Tunnel hopping in disordered systems

V. L. Nguen, B. Z. Spivak, and B. I. Shklovskiĭ

A. F. Ioffe Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad

(Submitted 21 May 1985)
Zh. Eksp. Teor. Fiz. **89**, 1770–1784 (November 1985)

In the region of variable-range hopping, electrons undergo tunnel hops over large distances and are scattered by other impurities. The effect of scattering on the wave function of a tunneling electron is studied by numerical simulation. The corrections to the localization length are found. They agree well with the predictions of the analytic theory. The behavior of the sign of the wave function is studied for a simple model of a binary alloy. As the fraction of scatterers with a negative scattering amplitude increases, a second-order phase transition occurs from a behavior such that the wave function retains a memory of its initial sign to a behavior such that the sign of the wave function at large distances becomes unpredictable. Aharonov-Bohm oscillations in the hopping conductivity are examined. The resistance may fluctuate by an order of magnitude. The sign transition observed in this study should be seen as a change from an oscillation with a normal flux quantum to an oscillation with a superconducting quantum. A new mechanism for a negative magnetoresistance is found in the region of variable-range hopping.

1. INTRODUCTION

A hopping conductivity with a variable hopping range (variable-range hopping) is a general mechanism for the low-temperature electrical conductivity of disordered systems with localized states, e.g., doped semiconductors.

In the variable-range hopping regime, electrons hop between donors which have energies close to the Fermi level and which are separated from each other by a distance r much greater than the average distance between donors, $N^{-1/3}$ (N is the donor concentration; for definiteness, we assume a lightly doped, compensated, n-type semiconductor). The probability for a tunnel hop between donors 1 and 2 with energies ε_1 and ε_2 is proportional to $I^2/(\varepsilon_1-\varepsilon_2)^2$, if I, the resonant overlap integral, determined by the repulsion of the levels of these donors, satisfies $I \ll |\varepsilon_1-\varepsilon_2|$. If there were no other donors between donors 1 and 2, the integral I would be of the form $I_0 = I_0^{(0)} e^{r/a}$, where $a = \hbar/(2m|\varepsilon|)^{1/2}$ is the localization length of the wave function of an isolated donor with energy ε, and $I_0^{(0)}$ is the coefficient of the exponential function. Actually, in the course of tunneling between donors 1 and 2 an electron is scattered by a large number of donors within a cigar-shaped surface of revolution of length r and diameter $(ra)^{1/2}$. As a result of the scattering by donor i, a scattered wave

$$\psi_{\text{inc}}(\mathbf{r}_i)\frac{\mu}{4\pi|\mathbf{r}-\mathbf{r}_i|}\exp\left(-\frac{|\mathbf{r}-\mathbf{r}_i|}{a}\right) \tag{1}$$

arises, where ψ_{inc} is the wave function of donor 1 at point $\mathbf{r}_i$ in the absence of donor i,

$$\mu = 8\pi a\varepsilon_i/(\varepsilon_1-\varepsilon_i) \tag{2}$$

is the amplitude for scattering by the donor potential, which we assume to be a short-range potential, and ε_i is the energy of the level of donor i. According to (2), we have $\mu>0$ at $\varepsilon_1<\varepsilon_i$ and $\mu<0$ at $\varepsilon_1>\varepsilon_i$. The integral I is formed by summing the contributions of multiple scattering along various paths along the donors of the cigar-shaped region between donor 1 and donor 2. An important point is that if the energy ε_1 is negative the backscattering which would elongate the tunneling path can be ignored.

The simplest manifestation of scattering during tunneling is an N-dependent correction to the length a. A random arrangement of impurities makes I a random quantity, and the value of this correction may depend on the nature of the averaging. For a hopping conductivity we are interested in the quantity[1,2]

$$L=\left\langle \ln\left|\frac{I}{I_0}\right|^2\right\rangle \equiv -\frac{2r}{a(N)}+\frac{2r}{a}\approx\frac{2r\Delta a}{a^2}, \tag{3}$$

where $\langle\ldots\rangle$ means an average over the positions of the scatterers. The correction $\Delta a = a(N)-a$ was calculated in Refs. 3 and 4 and was found to be different in two cases differing in the relative size $\delta I/I$ of the fluctuations: homogeneous ($\delta I/I \ll 1$) and fluctuational ($\delta I/I \gg 1$). If the scattering amplitudes of all donors are identical and equal to μ, in the homogeneous regime, which prevails at[3] $B \equiv (N\mu^2 a)^{1/2} \ll 1$, we would have

$$\Delta a = N\mu a^3/2, \tag{4}$$

while in the fluctuational regime[4] ($B \gg 1$) we would have

$$\Delta a = Ca(Na^3)^{1/3}(\ln B)^{2/3}, \tag{5}$$

where C is a numerical coefficient. Hydrogen-like impurities can be described[4] by expression (5); under the assumption $C \approx 1$, that expression can explain some of the increase in a which is observed in experiments on variable-range hopping long before the insulator-metal transition.

A topic of considerable interest is how the scattering of a tunneling electron affects the magnetoresistance in the variable-range hopping regime. It has been shown[2,5,6] that scattering changes the decay of $I(r)$ at large values of r in the direction perpendicular to the magnetic field $\mathbf{H}$. If the field is weak, so that we have $\lambda^2 = c\hbar/eH \gg ra$, then in the absence of scattering we have

$$I(r,H) \propto \exp\left(-\frac{r}{a}-\frac{r^3 a}{24\lambda^4}\right). \tag{6}$$

 0038-5646/85/111021-09$04.00

When scattering does occur, expression (6) holds only if $r^3a/24\lambda^4 \ll |\ln B|$; in the opposite limit we have

$$I(r, H) \propto \exp(-r/b), \tag{7}$$

where

$$b = a\left[1 - \left\{\frac{9}{32}\left(\frac{a}{\lambda}\right)^4 |\ln B|^2\right\}^{1/3}\right]. \tag{8}$$

Expressions (7) and (8) were found in Ref. 2 for scattering by a large number of weak scatterers ($|\mu| \ll a$). A similar picture of a transition from (6) to (7), (8), only with a replaced by $a(N)$, holds for the homogeneous regime with strong scatterers ($|\mu| \gg a$), which donors actually are in practice.[6] For the fluctuational regime, however, we have not been able to find even an estimate of $a(N) - b$ in the weak-field region. The behavior $I(r, H)$ in such fields is quite pertinent, however, since experiments on variable-range hopping are interpreted by means of equations derived with the help of (6), so that they determine the values of a and the critical indices near the insulator-metal transition.[7]

