Anomalous polarization of Raman scattering spectra from porous silicon

M. E. Kompan,[a] V. B. Kulik, and I. I. Novak
A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

J. Salonen
Turku University, FIN-20500, Finland

A. V. Subashiev
State Technical University, 194021 St. Petersburg, Russia

(Submitted 1 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 95–100 (25 January 1998)

A violation of the polarization selection rules for Raman scattering is observed in porous silicon. This effect is caused by a weak disorientation of the quasi-one-dimensional silicon wires, with the crystal structure of the wires themselves and the macroscopic homogeneity of the material in optical experiments remaining intact. © *1998 American Institute of Physics.* [S0021-3640(98)00202-3]

PACS numbers: 78.30.Am, 78.55.Mb

1. The initial model proposed in Ref. 1 for porous silicon (PSi) explains a number of unusual properties of this material, but facts which remain unexplained in the model of Ref. 1 are accumulating. Specifically, it is not understood why there is no correlation between the spectral position of the luminescence band and the magnitude of the shift of the characteristic band in the Raman scattering spectrum to lower energies, a fact which has been noted in a number of papers (see, e.g., Ref. 2).[3] These quantities should both be determined by the same factor — the smallness of the transverse cross section of the ''quantum wires.''

There is an obvious possible source of this conflict — a model explaining the phenomena occurring in individual wires cannot be adequate for explaining the properties of the macroscopic material. Specifically, it has been shown[4] that the aforementioned characteristic feature of the Raman scattering spectrum from porous silicon may not be observed and may even have the opposite sign if the structure of the macromaterial consisting of quantum wires has a certain type of structure.

In the present work we observed a characteristic feature of the polarization of Raman-scattered light from porous silicon. This feature was observed for all available types of samples (with one possible exception, see below). This attests to the fact that the causes of the indicated anomaly are of a fundamental nature.

The relative intensities of the differently polarized Raman components are determined by the fundamental selection rules for Raman scattering.[5] The existing approach to

the analysis of Raman scattering spectra in quantum-size objects is based on Ref. 6, where only the shift of the fundamental scattering band is studied. However, it can be expected *a priori* that a material consisting of sharply anisotropic objects will possess nontrivial polarization properties.

2. Raman scattering was investigated with a Spex-Ramalog 5 spectrophotometer with a dispersion of 3 cm^{-1}/mm and computerized recording of the spectra; the experiments were conducted at room temperature. Most experiments were performed with excitation by 50-mW 6328 Å light from a He–Ne laser. Argon laser radiation with a wavelength of 488 nm was used in some control experiments. The scattered light was detected in the backscattering geometry separately in XX and XY polarizations (for light propagating along the Z axis; the notation of Ref. 5 is used).

In addition, the luminescence was monitored for all PSi samples. Under short-wavelength excitation all the experimental samples exhibited bright luminescence in the orange-red region of the spectrum and a characteristic microsecond decay of the luminescence.

The experiments were performed on porous silicon samples produced by different technologies. Different types of initial silicon, production technology, and variants of surface treatment were represented in the experimental samples.

3. In experiments with PSi samples grown on [001] silicon it was found that the intensities of the Raman spectra for two polarizations of the scattered light — parallel and perpendicular to that of the exciting light (XX and XY) — differed by less than an order of magnitude for the case when one of the cubic axes of the material lay in the plane of polarization of the scattered light. According to the selection rules, in this case the intensity of the scattered light polarized in the same plane as the exciting light should equal zero.[5] To obtain more-detailed information we investigated the dependence of the degree of linear polarization of the Raman radiation (at the maximum of the band in the spectrum) on the angle α between the plane of polarization and the $\langle 100 \rangle$ axis in the sample. The degree of linear polarization ρ was determined in the standard way from measurements of the intensities of individual components: $\rho=(I_{xx}-I_{xy})/(I_{xx}+I_{xy})$.

The types of angular dependences $\rho(\alpha)$ obtained are shown in Fig. 1. The plot in Fig. 1a demonstrates the above-indicated dependence for the [001] cut of single-crystalline silicon. We note that four of the eight lobes of the pattern have a positive degree of polarization, while the four lobes in between them have a negative degree of polarization. This type of relation corresponds to the fundamental selection rules and is described by the well-known relation[5]

$$\rho = -\cos 4\alpha, \tag{1}$$

where α is the angle between the light vector **E** and the $\langle 100 \rangle$ axis.

For most of the porous silicon samples the pattern had the form shown in Fig. 1b and 1c. Four lobes are clearly distinguishable in these plots, while the four lobes in between are absent or strongly suppressed. It is obvious that such a dependence of the polarization properties, together with a change in the orientational dependence, does not reduce to a trivial depolarization of the Raman-scattered light.

The orientational dependence $\rho(\alpha)$ for a PSi sample obtained from a film of amorphous silicon a-Si differed from those described above. This sample also exhibits bright

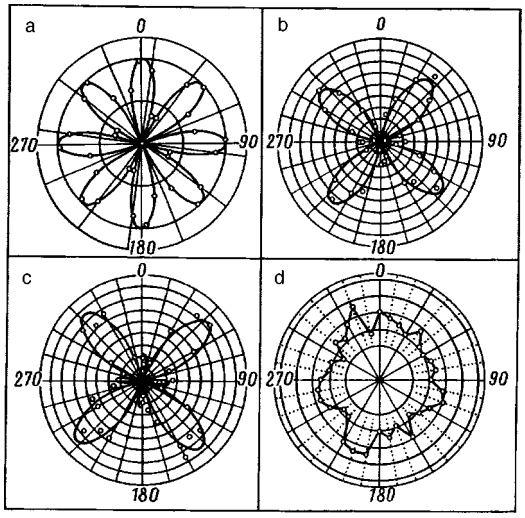

FIG. 1. Degree of linear polarization of Raman-scattered light versus the angle between the [100] axis and the plane of oscillations of the electric field vector of the light: a — For a polished sample of single-crystalline silicon, b — for a mirror film of porous silicon on a [100] surface of a single-crystalline substrate; c — same for a film sample separated from the substrate; d — for a porous-silicon film formed from an *a*-Si film. The parameters (A,B) of interpolation formulas of the type (2) for the indicated dependences equal: a) 0, 1; b) 0.15, 0.19; c) 0.27, 0.51.

orange luminescence, in accordance with the main experimental criterion for the presence of PSi. However, the dependence of the degree of polarization on the rotation angle of the sample in this case is not a regular rosette, and the degree of polarization is small, between 0.1 and 0.2 (Fig. 1d). This is a natural result for a material with no crystalline structure.

Taken together, the observations made on the main set of experimental samples show that the polarization dependences of the Raman scattering in porous silicon maintain a clear connection with the cubic symmetry of the initial material: The orientation of the rosettes in Fig. 1b and 1c corresponds to the crystallographic axes of the initial material, in agreement with existing results from x-ray crystallographic investigations (see, for example, Refs. 7 and 8). However, the mechanism responsible for the change in the orientational dependence of the Raman scattering cross section from the dependence for crystalline silicon is not obvious.

4. To determine this mechanism, the degree of polarization of light reflected from the PSi films and of light transmitted through a sample of free-standing PSi were measured for different orientations of the plane of polarization of the incident light. In all the cases measured the depolarization amounted to no more than several percent and did not have a regular angular dependence. This eliminates the possibility that the observed change in polarization is due to depolarization of the light on passage through the sample.

In addition, we made measurements of the angular dependence of the intensity of the elastic scattering. The three curves in Fig. 2 represent the elastic scattering pattern for polished crystalline silicon, porous silicon with a mirror surface, and, for comparison, a ground surface of crystalline silicon. It is evident from the curves in the figure that, just

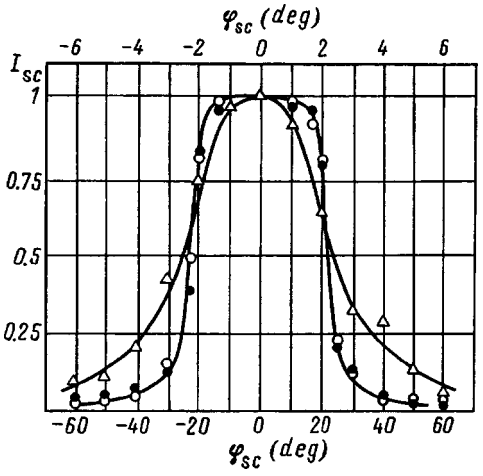

FIG. 2. Intensity of elastic scattering of light versus the scattering angle for samples of polished single-crystalline silicon (○), a PSi sample with a mirror surface (●) (top scale of angles), and the ground surface of a crystalline silicon sample (Δ) (bottom scale of angles).

like the sample of polished crystalline silicon, the porous silicon sample reflects light specularly with no scattering, but at the same time the polarizations are radically different in these two cases (see Figs. 1a, 1b, and 1c). This indicates that the porous silicon remains optically homogeneous and that light is not scattered by separate quantum wires.

In summary, the preliminary experiments established that the explanation of the anomalous polarization cannot be based on the assumption of scattering by inhomogeneities or of depolarization of the light. However, the intensity of Raman scattering at fixed polarizations of the exciting and scattered radiation should be sensitive to local deviations of the directions of the crystallographic axes from their average orientation. Hence it follows that if the quantum wires are disoriented, the intensities of the scattered components will differ from the case of an ideal crystal.[5]

It is obvious *a priori* that there can be two basic types of disorientation of the wires which could result in local changes of the directions of the crystallographic axes: twisting and inclination (bending). It is quite obvious (and confirmed by our calculations) that twisting of the filaments can only lead to an orientational dependence $\rho(\alpha)$ in the form of a superposition of a set of figures of the type in Fig. 1a and, as a result, in depolarization (decrease of scale) and rotation of the total orientational dependence, without causing any qualitative changes in it.

For further analysis, the data on the angular dependences $\rho(\alpha)$ were approximated by an expression similar to Eq. (1), i.e., containing a function of the angle 4α. It was found that the dependences of the type indicated in Fig. 1b and 1c can be described quite well by functions of the type

$$\rho = A + B \cos 4\alpha. \tag{2}$$

For negative phases of the cosine the contributions from A and B work against each other, causing degeneration of four of the eight lobes of the angular pattern.

The constants A and B in Eq. (2) have a simple physical meaning: In crystalline silicon a dependence of the type (1) arises only when light is scattered from a [100] surface (in this case $A=0$), while in the case of scattering from a [111] surface the function $\rho(\alpha)$ is a circle ($B=0$). This is the key to understanding the reasons why dependences of the type (1) appear: In porous silicon the local directions of the lattice vector and of the electric field vector of the light can be distorted on the scale of an individual filament and can alter the characteristics of the local Raman scattering process, but at the same time distortions on a nanometer scale will not show up in macroscopic experiments.

