a triad are connected equally to an external current-carrying circuit is relatively low. In the final analysis, this follows from the low probability of encountering triads which are effective sources of the Hall emf in a system of random centers (this will be demonstrated in greater detail in Sec. 4).

Transformation from a triangle to a star gives the voltage of an effective Hall source (see Fig. 2):

$$\mathscr{E}_i=\frac{R_{ki}\delta U_{ij}-R_{ij}\delta U_{ki}}{R_{ij}+R_{jk}+R_{ki}}. \tag{13}$$

Using Eq. (10), we obtain

$$\mathscr{E}_i=\frac{R_{ij}R_{ki}}{R_{ij}+R_{jk}+R_{ki}}(1-n_i^0)I_{jk}\mathscr{E}_{jk} \tag{14}$$

and similar expressions for $\mathscr{E}_j$ and $\mathscr{E}_k$.

Although in this section we shall not consider directly the procedure for averaging the Hall emf's in triads, we must mention here one circumstance which allows us to limit greatly the class of triads under investigation. It is clear from Fig. 2 that the currents I_{ki} and I_{ij} are the same. On the other hand, it follows from the symmetry of the problem in a magnetic field and directly from Eqs. (11) and (7) that the signs of the quantities $\mathscr{E}_{ki}$ and $\mathscr{E}_{ij}$ are opposite. This means that the signs of $\mathscr{E}_j$ and $\mathscr{E}_k$ are also opposite and after averaging their sum vanishes. Consequently, we shall consider only an emf $\mathscr{E}_i$ at a "dead" (i.e., not connected to a current-carrying circuit) site in a triad. Expressing the current I_{jk} in terms of the external current I, we finally obtain

$$\mathscr{E}_i=\frac{R_{ij}R_{ki}(R_{ij}+R_{ki})}{(R_{ij}+R_{jk}+R_{ki})^2}(1-n_i^0)I\mathscr{E}_{jk}. \tag{15}$$

It should be noted that in the absence of a magnetic field the potential at the central point of the star is equal to the potential at the site i. Therefore, the value of $\mathscr{E}_i$ given by Eq. (15) satisfies the adopted algorithm for the determination of the Hall emf and will be regarded as this emf in future.

It remains to give our expressions for the quantities R_{ij}, R_{jk}, R_{ki}, $\mathscr{E}_{jk}$ occurring in Eq. (15). In the case of resistances we shall use the expressions[2] known from an analysis of hopping conduction:

$$R_{ij}=R_0\exp\left(-\frac{2r_{ij}}{a}-\frac{E_{ij}}{T}\right)\equiv R_0\exp(\xi_{ij}), \tag{16}$$

$$E_{ij}={}^1/_2[|\varepsilon_i|+|\varepsilon_j|+|\varepsilon_i-\varepsilon_j|]. \tag{16a}$$

The expression (16a) must be refined if an allowance is made for the Coulomb correlations. In the case when $\varepsilon_i\varepsilon_j<0$ this expression acquires an additional term $-e^2/\varkappa r_{ij}$ (Ref. 2). However, we shall show that triads characterized by $\varepsilon_i\varepsilon_j<0$ are unimportant in the problem under discussion [see the comments after Eq. (28)]. The explicit form of the preexponential factor in Eq. (16) will not be used and we shall not give it here. The expression (11) for $\mathscr{E}_{jk}$ contains the change in the probability of an intercenter transition due to the application of a magnetic field. This change was obtained by Holstein[10] and is given by

$$P_{jk}^H=\frac{\hbar}{4W_{ij}W_{jk}W_{ki}}[W_{ik}^2P_{jk}P_{ji}+W_{kj}^2P_{ji}P_{ik}+W_{ji}^2P_{jk}P_{ki}]\frac{e\mathbf{HS}}{\hbar c}, \tag{17}$$