In the present paper we report a numerical simulation of tunneling with scattering in two- and three-dimensional systems. The results of simulations with $H = 0$ agree well with expressions (4) and (5) at $B \ll 1$ and $B \gg 1$, respectively; the coefficient C in (5) turns out to be 1.0 ± 0.1. Unexpected results emerge from a study of scattering with negative amplitudes. Analysis of the distribution function $f(I)$ of the quantity I shows that with increasing N a second-order phase transition occurs at $B \approx 1$. This phase transition can be summarized by saying that at small values of N a plus sign is encountered on the quantity I more frequently than a minus sign is, while at the phase transition $f(I)$ becomes an even function. In a search for an effect which is sensitive to this sign phase transition, we have settled on the Aharonov-Bohm effect, which has previously been studied only in metals. The resistance of a hollow thin-walled cylinder of a pure metal oscillates as a function of the magnetic flux Φ penetrating the cylinder with a period $\Phi_0 = ch/e$ equal to the normal flux quantum.[8] It was shown in Ref. 9 that the resistance of a cylinder made of a dirty normal metal should oscillate with a period equal to the "superconducting" flux quantum $\Phi_0/2$. Such oscillations were observed in Ref. 10. In the present paper we report the first study of Aharonov-Bohm oscillations in a material exhibiting a hopping conductivity. Our numerical simulation shows that in this case oscillations can occur with a period of either Φ_0 or $\Phi_0/2$. The transition between the two situations, as the concentration of scatterers is changed, is a second-order phase transition and is a consequence of a transition in the sign of structure of I. The results of the present study were summarized briefly in Ref. 11.

The Aharonov-Bohm effect was simulated for a sample surrounding a solenoid. To study the magnetoresistance we consider samples which are immersed entirely in a uniform magnetic field. At $B \lesssim 1$ the magnetoresistance is found to agree well with expressions (6)–(8). At $B \gg 1$ with $\mu > 0$, the function $I(r, H)$ is qualitatively the same as at $B \approx 1$, but in the presence of scatterers with $\mu < 0$ the magnetoresistance is negative in a very weak field; i.e., the quantity $\langle \ln |I|^2 \rangle$ increases with increasing field. We do not believe that this effect is a consequence of a shift of the mobility threshold in the magnetic field.[12] As we will show below, the causes of this effect are strong fluctuations of the quantity $|I|^2$ and the logarithmic nature of the averaging in (1). For low concentrations of scatterers with $\mu < 0$, the negative magnetoresistance rapidly gives way to a positive magnetoresistance, and all the results of the simulation are in good qualitative agreement with experimental data on germanium.[7] At high concentrations of scatterers with $\mu < 0$, the negative magnetoresistance becomes large, and it completely displaces the positive magnetoresistance from the weak-field region. This situation is usually not observed experimentally, possibly because of many-electron effects, which we have not considered here.

2. MODEL AND CORRECTIONS TO THE LOCALIZATION RADIUS

For the calculations we use an Anderson model with the Hamiltonian

$$\mathcal{H} = \sum_i \varepsilon_i a_i^+ a_i + \sum_{i \neq j} V_{ij} a_j^+ a_i, \tag{9}$$

where i and j are the sites of a square or simple cubic lattice, the operator a_i^+ creates an electron at site i in a state with an energy ε_i, and V_{ij} is equal to $V < 0$ for nearest neighbors and 0 otherwise. The energies of the leftmost and rightmost sites, ε_1 and ε_2, at the opposite ends of a diagonal of the square (Fig. 1) or of a body diagonal of the cube, are assumed to be approximately zero. Sites 1 and 2 represent donors with approximately equal energies, between which tunneling occurs in variable-range hopping. The distribution function of the energies of the other sites is

$$g(\varepsilon_i) = (1-x)\delta(\varepsilon_i - W) + x\delta(\varepsilon_i - W/A), \tag{10}$$

where $x \leqslant 0.5$, $W > 0$, $|A| \geqslant 1$, and $W/|A| \gg |V|$. The last two inequalities correspond to the situation deep in the insulating region or, physically, the case of a lightly doped semiconductor. We calculate the resonant integral I between sites 1 and 2:

$$I = V \sum_{\{\Gamma\}} \prod_{\{i_\Gamma\}} \left(\frac{V}{\varepsilon - \varepsilon_i}\right)\Big|_{\varepsilon=0} = V\left(\frac{|V|}{W}\right)^{kn-1} J, \tag{11}$$

where

$$J = \sum_{\{\Gamma\}} \prod_{\{i_\Gamma\}} \alpha_i, \tag{12}$$

$\{\Gamma\}$ is the set of oriented paths from site 1 and to site 2 (one such path is shown in Fig. 1); $\{i_\Gamma\}$ are the sites on this path other than sites 1 and 2; $\alpha_i = 1, A$ for $\varepsilon_i = W, W/A$; $2k = 4$ or 6 is the number of nearest neighbors in the corresponding lattice; and $n = 1$ is the number of sites on the side of the square or the edge of the cube. We ignore paths with returns, since they contain additional powers of the small parameters $A|V|/W$ and $|V|/W$ and contribute only small corrections to Δa.

The quantity I can be calculated not only at site 2 but also at an arbitrary site i. For the arbitrary case, we should take kn in (11) to mean the minimum number of steps lead-

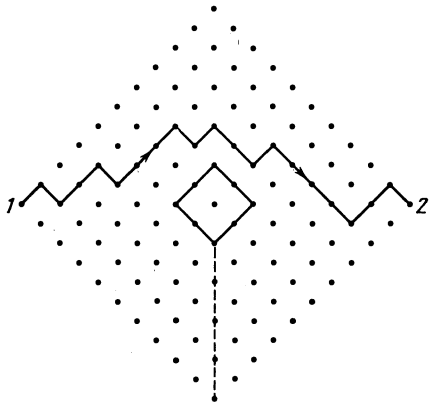

FIG. 1. Lattice used in the simulation of the quantity J. Sites 1 and 2 are the source and observation point. The arrow shows one of the oriented paths between these sites. The square at the center of the lattice is used in Section 3 to simulate an aperture in which a solenoid is placed. The dashed line is a "cut" on which the phase changes discontinuously.

ing from site 1 and site i. Let us examine the case $x = 0$, in which we have $\alpha_i = 1$ for all i; i.e., there is no scattering. In this case, J is equal to the total number of oriented paths from site 1 to site i, and it can easily be evaluated by a combinatorial approach. For definiteness, we focus on the three-dimensional case. For site i with coordinates (z, ρ, φ) in a cylindrical coordinate system with origin at point 1 and with polar axis z directed to point 2, under the condition $\rho \ll z$, we find

$$I_0(\rho, z) = \frac{6Vl}{4\pi z} \exp\left(-\frac{z}{a_1} - \frac{\rho^2}{2za_2} \right), \tag{13}$$

where

$$a_1 = l\left(\sqrt{3} \ln \frac{W}{3|V|} \right)^{-1}, \quad a_2 = \frac{l}{\sqrt{3}}, \tag{14}$$

l is the lattice constant, and the subscript 0 corresponds to an ordered lattice.