There is an obvious possible cause of the deviations of the indicated vectors — inclination of individual quantum wires with respect to the average direction. On account of the microscopic dimensions of the wires, strong scattering, refraction, or distortions of the wave front do not occur. However, the angle between the vector **E** of the light and the direction of the crystallographic axes can fluctuate in individual wires. Besides the obvious reasons, the angle will change because in an inclined wire the direction of the vector **E** should change relative to the initial direction. The same ideas were used in Refs. 9 and 10, where experiments on polarization luminescence of PSi are analyzed. The polarization dependences observed in this case can be explained well on the basis of the hypothesis that the local light field **E** in silicon wires is the same as if the wires were in a uniform external field in a medium with a much lower dielectric constant. Then, the component of the vector **E** along the axis of a wire will be greater than the component normal to the axis (as a result of the anisotropy of the polarization of the wire in the external field). As a result, the local deviations of **E** are substantial even for small angles of disorientation of the wires.

We performed calculations of the polarization dependences of the Raman scattering cross section for a model in which the porous silicon consists of randomly oriented wires in a medium with a lower dielectric constant. It was assumed that the crystallographic axis $\langle 001\rangle$ was directed along the axis of each wire. A similar model assumption was used in Refs. 9 and 10. The polarization dependences of the Raman scattering intensity for such a system are determined by the quantity

$$S=\eta\sum_{j=x,y,z}\langle(\mathbf{e}_i R_j \mathbf{e}_s)^2\rangle, \tag{3}$$

where $\mathbf{e}_i$ and $\mathbf{e}_s$ are the electric field vectors of the incident and scattered radiation in a separate wire, R_j is the Raman tensor for scattering of light by a phonon with polarization j, η is a constant, and the angle brackets in Eq. (3) denote averaging over all possible directions of disorientation of the wires. Let us give the result pf a computation of S for wires whose axes make an angle θ with the normal to the surface, while the transverse crystallographic axes partially retain the orientation of the initial crystal. It is convenient to write the expressions obtained for S in the form $S=S_\perp+S_z$, where S_z is the intensity of scattering by a phonon polarized in a direction along the axis of a wire and $S_\perp$ is the intensity of scattering by a phonon polarized in a direction perpendicular to the axis of the wire:

$$S_\perp=\eta a^2R^2\ \sin^2\theta\{\cos^2\theta[3/2\ \cos^2(\psi-\psi')+1/2]+1/2\}, \tag{4}$$

$$S_z = \eta R^2[(1+\cos\ \theta - 1/2\ \sin^2\theta)^2 \sin^2(\psi+\psi') + 1/2(1-\cos\ \theta - 1/2\ \sin^2\theta)^2 + 1/2\ \sin^4\theta \cdot 2\ \cos^2(\psi-\psi')]. \quad (5)$$

Here ψ and ψ' are the angles between the vectors $\mathbf{e}_i$ and $\mathbf{e}_s$ and the projection of the crystallographic axis $\langle 100\rangle$ of the wires onto the plane of the film, and a is a factor equal to the ratio of the electric field components parallel to and perpendicular to the axis of the wire and expressed in terms of the depolarization coefficients.[9] Since the degree of polarization is given by

$$\rho = [S(\psi=\psi') - S(\psi=\psi'+\pi/2)]/[S(\psi=\psi') + S(\psi=\psi'+\pi/2)], \quad (6)$$

a dependence of the form (1) arises for different random distributions of wires over the angles of inclination θ, with the transverse axes retaining their orientation. As a result of the large length-to-diameter ratio of a wire and the large difference between the dielectric constants of the wire and the medium, the pattern shown in Fig. 1c can be explained by the presence of a relatively small (less than 10%) fraction of wires which are inclined by substantial ($\sim 50°$) angles, while the remaining wires retain their initial orientation, or by the presence of a spread of orientations $\theta \sim 10°$ of the axes of all the wires relative to the normal to the surface. Our experiments do not give any grounds for favoring any one model over another, but the latter assumption is in better agreement with the data of Refs. 7 and 8, which indicate that the crystal structure of the substrate is retained in PSi. Irrespective of the specific structural model, the experimentally determined ratio A/B can characterize the average disorientation of the wires in porous silicon.

The observed effect manifests similarly to Raman scattering by a rough surface. When light passes through a rough surface, the light, being diffusely scattered by the surface, propagates in a wide range of angles, and components corresponding to different angles between the crystallographic axes and the direction of propagation should be present in the Raman-scattered light. For this reason, the total degree of polarization should also have a dependence on the orientation of the sample similar to Eq. (1). This assertion was checked experimentally. Indeed, the orientational dependence was similar to that shown in Fig. 1b and 1c. We note that the macroscopic character of the nonuniformity of the surface gives rise to much larger angles between the directions of propagation of the light in the material, but at the same time the field enhancement resulting from the shape of the filaments does not occur in the continuous medium. There is one other fundamental difference between the cases that we investigated and this latter case: The diffuse elastic scattering of light by a ground surface can be easily detected experimentally (Fig. 2), while in our case, on account of the local nature of the Raman scattering process, the nanometer-size nonuniformities are manifested only in the polarization of the scattered light.

5. In summary, in the present work a previously unknown property of porous silicon — a special type of orientational dependence of the degree of polarization of the Raman scattering — was observed. It was shown that this reflects the presence of a specific type of disordering in this material, wherein the macroscopic homogeneity of the material and the microscopic structure of the crystal lattice are preserved but disorder can occur on an intermediate (mesoscopic) spatial scale, the macromaterial being made up of individual

quantum wires. Investigation of the polarization of Raman scattering is a sensitive tool for studying this phenomenon, giving independent information about the mesostructure of the macromaterial of porous silicon.

This work was supported by the program ''Physics of Solid-State Nanostructures.''

[a)]e-mail: kompan@solid.pti.spb.su

[1]L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[2]I. I. Reshina and E. Guk, Fiz. Tekh. Poluprovodn. **27**, 732 (1993) [Semiconductors **27**, 401 (1993)].
[3]I. Gregora, B. Champagnon, and A. Halimaoui, J. Appl. Phys. **75**, 3034 (1994).
[4]M. E. Kompan, E. B. Kuz'minov, V. B. Kulik *et al.*, JETP Lett. **64**, 748 (1996).
[5]H. Poulet and J.-P. Mathieu, *Spectres de Vibration et Symmetries des Cristaux*, Gordon and Breach, Paris–New York–London, 1970.
[6]I. H. Campbell and P. M. Fauchet, Solid State Commun. **58**, 739 (1986).
[7]T. Ugami and M. Seki, J. Electrochem. Soc. **125**, 1339 (1978).
[8]A. Nakajiama, Appl. Phys. Lett. **62**, 2631 (1993).
[9]D. Kovalev, M. Ben-Chorin, J. Diener *et al.*, Appl. Phys. Lett. **67**, 1585 (1995).
[10]D. Kovalev, M. Ben-Chorin, J. Diener *et al.*, Phys. Rev. Lett. **79**, 119 (1997).

Translated by M. E. Alferieff

Effect of hole–hole scattering on the conductivity of the two-component 2D hole gas in GaAs/(AlGa)As heterostructures

S. S. Murzin and S. I. Dorozhkin
Institute of Solid State Physics, 142432 Chernogolovka, Moscow Region, Russia

G. Landwehr
Physikalisches Institut der Universität Würzburg, D-97074 Am Hubland, Würzburg, Germany

A. C. Gossard
University of California, Santa Barbara, CA 93106, USA

(Submitted 23 August 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 101–106 (25 January 1998)

The temperature dependences of the zero-magnetic-field resistivity ρ and magnetoresistance of the 2D hole gas in GaAs/(AlGa)As heterostructures are investigated in the temperature interval 0.4–4.2 K. As the temperature T is increased, (i) the resistivity ρ grows with a decreasing derivative $d\rho/dT$, and (ii) the positive magnetoresistance diminishes from about 40% at $T=0.4$ K to about 1% at $T=4.2$ K. The results are explained in terms of a temperature-dependent mutual scattering of the holes, accompanied by momentum transfer between two different spin-split subbands. © *1998 American Institute of Physics.*
[S0021-3640(98)00302-8]

PACS numbers: 73.40.Kp, 73.50.Jt

A positive magnetoresistance of up to 40% in weak magnetic fields has been observed at low temperatures in the high-mobility 2D hole gas of GaAs/(AlGa)As heterostructures in studies going back many years.[1–7] At first this magnetoresistance was attributed to two-carrier conduction.[3] It is known there are two groups of holes with different spectra and mobilities in the 2D hole systems of GaAs/(AlGa)As heterostructures. These two groups are formed from the heavy hole band as a result of the lifting of the spin degeneracy by the spin–orbit interaction in the absence of inversion symmetry.[8–10] In such systems a positive magnetoresistance should be observed[11] both in the case of the elastic scattering of holes by impurities and in the inelastic scattering of holes by phonons (even in the presence of inter-group scattering[12]).

However, in the experiments of Refs. 2, 4, and 7 the magnetoresistance was found to be strongly temperature dependent even at relatively low temperatures, when the electron–phonon scattering is unimportant. The magnetoresistance decreases with temperature, almost vanishing[2,7] at $T=4.2$ K. This has raised doubts that the effect is due to two-carrier conduction,[4,6] and new ideas have been put forward. In Refs. 2 and 4 it was

noted that a qualitatively similar effect can be caused by weak localization in a system with strong spin–orbit coupling.[13] However, the weak localization effects are too small to account for the large magnetoresistance in highly conductive heterostructures.[2] The authors of Ref. 6 hypothesized that the magnetoresistance could originate from quantum corrections due to the hole–hole interaction in disordered systems at values of the inverse screening length q_s that are large compared to the hole wave numbers at the Fermi level k_F. It can be shown that the magnetoresistance in this case is also small. The authors of Ref. 7 assumed that the magnetoresistance should be suppressed if the thermal energy k_BT (k_B is Boltzmann's constant) is much larger than the energy separation between the two bands at the wave vector of the smaller Fermi circle Δ_F. However, our calculations show that this factor alone cannot suppress the magnetoresistance but only leads to some changes in its value. Moreover, a drastic decrease in the magnetoresistance is observed when $k_BT \ll \Delta_F$. This factor alone obviously contradicts the idea of the authors of Ref. 7. Thus there is no satisfactory explanation of the strong temperature dependence of the magnetoresistance of the high-mobility 2D hole gas in GaAs/(AlGa)As heterostructures.

In this paper we propose a new idea which is capable of explaining this phenomenon: the mutual scattering of holes belonging to different groups. The equations derived here are compared with both the results of our detailed study of the temperature dependence of the zero-magnetic-field resistance and magnetoresistance and with all the available data; the results of this comparison demonstrates that the proposed effect gives a reasonable explanation of the data. It is important to note that the temperature dependence of the mutual scattering was found to be proportional to T^2, which supports the basic idea.