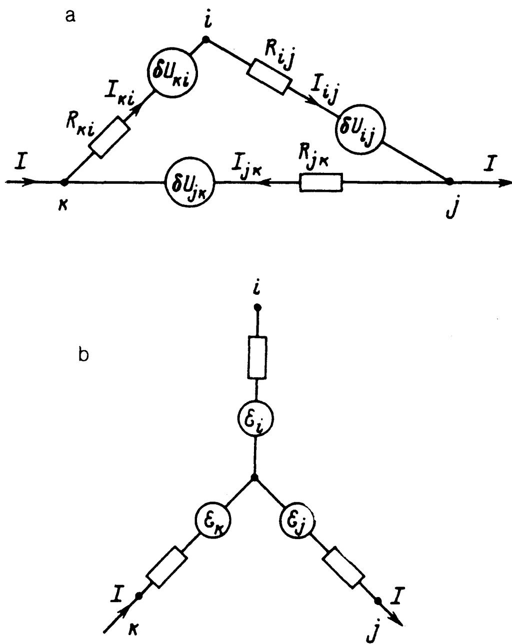

FIG. 2. Triad ijk (a) and its equivalent circuit in the form of a star (b). Here, I is the external current flowing through the triad.

where **S** is the vector area of the triangle ijk. It is convenient to represent the energy integral representing hopping between the sites in the form

$$W_{ik}\approx V_{ik}\exp(-r_{ik}/a), \tag{18}$$

where the spatial dependence of the amplitude V_{ik} can be ignored. The probability of an intercenter transition P_{ik} can be described in the usual way:

$$P_{jk}=\nu_{ph}\left(\frac{W_{jk}}{\varepsilon_j-\varepsilon_k}\right)^2\begin{cases}\left[\exp\left(\frac{\varepsilon_j-\varepsilon_k}{T}\right)-1\right]^{-1}, & \varepsilon_j<\varepsilon_k\\ \left[\exp\left(\frac{\varepsilon_j-\varepsilon_k}{T}\right)-1\right]^{-1}+1, & \varepsilon_j>\varepsilon_k.\end{cases} \tag{19}$$

The frequency factor $\nu_{\rm ph}$ may also exhibit a relatively weak (obeying a power law) dependence on the transition energy $|\varepsilon_j-\varepsilon_k|$ (see, for example, Refs. 2 and 15).

In view of the presence of exponentially strong energy dependences in Eqs. (16) and (17), we shall ignore the much weaker power-law dependences. This means that the preexponential factors in the subsequent expressions will be inaccurate, but it will simplify greatly our analysis. In this approximation Eqs. (17)–(19) allow us to derive an approximate expression for $\mathscr{E}_{ik}$ in the form

$$\mathscr{E}_{ik}=\frac{s\hbar\nu_{ph}V}{E_{jk}^2}\frac{\exp(\varepsilon_i/T)+\exp(\varepsilon_k/T)}{\exp(\varepsilon_i/T)+\exp(\varepsilon_k/T)+\exp(\varepsilon_j/T)}\times\exp\left(\frac{r_{jk}-r_{ij}-r_{ki}}{a}\right)\frac{e\mathbf{HS}}{\hbar c}, \tag{20}$$

where $s=\operatorname{sign}(V_{ij}V_{jk}V_{ki})$ and V is the characteristic value of the moduli of the quantities V_{ij}, V_{jk}, and V_{ki}.

Equation (20) ignores the Coulomb correlation, which is not important in the situations of interest to us. The solution of the problem of the Hall emf developed by an electron triad is given approximately (apart from numerical factors) by Eqs. (15), (16), and (20).

We shall conclude this section with a discussion of the sign of the Hall effect. The above analysis deals specifically with electron triads. Hole triads can be analyzed in exactly the same way. Once again we obtain Eq. (20). In particular, the signs of $\mathscr{E}_{jk}$ are the same for electron and hole triads with given values of V_{ij}, V_{jk}, and V_{ki}. In turn, the signs of the amplitudes V of a transition are governed by the nature of the localization centers, namely by the signs of the potentials acting on an electron. In the case of doped semiconductors these potentials are negative for donors and positive for acceptors. In amorphous semiconductors, where the localization of states is due to the structural disorder, the sign of V_{ij} is negative for the states which split off from the conduction band and positive for those that split off from the valence band.