The decay of I near the z axis is thus anisotropic. Since the theoretical expresssions (4) and (5) describe the isotropic problem, it is convenient in comparing with these expressions to transform to the coordinate system $r' = (\rho', z)$, in which the functional dependence of I on ρ' and z along the z axis is isotropic (an "I-system"). This transformation is described by $\rho'^2 = \rho^2 a_1/a_2$; near the z axis, it puts (13) in the form

$$I_0(\rho', z) = \frac{6Vl}{4\pi z} \exp\left[-\frac{(z^2+\rho'^2)^{1/2}}{a_1} \right], \tag{15}$$

as in isotropic systems. The following Green's function can be expressed in terms of I_0:

$$G_{1i}|_{\varepsilon=0} = \frac{I_0}{(\varepsilon-\varepsilon_1)(\varepsilon-\varepsilon_i)}\Big|_{\varepsilon=0} = \frac{I_0}{\varepsilon_1 \varepsilon_i}. \tag{16}$$

According to (15), the Green's function near the z axis in an I-system has the usual form for an ordered system at energies in the tunneling region. Up to this point we have been describing the case in which we have $\alpha_i = 1$ for all i. Let us assume that at the point i in a cubic lattice there is a scatterer with $\alpha_i = A \neq 1$. The Green's function can then be written as the sum of direct and scattered waves:

$$G_{12} = G_{12}^0 + G_{1i}^0 \frac{\mu}{4\pi|z-z_i|} \exp\left(-\frac{|r'-r_i|}{a_1} \right), \tag{17}$$

where the scattering amplitude μ is

$$\mu = 6(A-1)l. \tag{18}$$

We have calculated J numerically for finite values of x and various values of A, using a grid with $n \leqslant 100$ for the two-dimensional case and a grid with $n \leqslant 24$ for the three-dimensional case. Our approach is to calculate J_{1i} in succession for all sites beginning with site 1. For this purpose we use the expression $J_{1i} = \Sigma_k \alpha_k J_{1k}$, which follows from (12); here k is the index of the nearest neighbors on the left of site i. For each specified pair of x and A values, we studied up to 2000 realizations of the aggregate $\{\alpha_i\}$, and we calculated the distribution function $f(J)$ and the averages $\langle \ln|J|^2 \rangle$, $\langle J \rangle$, $\langle J^2 \rangle$ over the realizations.

Figure 2 shows L as a function of x according to calculations from (3) for the three-dimensional case with $A = 2$, 10, 20, -1, -8, and -18. For comparison with (4) and (5), here are the equivalent theoretical expressions:

$$L(x) = zN_1\mu a_1 = \frac{2z}{a_2} x(A-1), \quad B \ll 1, \tag{19}$$

$$L(x) = 2Cz(N_1a_1)^{1/2} \ln^{3/2} B = 2C \frac{z}{a} x^{1/2} 3^{-3/4} \ln^{3/2} B, \quad B \gg 1, \tag{20}$$

where

$$B = (N_1\mu^2 a_1)^{1/2} = 6 \cdot 3^{-1/4} (A-1) x^{1/2}. \tag{21}$$

Expressions (19)–(21) were origially written in the I-system, in which the concentration of scatterers is denoted by N_1. We then transform to the laboratory system, using the obvious replacement $N_1 a_1 = xl^{-3} a_2$. As a result, the equations and applicability limits take a form convenient for comparison with the results of the simulation. It can be seen from Fig. 2 that in the case $B \ll 1$ all the results agree with (19). At $B \gg 1$, the results for $A = 10$ and 20 are described well by (20) with $C = 1.0 \pm 0.1$. For $A = -8$ and -18, expressions (21) and (20) give rise to the same values of $L(x)$ as are found with $A = 10$ and 20. It can be seen from Fig. 2 that, in agreement with this prediction, the experimen-

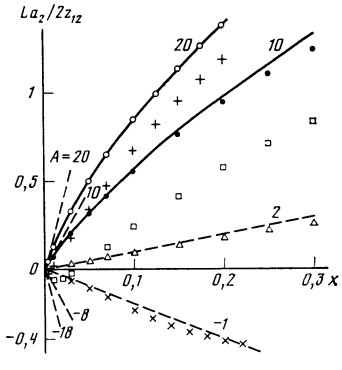

FIG. 2. The quantity $La_2/2z_{12}$ as a function of x for various values of the parameter A ($z_{12} = 24\sqrt{3}l$). The points show the results of simulations for the following values of A: ×—$A = -1$; △—$A = 2$; □—$A = -8$; ●—$A = 10$; +—$A = -18$; ○—$A = 20$. The dashed lines correspond to (19) and the homogeneous case; the solid lines correspond to the fluctuational case, i.e., expression (20) with $C = 1$.

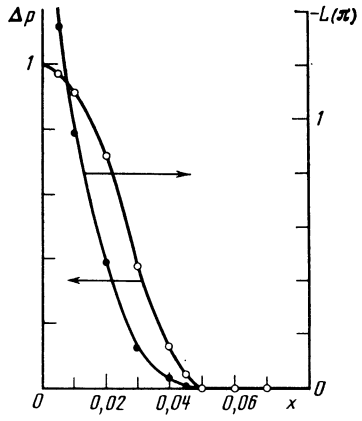

FIG. 3. The x dependence of $\Delta p = p_+ - p_-$ and of the quantity $L(\pi)$, defined in (25), demonstrating a sign transition at $x = 0.05$ (two-dimensional case, $A = -1$).

tal points for $A = -18$ and 20 at $B \gg 1$ are approximately the same. The same tendency is seen, less obviously, for $A = -8$ and 10.