1. EFFECT OF HOLE–HOLE SCATTERING

The positive magnetoresistance in a system with two groups of carriers is caused by the difference between their drift velocities $\mathbf{u}_i$ in an electric field. Intense mutual scattering of carriers should equalize the velocities leading to a vanishing magnetoresistance. Equations introducing the mutual scattering into the transport problem have been derived previously[15–18] for the case when the inter-group scattering is absent. Here we use these equations to calculate the zero-magnetic-field resistance and magnetoresistance for the case of carriers with like charges and different mobilities. Although they should not be expected to describe the magnetoresistance very accurately, we hope that they will describe rather well the main features of the phenomenon. The equation of motion in an electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ for particles of group 1, taking into account the collisions with particles of group 2, has the form[16,17]

$$m_1\mathbf{u}_1/\tau_1 + \eta n_2(\mathbf{u}_1 - \mathbf{u}_2) = e\mathbf{E} + (e/c)(\mathbf{u}_1 \times \mathbf{H}). \tag{1}$$

A similar equation can be written for the particles of group 2. Here m_i are the effective masses, τ_i are the momentum relaxation times for each group, and η is the mutual friction coefficient

$$\eta = \frac{m_1 m_2}{m_1 n_1 + m_2 n_2}\,\frac{1}{\tau_{e-e}}. \tag{2}$$

Since the relaxation time τ_{e-e} of the relative drift velocity $\mathbf{u}_1-\mathbf{u}_2$ due to the mutual scattering of carriers is proportional to T^{-2} (Ref. 21; see also the Appendix), η can be written as

$$\eta=\alpha T^2. \tag{3}$$

By solving the system of equations for $\mathbf{u}_i$ and substituting these velocities into the expression for the current density $\mathbf{j}=n_1e\mathbf{u}_1+n_2e\mathbf{u}_2$, we find the conductivities σ_{xx} and σ_{xy}:

$$\sigma_{xx}=\frac{[nw(He/c)^2+(\eta nw+w_1w_2)(\eta n^2+n_1w_2+n_2w_1)]e^2}{(He/c)^4+[n^2\eta^2+2\eta(n_1w_2+n_2w_1)+w_1^2+w_2^2](He/c)^2+(\eta nw+w_1w_2)^2}, \tag{4}$$

$$\sigma_{xy}=\frac{n(He/c)^2+(\eta^2n^3+2\eta n(n_1w_2+n_2w_1)+(n_1w_2^2+n_2w_1^2))}{(He/c)^4+[n^2\eta^2+2\eta(n_1w_2+n_2w_1)+w_1^2+w_2^2](He/c)^2+(\eta nw+w_1w_2)^2}\frac{e^3}{c}H. \tag{5}$$

Here $w_i=m_i/\tau_i=e/\mu_i$, $n=n_1+n_2$, and $w=(w_1n_1+w_2n_2)/n$. The longitudinal and Hall resistivities are $\rho_{xx}=\sigma_{xx}/(\sigma_{xx}^2+\sigma_{xy}^2)$ and $\rho_{xy}=\sigma_{xy}/(\sigma_{xx}^2+\sigma_{xy}^2)$. At low temperatures, when $\tau_{e-e}\gg\tau_i$ $(n\eta\ll w)$, the conductivity is the sum of the conductivities of each group. In this case our equation for the magnetoresistance coincides with the equation given in Ref. 11. The magnetoresistance is positive and saturates in high magnetic fields $\mu_iH/c\gg1$. At high temperatures, when $\tau_{e-e}\ll\tau_i$, the longitudinal resistivity ρ_{xx} and Hall resistivity ρ_{xy} are equal to $\rho_{xx}=1/ne\mu$, $\rho_{xy}=H/nec$. Here $\mu=e/w$ is the average mobility. In this case, the magnetoresistance is absent and the zero-magnetic-field resistivity ρ does not change with temperature if the μ_i are temperature independent. In the intermediate range $\tau_{e-e}\sim\tau_i$ the temperature dependence of the resistance exists only in weak magnetic fields $\mu_iH/c\leqslant1$. The resistivity ρ_{xx} increases with temperature and saturates at high temperatures. The difference $\rho(T\to\infty)-\rho(T=0)$ is equal to the difference $\rho_{xx}(H\to\infty,T=0)-\rho(T=0)$.

2. EXPERIMENT

The two samples used in the experiment were prepared by molecular-beam epitaxy. Sample 1 consisted of a GaAs (100) substrate overgrown with the following layers: undoped GaAs (0.2 μm), a GaAs(20Å)/$Al_{0.26}Ga_{0.74}As$(20Å) periodic structure (20 periods), undoped GaAs (1 μm), undoped $Al_{0.26}Ga_{0.74}As$ (250 Å), $Al_{0.26}Ga_{0.74}As$ doped with Be to $\sim2.7\times10^{18}$ cm^{-3} (300 Å), and undoped GaAs (50 Å). Sample 2 differed from sample 1 by the content of Al in $Al_xGa_{1-x}As$ layers ($x=0.3$), by the thickness of the doped AlGaAs layer, which was equal to 200 Å, and by the presence of a cap layer which consisted of 150 Åof undoped $Al_{0.3}Ga_{0.7}As$ and 100 Å of undoped GaAs.

The densities n_1 and n_2 for the two different groups of holes were determined from the Shubnikov–de Haas oscillations at a low temperature and are listed in Table I. This procedure is similar to the one used in Refs. 1 and 3. In fields $H<1$ T the period of the oscillations is determined by the density n_1 of the holes with the lower mass and density. Above 2 T the period is determined by the total hole density n. For samples 1 and 2 the total densities are 3.23×10^{11} and 3.43×10^{11} cm^{-2}, and the average mobilities μ at $T=4.2$ K are 7.4×10^4 and 9.3×10^4 cm^2/V·s, respectively.

TABLE I.

Sample	n_1, cm^{-2}	n_2, cm^{-2}	$\mu_{1,0}$, $cm^2/V\cdot s$	$\mu_{2,0}$, $cm^2/V\cdot s$	α, $g\cdot cm^2/s\cdot K^2$	β, 1/K
1	1.14×10^{11}	2.09×10^{11}	22×10^4	5.4×10^4	3.7×10^{-29}	0
2	1.27×10^{11}	2.16×10^{11}	24.7×10^4	7.5×10^4	2.85×10^{-29}	—
2	1.27×10^{11}	2.16×10^{11}	24.6×10^4	7.7×10^4	2.54×10^{-29}	0.02

The temperature dependence of the resistivity at $H=0$ is shown in Fig. 1. Both samples show qualitatively similar behavior. The resistivity increases with temperature by about 40%, with a derivative $d\rho/dT$ that is largest at low temperatures. Up to $T\approx3$ K the derivative $d\rho/dT$ decreases and then starts to increase slightly. The magnetoresistance at different temperatures is shown in Figs. 2 and 3. The main effect, common to both samples, is a positive temperature-dependent magnetoresistance with a tendency to saturation at high magnetic fields. The magnetoresistance strongly decreases as the temperature increases from 0.4 to 4.2 K.

3. DISCUSSION

The hole–hole scattering explains both the strong decrease of magnetoresistance at high temperatures and the temperature dependence of the zero-magnetic-field resistivity, with decreasing $d\rho/dT$ observed at $T<3$ K. The quantum corrections due to weak localization[13] and the hole–hole interaction[14] in our samples should be smaller than 1%. The large value of $\Delta/k_B\approx10$ K (Refs. 8 and 10) contradicts the explanation given in Ref. 7.

We fitted the experimental data by Eqs.(3)–(5) by varying three unknown parameters, namely, the temperature-independent mobilities $\mu_{1,0}=e/w_1$, $\mu_{2,0}=e/w_2$ and α, trying to reach the best accuracy at low temperatures. The results of the fitting are shown in Figs. 1, 2, and 3. The chosen values of parameters are listed in Table I. The main features of the experimental data are described well by the fitting curves.

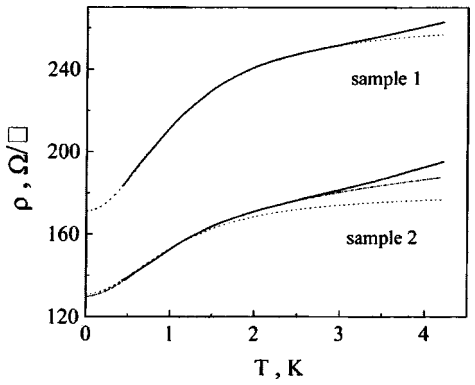

FIG. 1. Resistivity at zero magnetic field versus temperature. The solid lines are experimental curves, the dotted and dot-dashed lines show theoretical fits with temperature-independent and temperature-dependent hole mobilities, respectively.

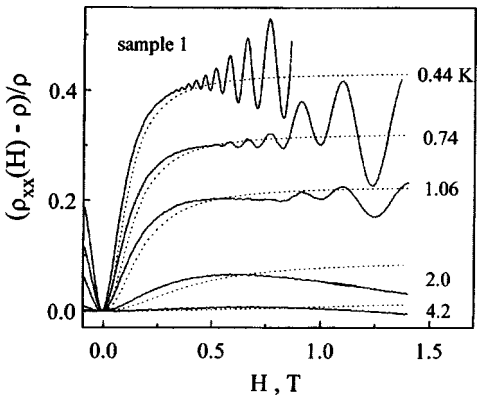

FIG. 2. Magnetoresistance $(\rho_{xx}-\rho)/\rho$ of sample 1 at different temperatures in a magnetic field perpendicular to the plane of the sample. The solid lines are experimental curves, the dotted lines represent the results of fitting.