3. AVERAGING. MACROSCOPIC HALL EFFECT

In this section we shall consider a macroscopic sample exhibiting hopping conduction and we shall find its effective Hall parameters: the Hall coefficient

$$\mathscr{R}_H=E_H/jH, \tag{21}$$

where E_H is the Hall field and j is the current density, and also the Hall mobility

$$\mu_H=c\mathscr{R}_H\sigma. \tag{22}$$

Our analysis will be based on percolation theory ideas,[2] according to which hopping transport of charge occurs between sites forming an infinite cluster, which penetrates the whole sample (Fig. 3). This infinite cluster has a correlation radius

$$L\approx a\xi_c^{1+\nu}, \tag{23}$$

where $\nu\approx0.9$ in the three-dimensional case and $\nu\approx1.3$ in the two-dimensional case.[2] Each cell (of size L) of this infinite

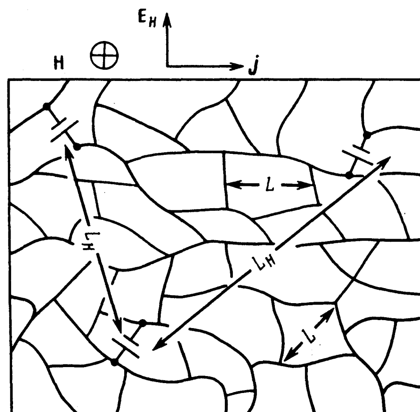

FIG. 3. Infinite cluster in a macroscopic sample. The directions of the magnetic field **H** and of the average values of the Hall field $\mathbf{E}_H$ and of the current **j** are shown; L is the correlation radius of the infinite cluster and L_H is the correlation radius of the Hall effect (representing a typical distance between Hall emf sources).

cluster represents a connected set of random resistances $R = R_0 \exp\xi$, where the random quantity ξ is bounded from above by

$$\xi \leqslant \xi_c. \qquad (24)$$

The order of magnitude of the resistance of each infinite-cluster cell is given by $R_c = R_0 \exp\xi_c$. Topologically, an infinite cluster is a network of twisted macrobonds. A typical distance between the sites connected by such macrobonds is of the order of L.

Since the voltages across triads considered in the preceding section are proportional to the current, it is clear that these triads should be "incorporated" in an infinite cluster where practically the whole current is concentrated. The resistance of each triad to the current flowing through it does not exceed R_c. Such triads can be located at nodes of an infinite cluster or, more likely, at its macrobonds (Fig. 3). A voltage developed by a triad located far from an infinite cluster node (i.e., located at its macrobond) is transferred to the Hall contacts through a circuit containing resistances R much higher than R_c. Therefore, this circuit is shunted by resistances R_c of the infinite cluster macrobonds, so that the transferred voltage is exponentially small, since it is $R_c/R \ll 1$ times less than the initial value. Therefore, triads located far from the infinite cluster nodes make only a small contribution to the net Hall emf. We shall therefore consider only the triads located at nodes of an infinite cluster where the shunting is unimportant.

We shall consider again Hall sources of the same type, for example, electron sources. First of all, we note that in spite of the random signs of the vector areas **S** in Eq. (20), the net Hall emf does not vanish as a result of averaging. This is because there is a preferred direction along which the source emf's are summed. Consequently, the Hall emf sources with opposite values of **S** are not canceled out but their absolute values are added. This is demonstrated in Fig. 4.

Another factor which could give rise to mutual balancing out of the emf's developed by elementary Hall sources is the twisted nature of the macrobonds in an infinite cluster. For this reason the current in some parts of an infinite cluster can flow opposite to the direction of the average current (Fig. 5a). Moreover, a random connection of elementary Hall emf's with one or another Hall electrode is possible (Fig. 5b). However, effects of this kind can only give rise to a comparatively weak (partial) compensation of the Hall effect. In fact, we can regard a macrobond as a path of a random walk consisting of $N = N_+ + N_-$ steps forward (N_+) and backward (N_-). The distance between the initial and final points of such a random walk is $l = N_+ b - N_- b > 0$ (Fig. 5). Hence, it follows that the number of steps in the "correct" direction (N_+) is greater, i.e., there is no compensation. Similar effects for two-phase systems were considered by Straley.[16]