3. SIGN PHASE TRANSITION AND AHARONOV-BOHM OSCILLATIONS

In the preceding section we analyzed results for $\langle \ln|J|^2 \rangle$. We begin this section with an analysis of the distribution function J. We first consider results obtained in the two-dimensional case for $A = -1$, in which the α_i take on the values ± 1. Analysis of 2000 realizations of the aggregate $\{\alpha_i\}$ shows that the nature of the distribution function J varies with x. At small x we have $J > 0$ in most of the realizations, while at $x \geqslant x_c = 0.05$ the probabilities for positive (p_+) and negative (p_-) values of J become equal. Figure 3 shows $\Delta p = p_+ p_-$ as a function of x for $n = 100$. The function $\Delta p(x)$ does not depend on n for $n > 20$; i.e., it corresponds to the limit $n \to \infty$. The nature of this function indicates the occurrence of a second-order phase transition. Accordingly, at $x < x_c$, in the limit $n \to \infty$, we can predict the sign of the Green's function $G_{12}(\varepsilon = 0) = I/\varepsilon_1 \varepsilon_2$ with some confidence, while at $x > x_c$ the two signs are equally probable. Figure 4 shows the distribution functions of the difference $\ln|J| - \langle \ln|J| \rangle$ separately for $J > 0$ and $J < 0$. At $x = 0.03$, these two peaks in the distribution function are very different, while at $x = 0.05$ the difference disappears. At $x > x_c$, the distribution function $f(J)$ of J is even. However, to say that the distribution function $f(J)$ is an even function does not in this case imply $\langle J \rangle = 0$. The quantity $\langle J \rangle$ can be calculated exactly; it is

$$\langle J \rangle = 2^{2n}(1-2x)^{2n}. \qquad (22)$$

The first factor in this expression is the total number of paths from point 1 to point 2, while the second is the contribution of one path, averaged over the realizations. We see that we have $\langle J \rangle > 0$ up to $x = 0.5$ and that this average has no structural feature of any sort at the transition point $x_c = 0.05$. We believe that because of the exponentially large scatter in the values of J from realization to realization the value of $\langle J \rangle$ at $x > x_c$ is determined by exponentially rare but very large positive values of J. In the limit $n \to \infty$ the contribution of these values to the normalization of the distribution function exponentially vanishes. It is in this sense that the distribution function is even at $x > x_c$. Support for this interpretation comes from the behavior of the theoretical value $\langle J \rangle_{\text{theo}}$. While at $x < x_c$ this theoretical value agrees well with expression (22) with $x > x_c$, it has a random sign and a random magnitude. The difference $\ln|\langle J_{\text{theo}} \rangle| - \ln\langle J \rangle$, where $\langle J \rangle$ is given by (22), is proportional to $(x - x_c)$. We believe that the reason for the discrepancy between $\langle J_{\text{theo}} \rangle$ and $\langle J \rangle$ is that the number of realizations studied is not adequate to "capture" those rare realizations which determine (22). We have also calculated the value of $\langle J^2 \rangle$. We found no features of any sort in this quantity at $x = x_c$; this was to be expected, since this quantity is also determined by very rare events. We believe that the same factor means that an analytic calculation of $\langle J^2 \rangle$ can provide no information about the sign transition. Less understandable is the result that there is no noticeable structural feature at $x = x_c$ in the values from the numerical simulation for the quantity $\langle \ln|J|^2 \rangle$, which is determined by typical realizations.

A similar sign transition is observed in the three-dimensional case. With $A = -1$, -4, and -8 we find $x_c = 0.135$, 0.025, and 0.01, respectively. All the qualitative results in the three-dimensional case are the same as in the two-dimensional case.

Can we explain the origin of the sign transition on the basis of expression (1)? Let us assume that in the three-dimensional case ψ_{inc} is a plane wave $e^{-r/a}$ which is incident from the half-space $z < 0$, and a donor 3 with $\mu < 0$ is at the origin of our cylindrical coordinate system (ρ, φ, z) (Fig. 5). Beyond this donor a region with $\psi < 0$ arises. The surface bounding this region is described by

$$\exp\left(-\frac{z}{a}\right) + \frac{\mu}{4\pi z}\exp\left[-\frac{(z^2+\rho^2)^{1/2}}{a}\right] = 0 \qquad (23)$$

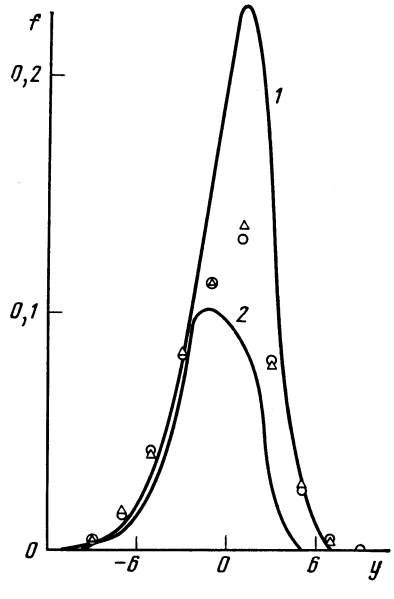

FIG. 4. Distribution functions $f(y)$ of the quantity $y = \ln|J| - \langle \ln|J| \rangle$, plotted separately for $J < 0$ and $J > 0$. Solid lines: $x = 0.03$. 1—$J > 0$; 2—$J < 0$. Points: $x = 0.05$. Δ—$J > 0$; ○—$J < 0$ (two-dimensional case, $A = -1$).

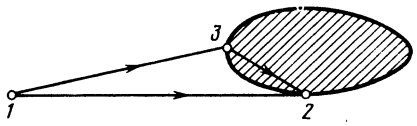

FIG. 5. Region beyond a scatterer with $\mu<0$ (donor 3) in which the condition $\psi\leqslant 0$ holds (hatched region).

and is cigar-shaped with a length on the order of $|\mu|$, a diameter on the order of $(|\mu|a)^{1/2}$, and a volume $\Omega=\mu^2 a/32\pi$. The fraction of the volume which is occupied by all regions with $\mu<0$ is $N\Omega=B^2/32\pi$, where N is the concentration of scatterers. When $B\ll 1$, in which case this fraction is small, a positive sign is clearly more probable than a negative sign for the incident wave function at large distances. It seems likely that the transition to an alternating-sign situation should occur at some critical value of $N\Omega$ on the order of unity, i.e., at

$$B^2=N\mu^2 a\equiv B_c^2, \tag{24}$$

where B_c is a critical value. Expression (24) relates N to μ at the transition point. According to (21), it follows from (24) that in our lattice model the quantity $x_c(A-1)^2$ should be identical at the transition point for different values of A. Indeed, we find that this quantity varies only from 11.2 to 17 as A is varied from -1 to -8.

How will this transition which we have discovered be manifested in Aharonov-Bohm oscillations? For the numerical simulation of this effect in the two-dimensional case we assumed $\alpha_i=0$ for all i in a 7×7 square at the center of the lattice; i.e., we set up an impenetrable aperture where the solenoid is located (Fig. 1). The magnetic flux (Φ) through the aperture is taken into account by multiplying α_i by $e^{i\varphi}$, where $\varphi=2\pi\Phi/\Phi_0$, at all the sites along the cut running from the lower corner of the aperture (the dashed line in Fig. 1).