There are several effects which were not taken into account by our simple model. These are inter-group scattering (for the case of elastic scattering this effect was considered in Ref. 12) and anisotropy of the hole Fermi surface. While the former effect can be suppressed in the case of elastic scattering by remote impurities (separated from the two-dimensional system by a spacer), it definitely exists in the case of the hole–hole scattering. These effects may be responsible for some discrepancies between the experimental and the theoretical magnetoresistance curves. The temperature dependence of the zero-magnetic-field resistivity should be much less sensitive to these factors. The differences between the fitting and the experimental curves observed in Fig. 1 at high temperatures can be explained by the temperature dependence of the mobilities μ_i due to electron–phonon scattering and to the finite value of k_BT/E_F (E_F is the Fermi energy, $E_F/k_B \approx 20$ K). The biggest correction caused by the latter effect is linear in k_BT/E_F because of the temperature dependence of the screening:[22]

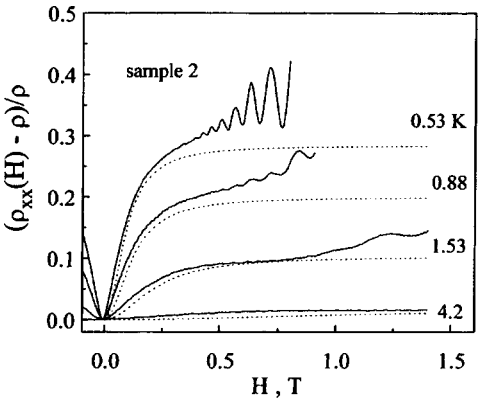

FIG. 3. Magnetoresistance $(\rho_{xx}-\rho)/\rho$ of sample 2 at different temperatures. The solid lines are experimental curves, the dotted lines represent the results of fitting.

$$\mu_i^{-1} = \mu_{i,0}^{-1}(1+\beta_i T). \tag{6}$$

This effect is important only for scattering with momentum transfer close to $2\hbar k_{F,i}$ ($k_{F,i}$ are the hole wave numbers at the Fermi level) and, therefore, is strongly dependent on the presence of the corresponding harmonics in a particular scattering potential. It can be very different for different samples even with similar structures. The fitting of the data with temperature-dependent mobilities given by Eq. (6), where we take $\beta = \beta_1 = \beta_2$, yields considerably better results for sample 2 (see Fig. 1). The results for sample 1 were not changed (for this sample β was found to be close to zero). New fitting parameters for sample 2 are also listed in Table I. The calculated magnetoresistance curves changed only slightly after taking into account the corrections to μ_i and we therefore do not present the new curves. The coefficients β have reasonable values smaller than $k_B/E_F \approx 0.05\ \mathrm{K}^{-1}$. It is worth noting that at $T = 4.2$ K the differences between the experimental curves and the new fitting curves in Fig. 1, which we ascribe to the electron–phonon scattering, are approximately equal for the two samples.

In order to verify whether η is proportional to T^2 we tried to fit the temperature dependence of the resistivity taking $\eta = \alpha T^p$ with $p = 1.5$ and 2.5 in the temperature range 0.4–3 K. In both cases the agreement with the experiment was noticeably worse in comparison with the case $p = 2$.

There are neither experimental nor theoretical data on η, α or τ_{e-e} in a two-component 2D electron (hole) gas. In Ref. 21, where the dependence $\tau_{e-e} \propto T^2$ was derived, the factor mutliplying T^2 was not calculated. In order to understand whether the values of α obtained from the fitting are reasonable or not, we have calculated τ_{e-e} and η for a simple model, following the approach of Ref. 18. This model neglects the anisotropy of the real energy spectrum and assumes the absence of hole transitions from one subband to the other. Although these conditions are not fulfilled in our system, we believe that the calculated value has the correct order of magnitude. Under the condition $q_s = e^2(m_1+m_2)/\kappa_0\hbar^2 \gg k_{iF}$ ($q_s/\max(k_{iF}) \approx 10$ in GaAs/(AlGa)As heterostructures with $n = 3\times 10^{11}\ \mathrm{cm}^{-2}$) we have

$$\eta = \frac{8}{3\hbar^3}\left(\frac{m_1 m_2}{m_1+m_2}\right)^2 \frac{1}{n_1 n_2} \ln\frac{\sqrt{n_1}+\sqrt{n_2}}{\sqrt{n_1}-\sqrt{n_2}} (k_B T)^2. \tag{7}$$

For the case of the effective masses $m_1 = 0.2m_e$ and $m_2 = 0.8m_e$ calculated in Refs. 8 and 10 we have obtained $\alpha \approx 7\times 10^{-29}\ \mathrm{g\cdot cm^2/s\cdot K^2}$ for our samples, which is in reasonable agreement with the experimental values.

We have checked that the published results on the temperature-dependent magnetoresistance for p channels in GaAs/AlGaAs heterostructures are consistent with our explanation. Unfortunately, a detailed comparison is not possible because, to the best of our knowledge, the only experimental data for which the temperature range was large enough to demonstrate strong variation of the magnetoresistance is given by Fig. 4 of Ref. 7. But in this paper only the total hole density $n = 2.08\times 10^{11}\ \mathrm{cm}^{-2}$ is presented. Nevertheless, we can approximately determine a coefficient $\alpha \approx 1\times 10^{-28}\ \mathrm{g\cdot cm^2/s\cdot K^2}$ for these data because it is not very sensitive to the n_1/n_2 ratio. The data presented in Fig. 5 of the same paper[7] for a low-mobility sample in the temperature range 0.3–1.3 K show only a weak temperature dependence of the magnetoresistance, which implies that τ_{e-e} is much less than the elastic scattering time and gives no chance of determining α.

Estimation of α is possible for the data presented in Fig. 2 of Ref. 4 ($n=3.8\times10^{11}$ cm^{-2}, $n_1=1.01\times10^{11}$ cm^{-2}), although the variation of the magnetoresistance there is not large there. This estimation gives $\alpha\approx1\times10^{-29}$ g·cm^2/s·K^2. The variation of α with hole density is consistent with the expected dependence (see Eq. (A8)) at least qualitatively.

In conclusion, we have shown that the temperature dependence of both the zero-magnetic-field resistance and the magnetoresistance of the 2D hole gas in GaAs/(AlGa)As heterostructures is governed by the hole–hole scattering at low temperatures. Similar effects can exist in other high-mobility semiconductor systems which contain several groups of carriers with different mobilities.

We acknowledge a helpful discussion with V. T. Dolgopolov and D. V. Shovkun. This work has been supported by Russian State Program ''Nanostructures'' (Grant 1-085/4) and by Deutsche Forschungsgemeinshaft.

[1] H. L. Stormer, Z. Schlesinger, A. Chang *et al.*, Phys. Rev. Lett. **51**, 126 (1983).
[2] E. E. Mendez, W. I. Wang, L. L. Chang, and L. Esaki, Phys. Rev. B **30**, 1087 (1984).
[3] J. P. Eisenstein, H. L. Stormer, V. Narayanamurti *et al.*, Phys. Rev. Lett. **53**, 2579 (1984).
[4] E. E. Mendez, Surf. Sci. **170**, 561 (1986).
[5] Y. Iye, E. E. Mendez, W. I. Wang, and L. Esaki, Phys. Rev. B **33**, 5854 (1986).
[6] Zhou Haiping, Zheng Houzhi, Yang Fuhua *et al.*, Chin. Phys. **10**, 223 (1990).
[7] M. Henini, P. A. Crump, P. J. Rodgers *et al.*, J. Cryst. Growth **150**, 446 (1995).
[8] U. Ekenberg and M. Altarelli, Phys. Rev. B **30**, 3369 (1984).
[9] D. A. Broido and L. J. Sham, Phys. Rev. B **31**, 888 (1985).
[10] T. Ando, J. Phys. Soc. Jpn. **54**, 1528 (1985).
[11] J. M. Ziman, *Principles of the Theory of Solids*, Cambridge University Press, 1972.
[12] E. Zaremba, Phys. Rev. B **42**, 14143 (1992).
[13] S. Hikami, A. I. Larkin, and Y. Nagaoka, Prog. Theor. Phys. **63**, 707 (1980).
[14] P. A. Lee and T. V. Ramakrishman, Rev. Mod. Phys. **57**, 287 (1985).
[15] C. A. Kukkonen and P. F. Maldaldague, Phys. Rev. Lett. **37**, 782 (1976).
[16] V. F. Gantmakher and Y. B. Levinson, Zh. Éksp. Teor. Fiz. **74**, 261 (1978) [Sov. Phys. JETP **47**, 133 (1978)].
[17] V. F. Gantmakher and Y. B. Levinson, *Carrier Scattering in Metals and Semiconductors*, Vol. 19 of Modern Problems in Condensed Matter Sciences, edited by V. M. Agranovich and A. A. Maradudin, Elsevier–North-Holland, 1987.
[18] J. Appel and A. W. Overhauser, Phys. Rev. B **18**, 758 (1978).
[19] A. N. Friedman, Phys. Rev. **159**, 553 (1967).
[20] A. H. Thompson, Phys. Rev. Lett. **35**, 1786 (1975).
[21] C. S. Ting, A. K. Ganguly, and W. Y. Lai, Phys. Rev. B **24**, 3371 (1981).
[22] A. Gold, V. T. Dolgopolov, J. Phys. C **18**, L463 (1985); Phys. Rev. B **33**, 1076 (1986).
[23] F. Stern, Phys. Rev. Lett. **18**, 546 (1967).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Optical phonons in quantum-wire structures

A. Milekhin, Yu. Pusep, Yu. Yanovskiĭ, V. Preobrazhenskiĭ, and B. Semyagin
Institute of Semiconductor Physics, 630090 Novosibirsk, Russia

(Submitted 10 June 1997; resubmitted 2 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 107–110 (25 January 1998)

The optical vibrational modes in GaAs/AlAs structures grown on a (311)*A*-oriented GaAs surface are investigated. It is found that the line corresponding to the fundamental *TO* vibrational mode localized in a GaAs quantum wire is split into two lines with different directions of the polarization vector. The dispersion of the *TO* phonons of GaAs in the (311) direction is determined from the IR spectra of periodic structures. © *1998 American Institute of Physics.*
[S0021-3640(98)00402-2]

PACS numbers: 63.22.+m, 78.30.Fs

Progress in molecular-beam epitaxy (MBE) technology has made it possible to grow perfect GaAs/AlAs superlattices (SLs) on high-index GaAs surfaces. Unlike the growth of SLs on (100)-oriented GaAs/AlAs, this leads to optical anisotropy in the plane of the SL layers.[2–6] The anisotropy of the optical and electronic properties of high-index SLs can be explained by periodic surface faceting, which has been found to be maximum for the (311)*A*-oriented SLs.[2,3] Surface faceting makes it possible to obtain quantum-wire structures directly during growth by MBE. The question of the presence of faceting and the height of a facet is now being debated in the literature. The first RHEED investigations of SLs grown on a (311)*A* surface of GaAs established the period of the faceting in the (011) direction to be $d=32$ Å and the facet height to be 10.2 Å (6 monolayers).[2,3] However, analysis of Raman scattering spectra gave a facet height of 2 monolayers.[7]

At present it is of great interest to study the electronic properties of (311) SLs. At the same time there exist only several works on the investigation of the vibrational properties of such structures.[7–10]

In the present letter we report the results of an investigation of the vibrational spectrum of (311)*A*-oriented GaAs/AlAs heterostructures by the method of IR Fourier spectroscopy.

We investigated $(GaAs)_n/(AlAs)_m$ structures ($n=8$, 10, 12, 28 and $m=12$, 16, 24 are the numbers of monolayers in corresponding layers) grown on a (311)*A*-oriented GaAs surface. The thicknesses of the GaAs and AlAs layers were monitored by observing the oscillations of the intensity of the specular reflection in the RHEED pattern using SLs grown on (100) GaAs substrates in the same process. The values of n and m were calculated for $(GaAs)_n/(AlAs)_m$ SLs with ideal heterointerfaces, neglecting the faceting. There were 10–200 repetitions of the layers for different structures.