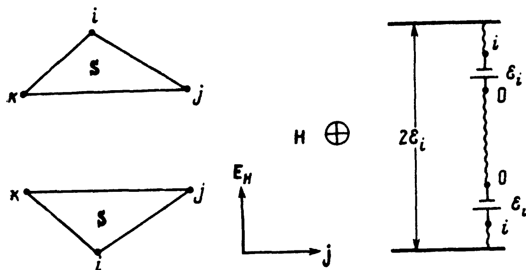

FIG. 4. Addition of emf's from two Hall sources with opposite signs of the vector areas.

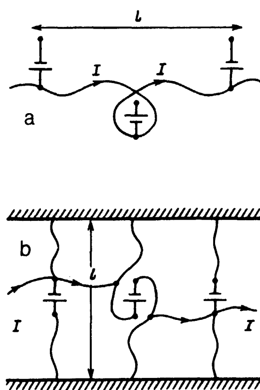

FIG. 5. Partial compensation of the emf's developed by Hall sources due to the twisted nature of the infinite cluster macrobonds: a) change in the direction of the current; b) change in the direction of connection of a source.

These conclusions about unimportance (with an exponential precision) of fluctuations of the signs of the vector areas **S** and of the twistedness of the infinite cluster macrobonds allow us to ignore the signs of $\mathscr{E}_i$ and carry out averaging of the absolute values alone. We shall carry out such averaging in two stages. First of all, we shall determine the optimal exponentially significant triads and the Hall emf's created by them. We shall then estimate the probability of formation of an optimal triad.

The existence of a factor $(1 - n_i^0)$ in the expression (15) for $\mathscr{E}_i$ allows us to consider only the case when $\varepsilon_i > 0$. Moreover, it follows from Eq. (15) that the value of $\mathscr{E}_i$ is not small only if $R_{jk} \leqslant R_{ki} + R_{ij}$. If we drop R_{jk} from Eq. (15), we find that $\mathscr{E}_i$ is governed by the smaller of the two remaining resistances and, therefore, it follows that the maximum value of $\mathscr{E}_i$ is reached when $R_{ik} \approx R_{ij}$. Therefore, apart from a term of the order of unity, the optimal triad should satisfy the condition

$$\frac{2r_{ij}}{a} - \frac{2r_{ki}}{a} + \frac{E_{ij}}{T} - \frac{E_{ki}}{T} = 0.$$

Using the last relationship, we find from Eqs. (15) and (20) that

$$\mathscr{E}_i \approx IR_0 \frac{s\hbar\nu_{ph}V}{E_{jk}^2} \frac{e\mathbf{HS}}{\hbar c} \frac{\exp(\varepsilon_j/T) + \exp(\varepsilon_k/T)}{\exp(\varepsilon_i/T) + \exp(\varepsilon_k/T) + \exp(\varepsilon_j/T)} \times \exp\left[\frac{r_{jk}}{a} + \frac{E_{ij} + E_{ki}}{2T}\right]. \qquad (25)$$

Equation (25) can be optimized in respect of the site energies. We can do this using the definitions given in Eq. (16) for E_{ij} and E_{ki} and consider the cases of different relationships between the level positions. It is necessary to allow also for the fact that $\varepsilon_i > 0$ and that only one of the other two

levels may be negative (because we are considering an electron triad). Such an analysis shows that in all cases Eq. (25) can be represented in the form

$$\mathscr{E}_i \approx I R_0 \frac{s\hbar v_{ph} V}{E_{jk}^2} \frac{e\mathbf{HS}}{\hbar c} \exp\left[\frac{\xi_{ik}}{2} + \frac{\varepsilon_\alpha}{2T}\right], \tag{26}$$

where ε_α is the energy of that of the levels ε_i, ε_j, and ε_k, which is between the other two. It follows from the inequality (24) that the maximum value of the quantities in the argument of the exponential function in Eq. (26) is $(\xi_{jk})_{\max} = \xi_c$. Therefore, the maximum emf developed by a Hall source is