We calculated the quantity $L(\varphi, x)=\langle\ln|J(\varphi)/J(0)|^2\rangle$. Figure 6 shows $L(x)$ as a function of φ for $A=-1$ and various values of x. At $x<x_c$ the magnetoresistance is seen to have a period of 2π. As x is increased the quantity $|L(\pi, x)|$ decreases, vanishing at $x=x_c$ (Fig. 3). At $x>x_c$ the magnetoresistance is negative for φ and has a period of π. The transition to a phase which is disordered in terms of the sign of J thus causes the period Φ_0 to be replaced by a period $\Phi_0/2$. To clarify the relationship between these effects, we write J in the form $J=J_1+J_2$, where J_1 and J_2 are the sums along paths running above and below the solenoid. We can then write

$$L(\varphi, x)=\int_{-\infty}^{\infty}\int dJ_1\, dJ_2 F(J_1, J_2)\ln\left|\frac{J_1+J_2 e^{i\varphi}}{J_1+J_2}\right|^2, \tag{25}$$

where $F(J_1, J_2)$ is the distribution function of the values of J_1 and J_2. At $x=0$ we have

$$F(J_1, J_2)=\delta(J_1-J/2)\,\delta(J_2-J/2) \text{ and } L(\varphi, 0)$$
$$=\ln[(1+\cos\varphi)/2]\leqslant 0,$$

in accordance with Fig. 6. As x is increased, the function $F(J_1, J_2)$ becomes progressively more nearly symmetric in terms of the sign of each variable. At the point $x=x_c$ the function $F(J_1, J_2)$ takes the form $F(J_1^2, J_2^2)$. From (25) we then find

$$L(\varphi, x)=2\int_0^{\infty}\int_0^{\infty} dJ_1\, dJ_2 F(J_1^2, J_2^2)$$
$$\times\ln\left[1+\frac{2J_1^2 J_2^2}{(J_1^2-J_2^2)^2}(1-\cos 2\varphi)\right]>0. \tag{26}$$

According to (26), at $x>x_c$ the oscillation period is π, and the magnetoresistance is negative for all φ. Clearly, the period π arises in (25) because when we add π to φ the phase factor $e^{i\varphi}$ changes sign; the effect is equivalent to a change in the sign of J_2. Since the distribution in J_2 is of even parity, however, there is no change in $L(\varphi, x)$.

The reason for the negative sign of the magnetoresistance can be seen most easily at $\varphi=\pi/2$, where a difference $\langle\ln(J_1^2+J_2^2)\rangle-\langle\ln(J_1+J_2)^2\rangle$ arises in (25). This difference is positive, so that taking an average of the logarithm always emphasizes small values of its argument, and when the magnetic flux is absent the argument varies far more than when a flux is present.

It can be seen from Fig. 6 that at small values of φ a negative magnetoresistance prevails even for $x<x_c$. To find the reason for this result, we rewrite (25), expanding in $\varphi\ll 1$ and introducing the variable $J=J_1+J_2$:

$$L(\varphi, x)=\int_{-\infty}^{\infty} dJ_1\int_{-\infty}^{\infty} dJ F(J_1, J-J_1)\ln\left[1+J_1(J_1-J)\frac{\varphi^2}{J^2}\right]. \tag{27}$$

At $\varphi\ll 1$, expression (27) is dominated by values $J\leqslant|\varphi||J_1|\ll|J_1|$. If $F(J_1, -J_2)\neq 0$, we find from (27) that at $|\varphi|\ll 1$ we have

$$L(\varphi, x)=4\pi|\varphi|\int_{-\infty}^{\infty} dJ_1|J_1|F(J_1, -J_1)>0. \tag{28}$$

Clearly, the negative magnetoresistance which we have found does not arise because localized corrections to the conductivity which result from traversing closed loops in two directions are suppressed,[9,13] since our simulation com-

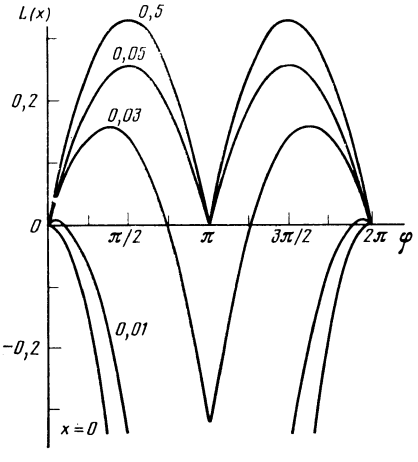

FIG. 6. The quantity $L(x)$, defined in (25), versus the phase φ for various values of x (the curve labels; two-dimensional case; $A=-1$).

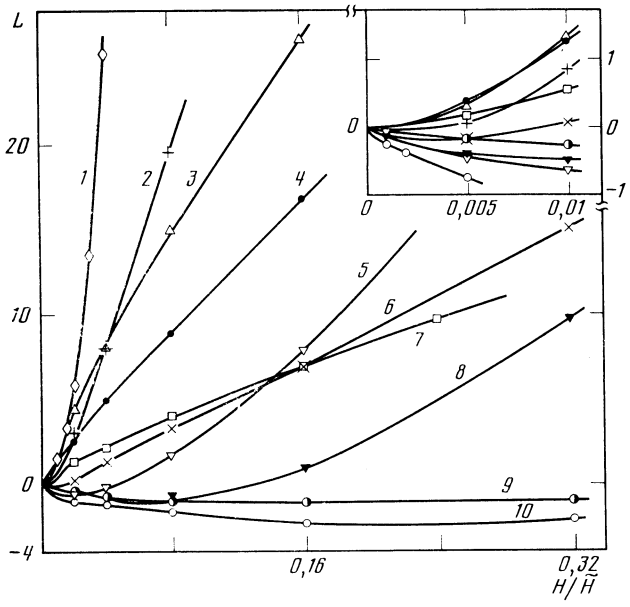

FIG. 7. The quantity $L_T + \langle \ln|J(H)/J(0)|^2\rangle$ as a function of the dimensionless magnetic field $H/\tilde{H}$ for the three-dimensional case and various values of A and x: 1—$x=0$; 2—$A=8$, $x=0.001$; 3—$A=2$, $x=0.5$; 4—$A=20$, $x=0.01$; 5—$A=-8$, $x=0.01$; 6—$A=-1$, $x=0.1$; 7—$A=20$, $x=0.1$; 8—$A=-8$, $x=0.02$; 9—$A=8$, $x=0.1$; 10—$A=-1$, $x=0.5$. Shown separately in the inset is the region of very weak fields, $H\leqslant 0.01\tilde{H}$.

pletely ignores these loops. It is interesting to see what expression (26) predicts if we arbitrarily assume $F(J_1, J_2) = f(J_1)f(J_2)$, where $f(t) = (\gamma\sqrt{\pi})^{-1}e^{-t^2/\gamma^2}$ is a Gaussian distribution function. In this case we find the result

$$L(\varphi, x) = \ln[1+(1-\cos\varphi)^{1/2}], \tag{29}$$

which is roughly twice as large as the experimental values of $L(\varphi, 0.5)$.