The IR reflection spectra were measured at 80 K with a Bruker IFS-113V IR Fourier spectrometer equipped with an Oxford Instruments cryostat. The resolution was equal to 0.5 cm^{-1} over the entire spectral range. The IR reflection spectra were recorded with normal incidence of light on the sample. The required component of the polarization vector was selected with a polarizer.

The lower symmetry of the (311) SL compared with the (001)-oriented SLs makes it more difficult to identify the vibrational modes. In a (311) SL the vibrational modes propagating in the (311) direction have either a purely transverse character (A″ modes), if they are polarized in the (011) direction, or a mixed longitudinal/transverse character (A′ modes).[8] Moreover, the surface faceting in a (311)*A* SL can result in splitting of the localized vibrational modes. For example, the splitting of localized *LO* modes in GaAs layers could be due to localization of phonons in the narrow and wide parts of the faceted layers.[7]

According to the selection rules, on account of their nonzero dipole moment all odd modes (A′ and A″) can be active in the IR spectra of (311)-oriented GaAs/AlAs SLs. The wave number of the A″ localized modes of long-period of SLs (where surface faceting can be neglected) can be determined as

$$q_m = mp/\{(n+d)d\}, \tag{1}$$

where n is the number of monolayers, $d = a\sqrt{11}$ is the thickness of one monolayer in the (311) direction, a is the lattice constant in the (100) direction, and m is the number of the localized mode. The parameter δ describes the penetration of localized modes into neighboring layers.

The identification of mixed localized modes with small wave numbers in (311)-oriented GaAs/AlAs SLs simplifies because the corresponding localized modes have mainly either *LO* or *TO* polarization.[10] Moreover, though the degeneracy of the *TO* modes polarized in the (233) and (011) directions is lifted, the magnitude of the splitting of these modes remains negligibly small all the way down to wave numbers $q \approx 0.3$ (Ref. 8).

The IR reflection spectra of GaAs/AlAs SLs recorded under conditions of normal incidence of the light make it possible to observe *TO* phonons localized in GaAs and AlAs layers. In addition, only odd localized modes, whose total dipole moment differs from zero, are manifested in the spectra. The IR spectra of the derivative of the reflectance in the region of the *TO* phonon frequencies in GaAs for long-period GaS/AlAs SLs grown on a (311)*A* GaAs surface are displayed in Fig. 1. To determine the frequencies of the localized modes, the IR reflection spectra calculated for a multilayer structure by the method described in Ref. 11 was fit to the experimental spectra. The arrows in the figure mark the spectral position of the frequencies of the higher-order odd localized TO_i modes, determined from the best agreement between theory and experiment. The TO_b line refers to a bulk *TO* phonon of the GaAs substrate. The values obtained for the frequencies of the localized modes from the IR reflection spectra for SLs with a different thickness of the layers and the corresponding wave numbers determined from the relation (1) were used to construct the dispersion curve for GaAs *TO* phonons in the (311) direction. The parameter δ was set equal to 1. The experimental values of the frequencies of the *TO* modes as a function of the wave vector are represented by triangles in the inset in Fig. 1. The circles represent the Raman scattering data taken from Ref. 5. For com-

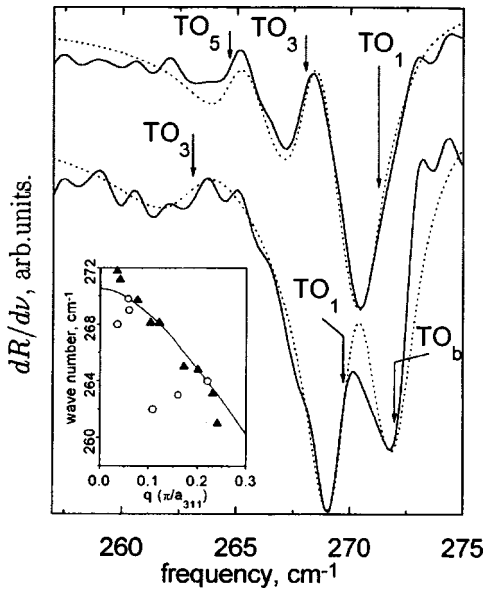

FIG. 1. IR reflection spectra of (311)*A*-oriented superlattices: $(GaAS)_{24}/(AlAs)_{12}$ (curve *1*) and $(GaAS)_{12}/(AlAs)_{12}$ (curve *2*). The dotted line shows the computed spectra. Inset: Dispersion curve of GaAs *TO* phonons in a (311)*A* superlattice. For comparison, the solid line shows the dispersion of GaAs *TO* phonons in GaAs in the (100) direction. The triangles represent the IR spectroscopy data[5] and the circles represent the Raman scattering data.

parison, the solid line is the dispersion curve of GaAs *TO* phonons in the (100) direction. As one can see from the figure, the data obtained by different methods are consistent with one another.

We note that the long-period (311)*A* GaAs/AlAs SLs studied in the present work and earlier in Ref. 12 did not exhibit anisotropy of the optical phonons in the plane of the SL layers as the sample was rotated relative to the growth axis with the polarization of the light held fixed.

We present the observation of anisotropy of the vibrational spectrum of short-period $(GaAs)_n/(AlAs)_m$ ($n=9$, 10 and $m=12$) structures, where surface faceting can lead to the formation of quantum wires. Figure 2 shows the IR spectra of the derivative of the reflectance of the structures, recorded with different directions of the polarization vector of the light. As one can see from the figure, the fundamental mode TO_1 splits into two modes with different directions of the polarization vector. In the case when the polarization vector of the light is directed along the quantum wires ((233) direction), a low-frequency $TO_\parallel$ mode propagating in the (311) direction dominates the IR spectra (curves *1*,*3*); here the polarization vector of the mode is directed in the direction of the light vector. If the polarization vector of the light is perpendicular to the quantum wires ((011) direction), a high-frequency $TO_\perp$ mode is clearly observed in the spectra. The polarization vector of this mode is perpendicular to the quantum wires, and for this reason this mode can interact effectively with IR radiation polarized in the (011) direction. The maximum splitting is obtained for $(GaAs)_8/(AlAs)_{16}$ structures and equals ~ 1.3 cm^{-1}.

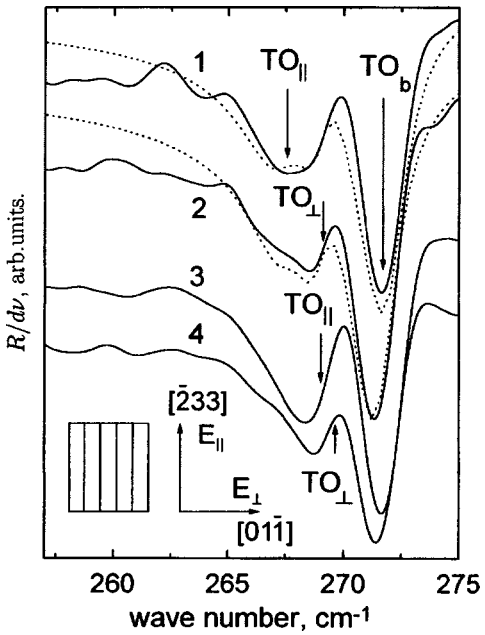

FIG. 2. IR reflection spectra of (311)*A* GaAs/AlAs SL measured with normal incidence of linearly polarized light in the spectral range of GaAs *TO* phonons: $(GaAS)_8/(AlAs)_{16}$ (curves *1* and *2*) and $(GaAS)_{10}/(AlAs)_{16}$ (curves *3* and *4*). The dotted lines show the computed spectra. Inset: Direction of the polarization vector of the light relative to the direction of the quantum filaments.

As the thickness of the GaAs layers in the SLs increases, the splitting of the *TO* modes decreases (curves *3* and *4* in Fig. 2).

The observed splitting of the TO_1 localized modes can be explained by the array of GaAs and AlAs quantum wires formed in the process of MBE growth. The calculations of the vibrational spectrum for rectangular GaAs wires separated by AlAs barriers[13,14] showed splitting of the TO_1 vibrational modes. In addition, investigations of the conductivity anisotropy[15] showed that an array of quantum wires could have formed in GaAs/AlAs structures grown on (311)*A*-oriented GaAs surfaces. It was found that the maximum splitting of the *TO* localized modes is observed in $(GaAs)_8/(AlAs)_{16}$ structures, where the largest conductivity anisotropy was observed. For smaller thicknesses of the layers, the intensity of the localized modes decreases sharply. This can be explained by the appearance of wire-like clusters[16] and, in consequence, an increase in the damping of localized modes in them.

In summary, in this work the localized optical vibrational modes of GaAs/AlAs structures grown on a (311)*A*-oriented GaAs surface were investigated by IR Fourier spectroscopy. It was found that the fundamental *TO* mode localized in the GaAs quantum wires splits. The dispersion of GaAs *TO* phonons in the (311) direction, as determined from the IR spectra of long-period of SLs, is in good agreement with the Raman scattering data.

This work was supported by the Russian Fund for Fundamental Research (Project

95-02-04431) and the State Committee of the Russian Federation on Higher Education (Project ZN-249-96/231 MP 2-5).

[1] L. N. Pfeiffer, K. W. West, H. Stormer *et al.*, Appl. Phys. Lett. **56**, 1697 (1990).
[2] R. Nötzel, N. Ledentsov, L. Däweritz, and K. Ploog, Phys. Rev. Lett. B **27**, 3812 (1991).
[3] R. Nötzel, N. N. Ledentsov, L. Däweritz *et al.*, Phys. Rev. B **45**, 3507 (1992).
[4] R. Nötzel, L. Däweritz, and K. Ploog, Phys. Rev. B **46**, 4736 (1992).
[5] C. Jouanin, A. Hallaoui, and D. Bertho, Phys. Rev. B **50**, 1645 (1994).
[6] G. Gershoni, I. Brener, G. A. Baraff *et al.*, Phys. Rev. B **44**, 1930 (1991).
[7] S. W. da Silva, Yu. A. Pusep, J. Galzerani *et al.*, Phys. Rev. B **53**, 1927 (1996).
[8] Z. V. Popovich, E. Richter, J. Spitzer *et al.*, Phys. Rev. B **49**, 7577 (1994).
[9] Yu. A. Pusep, S. W. da Silva, J. Galzerani *et al.*, Phys. Rev. B **51**, 5473 (1995).
[10] P. Castrillo and L. Colombo, Phys. Rev. B **49**, 10362 (1994).
[11] A. G. Milekhin, Yu. A. Pusep, V. V. Preobrazhenskii *et al.*, JETP Lett. **59**, 493 (1994).
[12] A. Milekhin, Yu. Pusep, D. Lubyshev *et al.*, in *Proceedings of the 22nd International Symposium on Compound Semiconductors*, Cheju Island, Korea, 1995; published in *Compound Semiconductors*, Institute of Physics Conference Series, 1996, No. 145, Chapter 3, IOP Publishing Ltd. Philadelphia, Bristol, p. 437.
[13] B.-F. Zhu, Phys. Rev. B **44**, 1926 (1991).
[14] F. Rossi, C. Bungaro, L. Rota *et al.*, Solid-State Electron. **37**, 761 (1994).
[15] V. Ya. Prints, I. A. Panaev, V. V. Preobrazhenskiĭ, and B. R. Semyagin, JETP Lett. **60**, 217 (1994).
[16] Zh. I. Alferov, A. Yu. Egorov, A. E. Zhukov *et al.*, Fiz. Tekh. Poluprovodn. **26**, 1715 (1992) [Sov. Phys. Semicond. **26**, 959 (1992)].