$$\mathscr{E}_i \approx I R_0 \frac{s\hbar v_{ph} V}{T^2} \frac{e\mathbf{HS}}{\hbar c} \exp\left[\frac{\xi_c}{2} + \frac{\varepsilon_{max}}{2T}\right], \tag{27}$$

where

$$\varepsilon_{max} = \max(|\varepsilon_i|, |\varepsilon_j|, |\varepsilon_k|).$$

The above analysis demonstrates, in fact, that the maximum voltage is obtained for triads with the same resistances close to the critical value, i.e., when

$$R_{ij} \approx R_{jk} \approx R_{ki} \approx R_c. \tag{28}$$

Moreover, it must be stressed that the maximum value of ξ_c in the argument of the exponential function (27) is obtained when all three levels ε_i, ε_j, and ε_k are on one side of the Fermi level and two of them are in a strip of width of the order of T near the maximum energy $\varepsilon_{\max}$. The expression for $\varepsilon_{\max}$ is governed by the hopping conduction mechanism and will be made more specific below. It also follows from Eq. (27) that the main contribution comes from triads with the minimum geometric dimensions ($r_{ij} \approx r_{jk} \approx r_{ki} \equiv r_{\min}$), which correspond to the maximum energy $\varepsilon_{\max}$. These requirements in respect of the energies and intersite distances determine completely the parameters of the optimal triad which develops an emf described by Eq. (27).

Since our analysis allows only for the exponentially significant dependences, the conditions of Eq. (28) are approximate. This means that in the optimal triad case the arguments ξ_{ij}, ξ_{jk}, ξ_{ki} of the exponential functions may differ from ξ_c by about unity. This allows us to estimate the probability of finding an optimal triad.

If this probability is P^3, then a typical distance between such triads is

$$L_H \approx L/P. \tag{29}$$

This distance corresponds to the Hall field

$$E_H \approx \mathscr{E}/L_H. \tag{30}$$

Using the expressions for the current density and the conductivity given by

$$j = I/L^2, \quad \sigma = 1/R_c L, \tag{31}$$

and also the definition (22) of the Hall mobility, we obtain

$$\mu_H = P \frac{c\mathscr{E}}{IHR_c} \approx P \frac{s v_{ph} V e r_{min}^2}{T^2} \exp\left(-\frac{\xi_c}{2} + \frac{\varepsilon_{max}}{2T}\right). \tag{32}$$

Equation (32) is the main result of the present section. We shall apply this result specifically to different hopping conduction cases.

4. HALL EFFECT IN VARIOUS LIMITING CASES

We shall begin with an analysis of doped crystalline semiconductors. In this case the process of hopping conduction takes place in an impurity energy band of donor or acceptor levels. The following types of hopping conduction can be distinguished.

a) The high-temperature hopping conduction characterized by a variable jump length, i.e., variable-range hopping (VRH), is described by the familiar Mott law

$$\xi_c = (T_0/T)^{1/4}, \quad T_0 = \beta/g_0 a^3, \tag{33}$$

where g_0 is the density of states at the Fermi level and β is a numerical coefficient.[2] The main contribution to the conductivity then comes from carrier jumps between centers separated by a distance $r \approx a\xi_c$ with single-site energies in a band of the order of $T\xi_c$ wide and located near the Fermi level. Typical values are $\xi_c \approx 20$–30. The VRH conditions are retained on increase in the temperature T until the characteristic energy $T\xi_c$ exceeds a typical width of the impurity energy band.

b) The low-temperature conduction characterized by a variable jump length is described by Eq. (1), but the parameters are now

$$\xi_c = (T_1/T)^{1/2}, \quad T_1 = \beta_1 e^2/\varkappa a, \tag{34}$$

where $\varkappa$ is the permittivity and β_1 is a numerical coefficient[2] (this is known as the Shklovskii–Éfros law). The nature of the dependence $\xi_c(T)$ of Eq. (34) is governed by the influence of electron–electron Coulomb correlations. They give rise to an energy dependence of the density of states $g \propto \varepsilon^2$ in the region of what is known as the Coulomb gap near the Fermi level, which is typically $\Delta = e^3 g_0^{1/2} \varkappa^{-3/2}$ wide. The dependence (34) applies until the characteristic jump energy $T\xi_c$ exceeds the width of the Coulomb gap Δ.