Similar numerical simulations were carried out in the three-dimensional case with $A=-1$. It was assumed that a cylinder was cut from a cube in the direction perpendicular to the body diagonal connecting sites 1 and 2. To simulate a solenoid placed in this cylinder, we multiply the contributions from all paths which pass below the cylinder by $e^{i\varphi}$. It is found that, as in two dimensions, the oscillation period changes and the positive magnetoresistance disappears completely where the sign changes. Curiously, at $x=0.5$ the theoretical function $L(\varphi, 0.5)$ is very close to (29), telling us that at $x=0.5$ the distribution function of J in three dimensions is described considerably better by a Gaussian distribution than in two dimensions.

4. SIMULATION OF THE MAGNETORESISTANCE

In this section we report the results of a simulation of the effect of a magnetic field on tunnel hopping. We consider only the three-dimensional case. In contrast with Section 3, we simulated a situation in which an entire $25\times25\times25$ cube is immersed in a uniform magnetic field directed perpendicular to the vector $\mathbf{r}_{12}$. There are no apertures of any sort inside the cube, of course. In a magnetic field H, the quantities V_{ij} in (9) acquire a phase factor

$$V_{ij} = V\exp(i\varphi_{ij}), \quad \varphi_{ij} = (e/2\hbar c)\mathbf{H}[\mathbf{r}_i, \mathbf{r}_j]. \tag{30}$$

The change in $|V_{ij}|$ in fields $H\ll\hbar c/el^2$ can be ignored. As a result, we can replace expression (12) for J by

$$J = \sum_{\{\Gamma\}}\exp(i\varphi_\Gamma)\prod_{\{i\Gamma\}}\alpha_i, \tag{31}$$

where the α_i have the same meaning as in (12), and φ_Γ is the sum of the phases φ_{ij} corresponding to all steps of the path Γ. The values of J are calculated in succession, beginning at the left. A new value of J_{1j} is found in the formula

$$J_{1j} = \sum_i J_{1i}\alpha_i\exp(i\varphi_{ij}),$$

where i runs over the three nearest neighbors on the left of site j. On the diagonal connecting points 1 and 2 we calculate the quantity $L(H, r) = \langle\ln|J(H)/J(0)|^2\rangle$, where r is the distance to point 1. Figure 7 shows results on $L_T(H)\equiv L(H, r_{12})$ for several values of the parameters A and x. The magnetic field is expressed in units of $\tilde{H} = 2c\hbar/el^2$, where l is the lattice constant.

At $x=0$, i.e., in the absence of scattering, the results of the simulation for $L(H, r)$ are in excellent agreement with (6); the latter predicts the following for $L(H, r)$:

$$L(H, r) \equiv -\frac{r^3a}{12\lambda^4} = -\frac{r^3ae^2H^2}{12c^2\hbar^2}. \tag{32}$$

The meaning of a is none other than the quantity $a_2 = l/\sqrt{3}$ defined in (13), (14). Curve 1 in Fig. 7 is in fact a parabola $L(H) = \beta(H/\tilde{H})^2$, and to compare the coefficient β with (32) we should note that we have $r_{12} = 24\sqrt{3}l$.

We begin our study of the effect of scattering with the case $A=2$, $x=0.5$ (curve 3). Noting that the average value of α_i in this case is 1.5, and the deviations from the average are ± 0.5, we can expect that this case is still within or at the boundary of the range of applicability of the weak-scattering theory.[2] As we mentioned above, this theory leads to (32) in the case $L(H)\ll\ln B$, i.e., as long as scattering does not play an important role. In the case at hand we have $\ln B\approx 1$, so that we should expect (32) to hold, i.e., that curves 3 and 1 will coincide at $L(H)\leqslant 1$. In fact, we find that these curves are very close together at $L<3$. As the field increases further, the scattering becomes important, and $|L(H)|$ increases in proportion to $H^{2/3}$, in agreement with (7) and (8). Simultaneously, the behavior changes from $L(H, r)\propto r^3$ to $L(H, r)\propto r$. Furthermore, we have not only a qualitative but a good quantitative agreement between curve 2 and expressions (7) and (8).

We now turn to the case of strong scatterers ($A\gg 1$). Figure 7 shows results for $A=20$ and for $x=0.01$ and 0.1 (curves 4 and 7). We see from Fig. 2 that at $x=0.01$ the system is at the beginning of the fluctuational region, while at $x=0.1$ it is deep inside it. For the fluctuational case we have not been able to construct an analytic theory for magnetoresistance in weak fields, with $a(N) - b\ll\Delta a$ [in stronger fields, with $a(N) - b\gg\Delta a$, the theory of Refs. 5 and 6 should hold for arbitrary B]. It can be seen from Fig. 7, however, that the behavior $L(H)$ is qualitatively the same as in the homogeneous case. In extremely weak fields, with

$L(H) = \beta(H/\tilde{H})^2$, and with $x = 0.01$, the coefficient is the same as in the absence of scattering, while at $x = 0.1$ it is smaller by a factor of about two. At $L(H) > 1$ we have $L(H) = qH^s$ in each case, where $s = 0.8$–0.9, and the coefficient q, like β, falls off slowly (apparently, logarithmically) with increasing x. Here we have $L(H, r) \propto r^3$ in the region $L(H, r) < 1$ and $L(H, r) \propto r$ in the region $L(H, r) > 1$. Up to this point we have been discussing the results for the case $A > 0$, in which the contributions from all paths to J are positive, and the magnetic field can only cause cancellation of the contributions of the individual paths, because of the factor $e^{i\varphi}$, so that a positive magnetoresistance will be established. This reduction will be most effective in the absence of a scattering, and it will become far weaker in the fluctuational case, where the contributions of the different paths are very different from each other.