Translated by M. E. Alferieff

On the nature of the $\gamma-\alpha$ phase transition in cerium

G. Eliashberg
Institut für Theoretische Physik, RWTH Aachen, D-52056 Aachen, Germany; Landau Institute of Theoretical Physics, 142432 Chernogolovka, Russia

H. Capellmann
Institut für Theoretische Physik, RWTH Aachen, D-52056 Aachen, Germany

(Submitted 15 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 111–117 (25 January 1998)

In 1964 Davis and Adams established that the large increase of the thermal expansion and compressibility in the critical region of the γ- to α-Ce phase transition occurs predominantly in the α phase. This provides strong evidence that a tricritical point is realized in Ce. This also means that the aforementioned transition is not isomorphic and that α-Ce should have a distorted fcc structure. A careful examination of Jayaraman's data (1965) shows that a second-order transition line continues beyond the tricritical point to the vicinity of a triple point on the melting curve. The phase boundary with the tricritical point and the minimum of the melting curve are reconstructed within the framework of Landau theory. © *1998 American Institute of Physics.*
[S0021-3640(98)00502-7]

PACS numbers: 64.70.Kb, 65.70.+y, 81.05.Bx

1. The $P-T$ phase diagram of Ce shows a multitude of phases. Except for the body-centered tetragonal phase appearing at room temperature above 120 kbar,[1,2] the other known phases fall into a relatively low-pressure domain. They are drawn in Fig. 1, which is to some extent schematic, reflecting a substantial experimental uncertainty. The data were taken from Refs. 2 and 3. The line YZ is a second-order transition boundary predicted in our paper. More precise data are available for the high-temperature part of Fig. 1, where the melting curve has a negative slope at ambient pressure and goes through a minimum at around 33 kbar and 935 K.[4] This remarkable feature will be discussed below. At the periphery of Fig. 1 one can see a body-centered cubic (bcc) δ phase, a double-hexagonal close packed (dhcp) β phase, and two low-symmetry phases α' (orthorhombic α-uranium structure) and α'' (body-centered monoclinic), which at room temperature coexist in a fragile equilibrium for pressures between 40 and 120 kbar. The main part of Fig. 1 is occupied by the γ and α phases. The widely accepted view is that both γ-Ce and α-Ce have a simple face-centered cubic lattice. Across the line XY separating these phases a first-order transition occurs. It was found that the volume and the entropy changes at the transition are large at room temperature and below ($\Delta v/v \gtrsim 15\%, \Delta s \gtrsim 1.5$ per atom) and tend to zero beyond 500 K, indicating the existence of a terminal point. In 1958 Ponyatovskiĭ had already proposed that this might be a critical

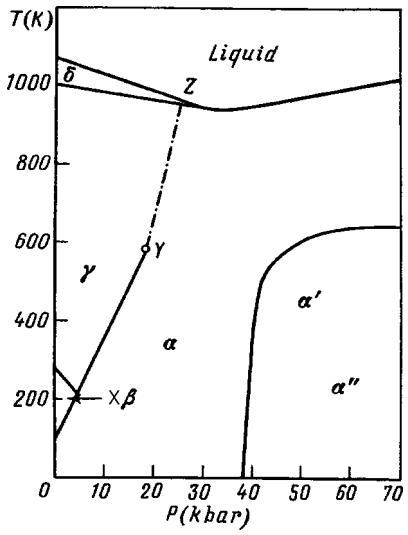

FIG. 1.

point like that in the vapor–liquid system.[5] This implies a singular behavior of the second derivatives of the thermodynamic potential (specific heat, compressibility, thermal expansion), which should tend to infinity at the critical point. Indeed in 1960 Beecroft and Swenson[6] observed a 10-fold increase of the thermal expansion in a critical region with respect to that at ambient conditions. A few years later Davis and Adams[7] in their elegant x-ray diffraction study confirmed this effect, and this was considered as a further justification for the critical point concept. But they also made an additional observation, the importance of which was not properly recognized: They were able to establish that singular behavior of the thermal expansion and compressibility occur only in the high-pressure phase, that is, in α-Ce. This was in fact an *experimentum crucis* which could long ago have lead to the unambiguous conclusion that instead of an ordinary critical point, the critical point in Ce is ''the critical point of a continuous phase transition'' (the general concept was developed by Landau in 1935–1937[8,9] and is now called a ''tricritical point'' (a name proposed by Griffiths[10]). According to Landau a first-order phase transition between two phases having different symmetry continues beyond a tricritical point as a second-order phase transition. In the vicinity of the tricritical point the compressibility, etc. diverge, but only in that phase which has the lower symmetry. This is exactly what was observed by Davis and Adams. We have to recognize therefore that α-Ce should have lower symmetry than γ-Ce. A phase transition from fcc to a distorted fcc phase has also been discovered in lanthanum and in praseodymium, the neighbors of Ce in the periodic table,[11–14] and within this systematics a distorted fcc structure of α-Ce is quite reasonably expected. The diffraction patterns for the distorted structures in La and Pr show weak superlattice reflections together with the set of strong reflections typical for the fcc structure. Due to the topology of the phase diagram good long-range-ordered crystals of α-Ce have not been available, and substantial line broadening masks the weak superlattice reflections. In the past this prevented a direct observation of the distorted structure of α-Ce.

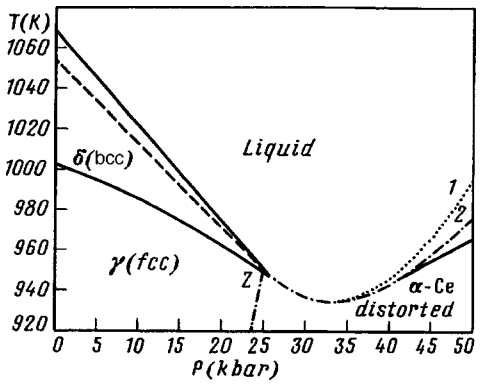

FIG. 2.

2. In this section we analyze the experimental information regarding the high-temperature part of the phase diagram.[4] Clear evidence will be found for a second-order $\gamma-\alpha$ transition in the vicinity of a triple point on the melting curve. In Fig. 2 the measured phase boundaries (solid lines) are shown together with our calculations (explained below). An important feature is the minimum of the melting temperature as a function of pressure. It follows from the Clausius–Clapeyron equation it follows that below the minimum ($P<33$ kbar) the solid is less dense than the liquid, but the situation reverses at $P>33$ kbar. Jayaraman[4] attributed this effect to a volume–pressure anomaly in the solid phase, supposing that at ambient pressure liquid Ce is already in a ''collapsed state,'' exhibiting the regular volume contraction with increasing pressure. The anomaly in solid Ce, being very strong in the critical region around point Y, should still be significant in the vicinity of the melting curve. This explanation is no doubt qualitatively correct, but within the tricritical point scenario we have to take into account that the anomalous properties manifest themselves in the low-symmetry phase. This, therefore, has to be the α phase, which exists in equilibrium with the liquid around the melting curve minimum. On the other hand it appears that the neighboring phase below the border with the δ phase is the high-symmetry phase, γ-Ce. This becomes apparent from an analysis of the corresponding boundary line, which we identify below as a $\gamma-\delta$ phase boundary. This line has a pronounced downward curvature (Fig. 2) and can be fitted by the equation

$$T=1003-1.4P-0.0305P^2, \tag{1}$$

where T is in kelvin and P in kilobar. The same equation can be written after expanding the equilibrium condition for the chemical potentials ($\mu_\delta=\mu_\gamma$ at the boundary) in powers of P and $T-T_0$, where $T_0=1003$ K, as

$$-(s_\delta-s_\gamma)(T-T_0)+(v_\delta-v_\gamma)P-\frac{1}{2}(\kappa_\delta-\kappa_\gamma)P^2=0. \tag{2}$$

Here s is the entropy per atom, v is the atomic volume, and $\kappa=-dv/dP$ is the compressibility. Other quadratic terms which are due to the thermal expansion and the heat capacity differences give negligible contributions. From the measured data we calculate the volume and entropy changes across the transition at $P=0$:

$$v_\delta - v_\gamma = -0.062 \text{ Å}^3; \quad s_\delta - s_\gamma = 0.044 \text{ Å}^3 \text{ kbar/K} \tag{3}$$

(1 Å^3 kbar/K corresponds to the dimensionless value 7.25 for the entropy per atom). Comparing Eqs. (1) and (2), we find the compressibility difference

$$\kappa_\delta - \kappa_\gamma \simeq 0.0027 \text{ Å}^3/\text{kbar}, \tag{4}$$

which is only $\sim 2\%$ of κ_γ ($\kappa_\gamma \approx 0.15$ Å^3/kbar for a bulk modulus $K \approx 230$ kbar and $v_\gamma \approx 34$ Å^3/kbar): There is no evidence for any anomalous increase of κ_γ. This means that Ce indeed remains in its ''normal'' high-symmetry γ phase along most of the γ–δ line. We come to the conclusion that the possible positions of point Z are limited to an interval of a few kilobar around the triple point (see Fig. 2), lying either on the γ–δ line or on the melting curve.