c) The ε_3 conduction case is observed at such high temperatures that the energy $T\xi_c$ exceeds the width of the impurity energy band. Then, thermal activation has relatively little effect on the jump probabilities and the jump lengths are of the order of the average intercenter distance. In this case the activation energy of hopping is close to the energy gap μ between a relatively narrow density-of-states peak and the Fermi level. This level lies within a density-of-states tail: above the peak in the case of weak compensation and below the peak if the compensation is strong. Therefore, in the ε_3 conduction case we have

$$\xi_c \approx \xi_c^0 + \mu/T, \quad \xi_c^0 \approx \alpha/N^{1/3} a \gg \mu/T, \tag{35}$$

where N is the total concentration of sites in the impurity energy band and α is a numerical coefficient.[2] The finite width of the density-of-states peak and the profiles of its tails are practically of no importance in the determination of ξ_c.

We shall find for these three cases the probabilities P of encountering resistances forming optimal triads of interest to us. In this procedure it is convenient to introduce a probability density $\rho(\xi, r; \xi_c)$ and to find in an infinite cluster a pair of states separated by the distance r, such that $\ln(R/R_0) = \xi$. The required probability is then described by

$$P \approx \int_{\xi_c - 1}^{\xi_c} d\xi \int_{r_{min} - a/2}^{r_{min} + a/2} d^3 r\, \rho(\xi, r; \xi_c). \tag{36}$$

By definition, we have

$$\rho(\xi, r; \xi_c) = Z \cdot 4\pi r^2 \int\int d\varepsilon_i \, d\varepsilon_j \, g(\varepsilon_i) g(\varepsilon_j) \delta\left[\xi - \frac{2r}{a} - \frac{\max(\varepsilon_i, \varepsilon_j)}{T}\right], \quad (37)$$

where the normalization coefficient Z is described by

$$\int_0^{\xi_c} d\xi \int_0^{\xi a/2} dr \, \rho(\xi, r; \xi_c) = 1. \quad (38)$$

In the three cases described above our calculations follow exactly the same procedure and are based on the relevant distribution functions:[17]

$$\text{a)} \quad g(\varepsilon) = g_0 = \text{const},$$
$$\text{b)} \quad g(\varepsilon) \propto \varepsilon^2,$$
$$\text{c)} \quad g(\varepsilon) \propto \delta(\varepsilon - \mu). \quad (39)$$

The results for the cases a), b), and c) are, respectively:

$$\rho(\xi, r; \xi_c) = \frac{480}{\xi_c^5 a^3} r^2 \left(\xi - \frac{2r}{a}\right),$$
$$\rho(\xi, r; \xi_c) = \frac{40\,824}{\xi_c^9 a^3} r^2 \left(\xi - \frac{2r}{a}\right)^5, \quad (40)$$
$$\rho(\xi, r; \xi_c) = \frac{24}{\xi_c^3 a^3} r^2 \delta\left(\xi - \frac{2r}{a} - \frac{\mu}{T}\right).$$

Substitution of these expressions into Eq. (36) allows us to determine the probability P for each of these cases and this probability occurs in Eq. (32) for the Hall mobility.

In writing down the final expressions we must bear in mind that in cases a) and b), which represent VRH, the minimum distance $r_{\min}$ is bounded from below by the condition of smallness of quantum repulsion between the levels $V\exp(-r/a)$ compared with the maximum jump energy $T\xi_c$. Therefore, in cases a) and b), we have

$$r_{min} = a \ln(V/T\xi_c) \ll a\xi_c, \quad (41)$$
$$\varepsilon_{max}/T = \xi_c - 2r_{min}/a = \xi_c - 2\ln(V/T\xi_c).$$

(It should be noted that the condition $V \gg T\xi_c$ is satisfied always for VRH.) In the case c) which represents the ε_3 conduction the scatter of the site energies is relatively small and in the first approximation we can assume that

$$\varepsilon_{max} = \mu,$$
$$r_{min} = (\xi_c - \mu/T) a/2 \approx \xi_c^0 a/2. \quad (42)$$