How does this situation change in the presence of scatterers with $A < 0$? Figure 7 shows results for the cases $A = -1$ (curves 6 and 10) and -8 (curves 2, 5, 8, and 9) for various values of x. In both cases we see a new effect: an effective increase in J in a magnetic field, corresponding to a negative magnetoresistance. At small values of x the negative magnetoresistance is found only in weak fields, where it is linear in H and thus greater than the positive magnetoresistance. We believe that this negative magnetoresistance is of the same nature as the Aharonov-Bohm effect. In those rare realizations in which, with $H = 0$, cancellation at the point r of the contributions of paths coming in from the left is anomalously strong because of scatterers with $A < 0$, even a small phase difference φ between these paths, caused by the magnetic field, will eliminate this cancellation. The logarithmic averaging emphasizes the role of these events and gives rise to a result analogous to (28). To estimate the phase φ, which plays the role of the phase set by the solenoid in the Aharonov-Bohm effect, we return to the case in Fig. 5, where in a continuous space there is a single scatterer with a negative amplitude μ, on which a wave is incident from point source 1. Let us assume that at observation point 2 the wave function is approximately zero because of the direct wave and that scattered by donor 3 cancel. In a magnetic field directed perpendicular to the plane of the paper, a phase difference $\varphi \approx S/\lambda^2$ arises between these waves, where $S \approx r(|\mu|a)^{1/2}$ is the area of the triangle with vertices at the points 1, 2, and 3; and $r = r_{12}$. When the concentrations N of scatterers with $\mu < 0$ and the values of φ are small, an average over realizations analogous to (27) leads to

$$L(H, r) \approx N\mu^2 a|\varphi| \approx N\mu^2 a(|\mu|a)^{1/2} rH. \quad (33)$$

This estimate of the negative magnetoresistance appears to be good for all N in the homogeneous case, since in this case other scatterers only slightly distort the spherical wave coming from point 1, simply making a correction to the localization radius. We cannot evaluate the negative magnetoresistance in the fluctuational case, but in the numerical simulations a linear dependence of L on H and r at small H is observed in all cases. It should be kept in mind, however, that the accuracy of the negative magnetoresistance results, despite the averaging over 50 realizations, is low because of the very strong fluctuations in $\ln|J(H)/J(0)|^2$ from one realization to another. From the standpoint of our explanation of the negative magnetoresistance, based on the role of rare realizations, this seems a natural result. At $x \ll x_c$, where x_c is the critical concentration for the sign transition, the negative magnetoresistance in (33) should be summed with the positive magnetoresistance in (32). As a result, a minimum should appear in the function $L(H)$ at some field. This behavior is indeed observed at $x \ll x_c$. At $x \lesssim x_c$ (curves 5 and 6 in Fig. 7), the picture is qualitatively the same, but the positive magnetoresistance is considerably weaker than at $x \ll x_c$, while the negative magnetoresistance has managed to grow to large values. For $x \gtrsim x_c$ the positive magnetoresistance decreases sharply, remaining only in very strong fields (curve 8 in Fig. 7). This positive magnetoresistance can be attributed to the effect of the field on the wave functions in spatial regions between scatterers with $A < 0$. At sufficiently large values of x, the positive magnetoresistance completely gives way to the negative magnetoresistance over the entire region $H < \tilde{H}$ (curves 9 and 10 in Fig. 7). In strong fields, a positive magnetoresistance should of course persist, but in order to find it in (30) we would have to take into account the dependence of $|V_{ij}|$ on H.

Again, we wish to stress that the negative magnetoresistance arises from the averaging of the logarithm $|J(H)/J(0)|^2$. If we average $|J(H)/J(0)|^2$, there will be no magnetoresistance at $x = 0.5$, while at $x \neq 0.5$ there will be a positive magnetoresistance. Averaging $|J(H)/J(0)|^\alpha$ leads to a positive magnetoresistance at $\alpha > 2$ or a negative magnetoresistance at $0 < \alpha < 2$.

We believe that the negative magnetoresistance described above is related to the negative magnetoresistance observed by Lee and Fisher[14] in a numerical calculation of the conductance of a square lattice in the Anderson model. Lee and Fisher did not restrict their study to oriented paths, so that they were able to study regions of both strong and weak localization. In the strong localization region they observed a substantial increase in the average logarithm of the conductance of a 32×32 lattice with increasing magnetic field, but they did not explain this negative magnetoresistance. We believe that it is due primarily to the logarithmic averaging, not the paths with returns which were considered in Ref. 14. To check this point, we carried out some calculations analogous to those in Ref. 14, considering only oriented paths. We studied a square lattice rotated through 45° (Fig. 8), with a side containing 32 sites. We calculated the transparency of the square:

$$T = \sum_{k,l} J_{kl}^2, \quad (34)$$

where k and l specify indices on respectively the left and right sides of the square, and the j_{kl} are the overlap integrals between sites k and l. They can be found from (31) by treating site k as a "source" and site l as an "observer." Periodic boundary conditions are imposed on the upper and lower boundaries of the square. As in Ref. 14, we assumed that the energies ε_i are distributed uniformly from $-W$ to W. An average is taken over 200 realizations on the basis of

$$L_T(H) = \langle \ln(T(H)/T(0)) \rangle. \quad (35)$$

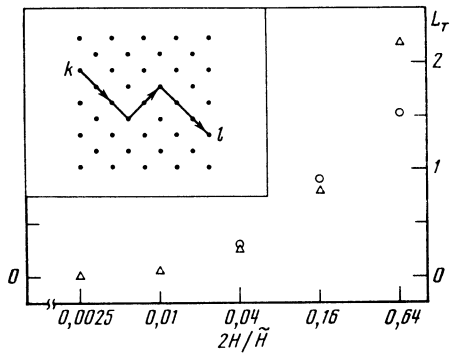

FIG. 8. Results of a calculation of the transparency of the square lattice shown in the inset. ○—Reference 16; △—Present study.

Figure 8 shows our results along with Lee and Fisher's results for the largest value of their parameter W, $W = 8$. We see that the results are similar, although the errors are very large. The implication is that logarithmic averaging plays a major role in the origin of the negative magnetoresistance. In weak fields, $L_T(H)$ is proportional to H^2, not H, as in (33). The reason is that, because of the summation of a large number of positive values in (34), the probability for extremely small values of L_T is low.

5. CONCLUSION

We have a few comments regarding the relationship between these calculations and experiments.

1. In this paper we have been concerned primarily with the scatterer energy distribution (10), which is characteristic of a binary alloy, not of doped semiconductors, which would have a continuous energy spectrum. It is thus natural to ask whether a sign transition occurs when there is a continuous energy spectrum. The answer to this question is that if the state density at the Fermi level, $g(\varepsilon_F)$, is nonzero, and if condition (24) holds, then the sign of I is not conserved in the limit $r \to \infty$ at any value of N. The reason is the special role played by scatterers with energies ε_i close to ε. Let us consider an energy interval ε_i such that $0 < \varepsilon - \varepsilon_i < \Delta$. The concentration of scatterers in it is $g(\varepsilon_F)\Delta$, and a typical amplitude is on the order of $\mu(\Delta) \approx -|\varepsilon/\Delta|$. Clearly, in the limit $\Delta \to 0$ condition (24) will hold for an arbitrarily small $g(\varepsilon_F)$. The situation changes if we take into account the Coulomb gap at the Fermi level, in which we have $g(\varepsilon) = \alpha(\varepsilon - \varepsilon_F)^2$. In this case the state density outside the Coulomb gap dominates. If this state density is low, there is a constant-sign situation; otherwise, an alternating-sign situation arises. Consequently, the presence of a Coulomb gap raises the hope that it will be possible to observe a sign change when the position of the Fermi level changes, by compensation, for example. An important question is the stability of the results obtained above with respect to the incorporation of paths with returns. As long as returns have only a slight effect on the localization radius, we see no reason why they should annihilate the sign transition. It seems natural that when the Fermi level moves upward out of the deep tail in the state density there will be a transition from a constant-sign to an alternating-sign situation at a certain energy, and only after this transition will there be a transition from localized states to delocalized states.