We can now calculate the coordinates (P_Z, T_Z) of the point Z and the compressibility jump at the γ–α transition. Suppose first that $P_Z > 26$ kbar. Beyond Z the liquid is therefore in equilibrium with the α phase: $\mu_{\text{liq}} = \mu_\alpha$, $\mu_\alpha = \mu_\gamma + \Delta\mu$. Expanding $\mu_{\text{liq}} - \mu_\gamma$ in powers of $P - P_Z$ and $T - T_Z$, we get the equation for the melting curve, $P > P_Z$:

$$-(s_{\text{liq}} - s_\gamma)(T - T_Z) + (v_{\text{liq}} - v_\gamma)(P - P_Z) - \frac{1}{2}(\kappa_{\text{liq}} - \kappa_\gamma)(P - P_Z)^2 - \Delta\mu = 0. \tag{5}$$

The main contribution to $\Delta\mu$ is due to the compressibility jump $\Delta\kappa$:

$$\Delta\mu = -\frac{1}{2}\Delta\kappa \cdot (P - P_Z)^2. \tag{6}$$

From the linearity of the border between the liquid and the δ phase we find that $\kappa_{\text{liq}} = \kappa_\delta$, and therefore $\kappa_{\text{liq}} - \kappa_\gamma = \kappa_\delta - \kappa_\gamma$, Eq. (4). Using the measured data we also get

$$v_{\text{liq}} - v_\delta = -0.38; \quad s_{\text{liq}} - s_\delta = 0.081 \tag{7}$$

(the units are as in Eq. (3)).

These values remain constant along the melting line. According to Eqs. (3) and (4) $v_\delta - v_\gamma = -0.13$ at the triple point, while $s_\delta - s_\gamma$ does not change appreciably. Using the values of Eq. (6) we get finally the coefficients in Eq. (5):

$$v_{\text{liq}} - v_\gamma = -0.51; \quad s_{\text{liq}} - s_\gamma = 0.125. \tag{8}$$

The initial slope of the melting curve on the high-pressure side of the triple point is therefore equal to 4.1 K/kbar. Defining a straight line having this slope and going through the triple point (dashed line in Fig. 2),

$$T = 1053 - 4.1P, \tag{9}$$

we impose three conditions necessary for calculating T_Z, P_Z, and $\Delta\kappa$: T_Z and P_Z satisfy Eq. (9); the coordinates of the melting curve minimum, viz., $T_m = 935$ K and $P_m = 33$ kbar, satisfy Eq. (5), and $dT/dP = 0$ at the minimum of the melting curve, Eq. (5). Solving the corresponding equations, we find that

$$\Delta\kappa \approx 0.065 \text{ Å}^3/\text{ kbar} \tag{10}$$

and $P_Z \approx 25$ kbar. This value is indeed very close to (but ~ 1 kbar below) the triple point. As a result, the calculated point is not the true point Z, but a point Z' which falls on the

line specified by Eq. (9) at a temperature $T'_Z \simeq 951$ K about 4 K above the $\gamma-\delta$ phase boundary. Estimates show that to an accuracy of ~ 1 K and a few tenth of kilobar we can simply take the projection of Z' on the $\gamma-\delta$ line to get the correct position of point Z:

$$P_Z \approx 25 \text{ kbar}, \quad T_Z \approx 947 \text{ K}. \tag{11}$$

The melting curve calculated using Eqs. (5) and (6) (with T'_Z instead of T_Z!) and shown as curve *1* in Fig. 2. coincides with the measured one along an interval of about 10 kbar beyond the triple point.

The discrepancy at higher pressure is, of course, not surprising for an approximate version of $\delta\mu$, Eq. (6), and is diminished within a more general description (curve *2* in Fig. 2; see next section).

3. It is most natural to expect that a second-order transition line continues from point Z to the tricritical point Y. Beyond Y a well-known first-order phase transition occurs (Fig. 1), and to get a quantitative description of this peculiar situation we expand (following Landau[8,9]) the chemical potential $\mu(P,T,u)=\mu(P,T,0)+\Delta\mu(P,T,u)$ in powers of some amplitude u, related to the lattice distortion:

$$\Delta\mu = Au^2 + Bu^4 + Cu^6, \quad C>0. \tag{12}$$

To avoid misunderstandings we note that the transition to the distorted structure is, most probably, driven by the softening of some phonon mode at a high-symmetry point in the Brillouin zone (see Sec. 4). The resulting distortion is some definite superposition of displacements (corresponding to several points with the same symmetry) which provides the minimum value of the fourth-order term. The expansion (12) is written for the amplitude u of this already selected superposition.

For $B(P,T)>0$ a second=order transition occurs along the line ZY, defined by the equation $A(P,T)=0$; $A>0$, $u=0$ for a ''normal'' γ phase, and $A<0$, $u\neq 0$ for a distorted α phase. Let $(P'T')$ be some point on ZY. Using the expansion

$$A(P,T) = \alpha(P'T')(T-T') - \beta(P'T')(P-P'), \tag{13}$$

we find the discontinuities of the compressibility $\kappa=-dv/dP$, the thermal expansion $\eta=dv/dT$, and the heat capacity c (per atom):

$$\Delta\kappa = \frac{\beta^2}{2B}, \quad \Delta\eta = \frac{\alpha\beta}{2B}, \quad \Delta c = T'\frac{\alpha\beta}{2B} \tag{14}$$

where $\Delta\kappa=\kappa_\alpha-\kappa_\gamma$, etc., and all quantities depend on P',T'.

The tricritical point Y appears as a crossing point of the lines $A(P,T)=0$ and $B(P,T)=0$. In the vicinity of Y

$$B(P,T) \approx B_0(P,T) = \gamma(T-T_Y) - \beta(P-P_Y) \tag{15}$$

and $\Delta\kappa$, $\Delta\eta$, and Δc therefore tend to infinity as $(T'-T_Y)^{-1}$.

Below T_Y the $\gamma-\alpha$ phase boundary continues as a line of first-order transitions: $B^2=4AC$ (Ref. 9). Within the scope of Eq. (13), taken around Y, and Eq. (16) the phase boundary is defined by a simple quadratic equation. In terms of the dimensionless variables t and p,

$$t=\frac{T-T_Y}{T_Z-T_Y},\quad p=\frac{1}{P_0}\left(P-P_Y-\frac{T-T_Y}{k_A}\right) \tag{16}$$

this equation is (for $t<0$):

$$p+\frac{3}{4}(t-\nu p)^2=0,\quad \nu=\frac{k_Ak_B}{k_A-k_B}\frac{P_0}{T_Z-T_Y}, \tag{17}$$

where

$$P_0=\frac{B_0^2(Z)}{3\beta C},\quad k_A=\frac{\beta}{\alpha},\quad k_B=\frac{\delta}{\gamma}, \tag{18}$$

and $B_0(Z)=\gamma(1-k_B/k_A)(T_Z-T_Y);\alpha=\alpha(Y),\beta=\beta(Y)$ (see Eq. (13)). Although the information concerning the position of the singular point Y is vague, the measured temperature evolution of the pressure-resistance isotherms[4] shows that 530–560 K is the most plausible interval for T_Y. Taking $T_Y=550$ K and $k_A=(T_Z-T_Y)/(P_Z-P_Y)$ (supposing ZY is close to a straight line), we have calculated P_Y and the parameters P_0 and ν, fitting Eq. (17) to the measured characteristics of the phase boundary. For the values

$$P_Y=16.7\text{ kbar},\quad P_0=28\text{ kbar},\quad \nu=1.12 \tag{19}$$

the theoretical curve

$$p(t)=-\frac{2}{3\nu^2}\left(1-\frac{3}{2}\nu t-\sqrt{1-3\nu t}\right),\quad t<0 \tag{20}$$

passes through the point $P=7.2$ kbar at $T=300$ K and is nearly linear up to 450 K, with a slope $k(T)=dT/dP$ that changes slowly from 22 to 27 K/kbar. These characteristics reproduce the measured data well within the experimental uncertainties. Above 450 K $k(T)$ increases to the value $k_A\approx 48$ K/kbar at point Y. A substantial increase of $k(T)$ above 450 K has been reported,[15] but, all in all, the data in this region are controversial.

Provided that $\Delta\mu$ in Eq. (12) is minimal,[9] u^2 is given in terms of p,t by the equation:

$$\beta u^2=2P_0(\Delta\kappa)_0 f;\quad (\Delta\kappa)_0=\beta^2/2B_0(Z), \tag{21}$$

where

$$f=f(t,p)=-t+\nu p+\sqrt{(t-\nu p)^2+p}. \tag{22}$$

Inserting Eq. (21) with (22) into Eq. (12), we get $\Delta\mu(p,t)$:

$$\Delta\mu=-\frac{2}{3}(\Delta\kappa)_0P_0^2[2p-(t-\nu p)f]f \tag{23}$$

and for $\Delta v=\partial\Delta\mu/\partial P$ and $\Delta s=-\partial\Delta\mu/\partial T$ we find:

$$\Delta v=-2(\Delta\kappa)_0P_0(1+\nu f)f;\quad \Delta s=-2(\Delta\kappa)_0P_0\left(\frac{1}{k_A}+\frac{\nu}{k_B}f\right)f. \tag{24}$$

In particular, at the first-order transition line Eq. (20), $t<0$, we have

$$f(t,p(t))=\frac{1}{\nu}(\sqrt{1+3\nu|t|}-1), \tag{25}$$

and, using the measured value of the volume jump at $T=300$ K, $\Delta v\approx 4.5$–4.8 Å^3 (Ref. 3), we find (for P_0 and ν given by Eq. (19)):

$$(\Delta\kappa)_0\approx 0.067\text{–}0.072 \text{ Å}^3/\text{kbar}. \tag{26}$$

Using $k_A=48$ K/kbar and $k_B=12$ K/kbar (for $\nu=1.12$, Eq. (17)), we find from Eq. (24) that $\Delta s=-(1.61\text{–}1.73)$ per atom at $T=300$ K, in agreement with the measured value.[3]

We emphasize that $(\Delta\kappa)_0$ in Eq. (26) is close to $\Delta\kappa$ in Eq. (10). Comparing Eq. (14) for $\Delta\kappa$ and Eq. (21) for $(\Delta\kappa)_0$ we see that the similarity obtained is consistent with a situation in which the characteristic temperature scale in the expansion (15) of B is much larger than T_Z-T_Y and in which the coefficients $\alpha(P'T')$ and $\beta(P'T')$ in Eq. (13) are only weakly varying along the line ZY. This does not seem surprising: within the interval considered T is several times the Debye temperature $\Theta\approx 130$ K and the aforementioned scale should be of electronic origin.