Consequently, the expression for the Hall mobility in the cases a) and b) becomes

$$\mu_H \approx s \frac{ea^2}{\hbar} \frac{\hbar\nu_{ph}}{T} \xi_c \left[\frac{\ln(V/T\xi_c)}{\xi_c}\right]^4, \quad (43)$$

whereas in the case c), we have

$$\mu_H \approx s \frac{ea^2}{\hbar} \frac{\hbar\nu_{ph}}{T} \xi_c^0 \exp\left(-\frac{\xi_c}{2}\right). \quad (43a)$$

The expression (43) is identical, with exponential precision, with the results of Refs. 11 and 13, obtained ignoring the Coulomb correlation of the occupancy numbers. For example, it is shown that this correlation does not affect the Hall coefficient within the limits of the experimental precision. The expression (43a) agrees with the results of Ref. 11 (again with exponential precision) but not with the results of Refs. 12 and 13.

The corresponding expressions for a typical distance between the optical triads in the cases a) and b) are

$$L_H \approx L\left[\frac{\xi_c^2}{\ln(V/T\xi_c)}\right]^2, \quad (44)$$

whereas in the case c), we have

$$L_H \approx L\xi_c. \quad (44a)$$

We considered above the case of doped semiconductors. The concept of an impurity energy band is meaningless in the case of amorphous semiconductors and localized states in such semiconductors fill almost completely the whole of the mobility gap, so that hopping conduction can occur only in the VRH case. Using the above expressions to describe VRH, we have to allow for the presence of two types of localized states: those split off from the valence band and those split off from the conduction band.[1] These states correspond to opposite signs of the matrix elements V_{ij} and, consequently, to opposite signs of s in the expressions for μ_H. Therefore, the resultant Hall mobility

$$\mu_H = \frac{\mu_{H1}\sigma_1 + \mu_{H2}\sigma_2}{\sigma_1 + \sigma_2} \quad (45)$$

can be positive or negative, depending on the nature of carriers (electrons or holes) that dominate the conduction process.

5. CONCLUDING COMMENTS

We shall estimate the Hall mobility in the cases discussed above. Such estimates are quite rough, in view of indeterminacy of the numerical coefficients in Eqs. (43) and (43a).

We shall consider amorphous materials and assume that $\xi_c \approx 10$, $a \approx 10$ Å, $\hbar\nu_{ph}/T \approx 0.1$, $\ln(V/T\xi_c) \approx 2$. Then, the estimate represented by Eq. (43) yields $\mu_H \approx 10^{-2}$ $cm^2 \cdot V^{-1} \cdot s^{-1}$. Experiments on amorphous semiconductors usually give values $|\mu_H| \approx 10^{-2}$–10^{-1} $cm^2 \cdot V^{-1} \cdot s^{-1}$ at $T = 200$–500 K (Ref. 1).

We must bear in mind that Eqs. (43) and (43a) are based on the assumption that the size q of a sample is much greater than a typical distance L_H between Hall emf sources. The distance L_H can be called the correlation radius of the Hall effect. If $q < L_H$, a sample as a rule does not contain optimal Hall sources of the emf described by Eq. (27). In

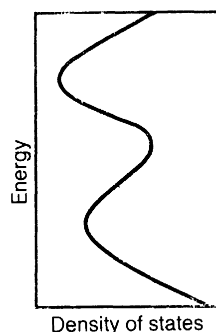

FIG. 6. Schematic representation of the behavior of the density of states in an amorphous semiconductor.[18]

this case the most effective (and not optimal) sources are active. The emf's developed by them are exponentially smaller than the emf of an optimal triad given by Eq. (27). The parameters of such Hall sources are random and self-averaging does not occur in the dimensions of a sample. We can therefore expect significant fluctuations of the Hall coefficient from sample to sample. It follows from Eqs. (26) and (27) that the Hall mobility has a random activation energy [$\approx(\varepsilon_{\max} - T\xi_{\max})/2$ for a given sample]. The size effects become more pronounced at low temperatures because, in accordance with Eqs. (44) and (44a), the value of L_H rises steeply on increase in ξ_c. This allows us to interpret the appearance of the activated temperature dependence of the Hall mobility at low temperatures.[1] It is possible that this applies also to the activated temperature dependence of μ_H of crystalline Si:P (Ref. 5). It should be pointed out that the correlation radius L_H may be quite large. Estimates based on Eqs. (44) and (44a) give for this quantity a fraction of a millimeter at low temperatures, which is comparable with the distances between the Hall contacts in the case of real samples.