2. How would it be possible to observe Aharonov-Bohm oscillations in variable-range hopping? Let us first consider an experiment on a plane lattice constructed of narrow strips of the material of interest. A metal lattice of this sort was used in Ref. 15 to study Aharonov-Bohm oscillations with a charge of $2e$. The lattice was immersed in a transverse magnetic field, and oscillations of the resistance were observed when the flux penetrating a single cell of the "honeycomb" changed by $\Phi_0/2$. Imagine a lattice of a material in which the conductivity is of the nature of variable-range hopping, and the hopping length is comparable to or slightly greater than the lattice period. For each hop the value of I will then be the sum of two or several comparable contributions from sheafs of paths going around different sides of the apertures. The results derived above are thus applicable in a qualitative way to such an entity, and oscillations of significant amplitude should be observed.

Another observtion method is analogous to that of Ref. 10 and consists of a study of the resistance of a long-hollow cylinder as a function of the flux penetrating it along the axis when the hopping length is greater than the diameter of the cylinder. In this case, sheafs of paths going on different sides around the flux contribute to each hop along the cylinder. This circumstance is the reason for the oscillations. This method, however, has a distinctive feature because of the one-dimensional nature of the conductivity. In the one-dimensional case, the hopping conductivity is determined by the sparsely distributed elements of a Miller-Abrahams equivalent chain with very large resistances. In a long sample there are many such elements, and we need to take the average of the resistance of a Miller-Abrahams network. In shorter samples the resistance is determined by a single element with a maximum resistance. In this case, an averaging does not arise at all, and the resistance fluctuates markedly from sample to sample.

3. The magnetoresistance of n-type Ge in the variable-range hopping region was studied in Ref. 7. The field dependence of the resistance is qualitatively reminiscent of the curves found in our numerical simulations for $x \lesssim x_c$, e. g., curve 6. In weak fields there is a negative magnetoresistance, in agreement with (33), and it is linear in H and r. The magnetoresistance then goes through a minimum and becomes positive and proportional to H^2. In stronger fields, the H dependence of the positive magnetoresistance weakens. These points of agreement suggest that a constant-sign situation prevailed in the samples under study. On the other hand, we are talking about samples which are not very far from a metal-insulator transition, where it would appear that an alternating-sign situation should prevail. We do not rule out the possibility that many-electron effects, which we have not considered here, disrupt the sign transition. Since a tunneling electron has an energy near the Fermi level, donors with low energies are filled. Although it would appear at first glance that this circumstance would not greatly change the situation,[4] it may be that spin flip electrons are scattered by filled donors cause the scattering to be incoherent and sharply reduce the role played by scatterers with

$A < 0$ (Ref. 16).

We wish to thank B. L. Al'tshuler, A. I. Aronov, E. I. Levin, S. P. Obukhov, M. É. Raĭkh, D. E. Khmel'nitskiĭ, D. B. Shklovskiĭ, and A. L. Éfros for useful advice and discussions.

[1]B. I. Shklovskiĭ and A. L. Éfros, Élektronnye svoĭstva legirovannykh poluprovodnikov (Electronic Properties of Doped Semiconductors), Nauka, Moscow, 1979.
[2]B. I. Shklovskiĭ and A. L. Éfros, Zh. Eksp. Teor. Fiz. **84**, 811 (1983) [Sov. Phys. JETP **57**, 470 (1983)].
[3]I. M. Lifshits and V. Ya. Kirpichenko, Zh. Eksp. Teor. Fiz. **77**, 989 (1979) [Sov. Phys. JETP **50**, 499 (1979)].
[4]B. I. Shklovskiĭ and B. Z. Spivak, J. Stat. Phys. **38**, 267 (1985); B. Z. Spivak, V. A. Kharchenko, and B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **19**, 799 (1985) [Sov. Phys. Semicond. **19**, 493 (1985)].
[5]B. I. Shklovskiĭ, Pis'ma Zh. Eksp. Teor. Fiz. **36**, 43 (1982) [JETP Lett. **36**, 51 (1982)].
[6]B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **17**, 2055 (1983) [Sov. Phys. Semicond. **17**, 1311 (1983)].
[7]I. S. Shlimak, A. N. Ionov, and B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **17**, 503 (1983) [Sov. Phys. Semicond. **17**, 314 (1983)]; A. N. Ionov, I. S. Shlimak, and M. N. Matveev, Solid State Commun. **47**, 763 (1983).
[8]I. O. Kulik, Pis'ma Zh. Eksp. Teor. Fiz. **5**, 423 (1967) [JETP Lett. **5**, 345 (1967)]; Zh. Eksp. Teor. Fiz. **63**, 1838 (1972) [*sic*].
[9]B. L. Al'tshuler, A. G. Aronov, and B. Z. Spivak, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 101 (1981) [JETP Lett. **33**, 94 (1981)].
[10]D. Yu. Sharvin and Yu. V. Sharvin, Pis'ma Zh. Eksp. Teor. Fiz. **34**, 285 (1981) [JETP Lett. **34**, 272 (1981)].
[11]V. L. Nguen, B. Z. Spivak, and B. I. Shklovskiĭ, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 35 (1985) [JETP Lett. **41**, 42 (1985)].
[12]B. L. Al'tshuler, A. G. Aronov, and D. E. Khmel'nitskiĭ, Pis'ma Zh. Eksp. Teor. Fiz. **36**, 157 (1982) [JETP Lett. **36**, 195 (1982)].
[13]B. L. Al'tshuler, A. G. Aronov, A. I. Larkin, and D. E. Khmel'nitskiĭ, Zh. Eksp. Teor. Fiz. **81**, 768 (1981) [Sov. Phys. JETP **54**, 411 (1981)].
[14]P. A. Lee and D. S. Fisher, Phys. Rev. Lett. **47**, 882 (1981).
[15]B. Pannetier, J. Chaussy, R. Rammal, and P. Gandit, Phys. Rev. Lett. **53**, 718 (1984).
[16]B. Z. Spivak, Zh. Eksp. Teor. Fiz. **87**, 1371 (1984) [Sov. Phys. JETP **60**, 787 (1984)].

Translated by Dave Parsons