Hopping conduction can explain also what is known as the double anomaly of the sign of the Hall effect exhibited by amorphous semiconductors. In fact, in accordance with the Cutler–Mott expression,[1] the thermoelectric power is governed by the energy derivative of the density of states at the Fermi level. In the case of the density of states usually assumed for amorphous semiconductors (Fig. 6) the signs of the Hall coefficient and thermoelectric power may be opposite, as can easily be demonstrated.

From our point of view, it therefore follows that the available experimental data are not in conflict with our hypothesis about the hopping mechanism of the Hall effect in amorphous semiconductors. Further progress in the identification of this mechanism should follow from investigations of fluctuations of the Hall coefficient from sample to sample and of the size effects (dependences of this coefficient on the dimensions of a sample, on the thickness of an amorphous film, etc.).

The authors are grateful to G. E. Pikus for discussing the work and to V. L. Gurevich for discussions and reading the manuscript.

[1] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, Clarendon Press, Oxford (1971).
[2] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors*, Springer Verlag, Berlin (1984).
[3] M. Amitay and M. Pollak, *Proc. Eighth Intern. Conf. on Physics of Semiconductors*, Kyoto, 1966, in J. Phys. Soc. Jpn. **21**, Suppl., 549 (1966).
[4] C. Yamanouchi, K. Mizuguchi, and W. Sasaki, J. Phys. Soc. Jpn. **22**, 859 (1967).
[5] D. W. Koon and T. G. Castner, Solid State Commun. **64**, 11 (1987).
[6] P. G. Le Comber, D. I. Jones, and W. E. Spear, Philos. Mag. **35**, 1173 (1977).
[7] J. Dresner, Appl. Phys. Lett. **37**, 742 (1980).
[8] C. H. Seager and M. L. Knotek, *Proc. Fifth Intern. Conf. on Amorphous Semiconductors, Garmisch-Partenkirchen, 1973* (ed. by J. Stuke and W. Brenig), Taylor and Francis, London (1974), p. 1133; A. H. Clark, M. M. Cohen, and H. P. D. Lanyon *ibid.*, p. 1185.
[9] W. Beyer and H. Mell, Solid State Commun. **38**, 891 (1981).
[10] T. Holstein, Phys. Rev. **124**, 1329 (1961).
[11] H. Bottger and V. V. Bryksin, Phys. Status Solidi B **81**, 433 (1977).
[12] P. N. Butcher and A. A. Kumar, Philos. Mag. B **42**, 201 (1980).
[13] L. Friedman and M. Pollak, Philos. Mag. B **38**, 173 (1978); **44**, 487 (1981).
[14] R. Nemeth and B. Muhlschlegel, Solid State Commun. **66**, 999 (1988).
[15] Yu. M. Gal'perin, V. L. Gurevich, and D. A. Parshin, Zh. Eksp. Teor. Fiz. **85**, 1757 (1983) [Sov. Phys. JETP **58**, 1023 (1983); M. Pollak and B. Shklovskii (eds.), *Hopping Transport in Solids*, North-Holland, Amsterdam (1990).
[16] J. P. Straley, J. Phys. C **13**, L773 (1980).
[17] E. I. Levin, Fiz. Tekh. Poluprovodn. **18**, 255 (1984) [Sov. Phys. Semicond. **18**, 158 (1984)].
[18] J. D. Joannopoulos and G. Lucovsky (eds.), *The Physics of Hydrogenated Amorphous Silicon II: Electronic and Vibrational Properties*, Springer Verlag, Berlin (1984) [Topics in Applied Physics, Vol. 56].

Translated by A. Tybulewicz