The line $A(P,T)=0$, which is confined to a relatively narrow pressure interval ($P_Z-P_Y\approx 8$ kbar), is therefore close to a straight line (as was supposed above). Another basic line, $B(P,T)=0$, which has a much smaller slope, continues to the higher-pressure region. On very qualitative grounds one can expect the bulk modulus (~ 200 kbar) to be an appropriate pressure scale in the expansion (15) of B, but because the linear term is relatively small ($k_B\approx 0.25k_A$ at $P=P_Y$), the nonlinear dependence on P may become important at much lower pressure. Indeed, inserting $\Delta\mu$ from Eq. (23) into the equation (5) of the melting curve, we find (for the calculated values of the parameters) a substantial deviation from the measured curve already at $P\sim 33$ kbar. We therefore take into account the quadratic term $\sim(P-P_Y)^2$ in Eq. (15):

$$B(P,T)\approx\gamma\left[T-T_Y-k_B(P-P_Y)\left(1-\frac{P-P_Y}{Q}\right)\right]. \tag{27}$$

Making the corresponding changes in $\Delta\mu$, we get a reasonably good fit to the melting curve for $Q=40$ kbar (curve *2* in Fig. 2). It is seen that for this Q the line $B(P,T)=0$, which has a slope $k_B=12$ K/kbar at $P=P_Y$ (≈ 17 kbar), approaches a maximum of $T=670$ K at $P=37$ kbar. Within this model only small (5–10%) corrections to the above-determined parameters P_0, ν, and $\Delta\kappa$ are needed to maintain an equally good quantitative description simultaneously in both domains of the phase diagram for which reliable data are available: around the melting curve minimum and at $T<T_Y$, where the first-order transition occurs.

4. The coefficient β in Eq. (21) for u^2 does not appear explicitly in Eq. (23) for $\Delta\mu$, and microscopic information must be invoked to determine β. In La and Pr the observed distortion of the fcc lattice is related to the softening of the zone-boundary transverse mode (the so-called point L in the Brillouin zone).[11,13] The dispersion anomalies of this mode and the frequencies $\omega(L)$ measured in Ce (Ref. 16) and in La (Ref. 17 and 18) are almost identical, and one can expect that the same mechanism drives the lattice distortion in both metals. Supposing this is the case, we can estimate β using the value of $\omega(L)$ measured at $P=0$. Extrapolating Eq. (13) (for $P'=P_Y, T'=T_Y$) to $P=0$, we have in the

harmonic approximation: $A=\beta P_Y=M\omega^2/2$ (M is the atomic mass). An appreciable softening of the transverse L mode in Ce from $T\sim 900$ K to room temperature has been observed to occur.[19] We take roughly $\omega=2\pi\cdot 10^{12}$ sec^{-1} for $T=T_Y$. For $P_Y\approx 17$ kbar we get $\beta\approx 30$ Å, and

$$u^2=\ell_0^2 f(p,t),\quad \ell_0^2=\frac{2P_0(\Delta\kappa)_0}{\beta}\approx 0.1\ \text{Å}^2, \tag{28}$$

which is comparable with the displacements measured in Pr.[13]

It is worth pointing out, that the aforementioned instability results in a distorted lattice having at least two nonequivalent atomic positions in the unit cell. Therefore each atom is displaced from the centrosymmetric position; this indicates that an ionic core polarization is the source for the lattice instability. Followed by a mixing of electronic orbitals (having different local parity) this polarization gives rise to the softening of the phonon modes and to a strong modification of the electronic properties.

We wish to thank A. S. Ioselevich, I. Luk'yanchuk, K. U. Neumann, Ye. G. Ponyatovskii, and K. R. A. Ziebeck for fruitful discussions.

[1] J. Staun Olsen, L. Gerward, U. Benedict and J.-P. Itié, Physica B **133**, 129 (1985).
[2] Y. C. Zhao and W. B. Holzapfel, J. Alloys Compounds **246**, 216 (1997).
[3] D. A. Koskenmaki and K. A. Gschneidner Jr., in *Handbook on the Physics and Chemistry of Rare Earths*, edited by K. A. Gschneidner Jr. and L. Eyring, North-Holland, Amsterdam, 1978, Chap. 4.
[4] A. Jayaraman, Phys. Rev. **137**, A179 (1965).
[5] Ye. G. Ponyatovskii, Dokl. Akad. Nauk SSSR **120**, 10221 (1958).
[6] R. I. Beecroft and S. A. Swenson, J. Phys. Chem. Solids **15**, 234 (1960).
[7] B. L. Davis and L. H. Adams, J. Phys. Chem. Solids **25**, 379 (1964).
[8] L. D. Landau, Phys. Z. Sowjetunion **8**, 113 (1935); **11**, 26 (1937).
[9] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, Oxford, 1958.
[10] R. B. Griffiths, Phys. Rev. Lett. **24**, 715 (1970).
[11] W. A. Grosshans, Y. K. Vohra, and W. B. Holzapfel, Phys. Rev. Lett. **49**, 1572 (1982).
[12] F. Porsch and W. B. Holzapfel, Phys. Rev. Lett. **70**, 4087 (1993).
[13] N. Hamaya, Y. Kakamoto, H. Fujihisa *et al.*, J. Phys.: Condens. Matter **5**, L369 (1993).
[14] Y. C. Zhao, F. Porsch, and W. B. Holzapfel, Phys. Rev. B **52**, 134 (1995).
[15] L. D. Livshitz, Yu. S. Genshaft, and V. K. Markof, Zh. Eksp. Teor. Fiz. **43**, 1262 (1962) [Sov. Phys. JETP **16**, 894 (1963)].
[16] C. Stassis, T. Gould, O. D. McMasters *et al.*, Phys. Rev. B **19**, 5746 (1979).
[17] C. Stassis, C.-K. Loong, and J. Zaretsky, Phys. Rev. B **26**, 5426 (1982).
[18] C. Stassis, G. Smith, B. N. Harmon *et al.*, Phys. Rev. B **31**, 6298 (1985).
[19] C. Stassis, C.-K. Loong, G. D. MacMasters, and R. M. Nicklow, Phys. Rev. B **25**, 6485 (1982).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Weak antilocalization in a 2D electron gas with chiral splitting of the spectrum

M. A. Skvortsov

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117940 Moscow, Russia

(Submitted 15 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 118–123 (25 January 1998)

Motivated by the recent observation of the metal–insulator transition in Si MOSFETs, a study is made of the quantum interference correction to the conductivity in the presence of the Bychkov–Rashba spin splitting. For a small splitting, a crossover from the localizing to antilocalizing regime is obtained. The antilocalization correction vanishes, however, in the limit of a large separation between the chiral branches. The relevance of the chiral splitting for the 2D electron gas in Si MOSFETs is discussed. © *1998 American Institute of Physics.*
[S0021-3640(98)00602-1]

PACS numbers: 72.15.Rn, 73.50.Bk, 71.10.Ca

Since the appearance of the scaling theory of localization[1] in 1979, it has been a common belief that there can be no metal–insulator transition (MIT) in 2D electron systems, since all the states are localized at arbitrary weak disorder. Recent experiments on high-mobility Si MOSFETs by Kravchenko *et al.*,[2] however, showed evidence for a zero-magnetic-field MIT which is controlled by the density n_s of 2D carriers. For small densities $n_s<n_c\simeq 10^{11}$ cm^{-2} the system is insulating, with an exponentially diverging resistivity in the limit $T\to 0$, whereas for n_s higher than the critical density a strong drop in resistivity (by one order of magnitude) is observed for $T<2$ K.

The origin of the new metallic phase is not yet understood. Nevertheless, it is evident that the electron–electron interaction plays an important role, as the critical density n_c is quite low, so that the Coulomb interaction dominates the kinetic energy. Their ratio is $r_s\simeq 10$ at the transition point and decreases $\propto n_s^{-1/2}$ deep in the metallic phase. Several theoretical approaches to the treatment of the strong Coulomb interaction, such as p-wave,[3] triplet,[4] or anyon[5] superconductivity and superconductivity resulting from a negative dielectric function[6] have been suggested during the last year.

Besides a strong Coulomb interaction, Si MOS structures are characterized by a spin–orbit splitting of the spectrum.[7] It originates from a strong asymmetry of the confining potential $V(z)$ of the quantum well. The corresponding term in the Hamiltonian of a 2D electron gas, the so-called Bychkov–Rashba term, is given by[8]

$$H_{so}=\alpha[\hat{\boldsymbol{\sigma}}\times\hat{\mathbf{p}}]. \tag{1}$$

Here $\hat{\vec{\sigma}}$ is the vector of the Pauli matrices, $\hat{\mathbf{p}}$ is the 2D momentum operator, α is a constant of the spin–orbit symmetry breaking measured in the units of velocity, and $[\cdot\times\cdot]$ stands for the z component of the vector product. This term lifts the spin degeneracy at zero magnetic field and results in the splitting of the spectrum into two chiral branches:

$$\epsilon_{\pm}^{(0)}(p)=\frac{p^2}{2m}\pm\alpha p, \tag{2}$$

with the splitting growing linearly with p.

For a Si MOSFET, the minimum of the spectrum (2), $-\epsilon_0=-m\alpha^2/2$, is estimated as 1 K,[7,9] while the Fermi energy is $\epsilon_F\simeq 6$ K at the transition. Then the ratio of the concentrations of left- and right-chiral fermions is $n_+/n_-=(\sqrt{\epsilon_F+\epsilon_0}+\sqrt{\epsilon_0})^2/(\sqrt{\epsilon_F+\epsilon_0}-\sqrt{\epsilon_0})^2\simeq 5$. Thus we conclude that the spin splitting results in a drastic change of the internal properties of the system even without allowing for the Coulomb interaction. This observation may call into question the remark by Belitz and Kirkpatrick[4] that the spin–orbit scattering is irrelevant due to the presence of a long-ranged Coulomb interaction. The latter should be strongly modified by the predominance of one type of chirality.

The relevance of the spin correlations was also demonstrated in magnetic measurements.[10] Magnetic field applied in the 2D plane was shown to suppress the metallic state, leading to a huge increase in resistivity. Measurements in a perpendicular magnetic field show a large positive magnetoresistance at high densities $n_s>2n_c$, also indicating the spin-related origin of the conducting phase.

We argue that an understanding of the new conducting phase and of the MIT itself can hardly be obtained without taking the strong chiral splitting into account. Thus the theory of the metallic state should be the theory of Coulomb-interacting chiral fermions. The necessary first step, then, is to consider noninteracting particles with a chiral splitting of the spectrum.

In this letter we study the first quantum correction to the conductivity for noninteracting particles in the presence of the Bychkov–Rashba term (1) and obtain it as a function of the spin–orbit splitting. There are three energy scales in the problem: the first is the Fermi energy ϵ_F, the second is the chiral splitting $\Delta=2\alpha p_F$ between the two branches (2) at the Fermi level, and the third is the inverse elastic mean free time τ^{-1} introduced by disorder. We will assume ϵ_F to be the largest energy scale:

$$\epsilon_F\gg\frac{1}{\tau},\quad \epsilon_F\gg\Delta. \tag{3}$$

The relationship between Δ and τ^{-1} is not specified, so that the variable

$$x=\Delta\tau \tag{4}$$

that controls the strength of the chiral splitting may vary from 0 to ∞ provided that the relations (3) are fulfilled. At the critical density, the ratio Δ/ϵ_F is of the order of 1, but it decreases as n_s^{-1} into the metallic phase. The experimental value of the parameter x depends slightly on the density, varying from 5 to 10 as n_s varies from 10^{11} cm^{-2} to 3×10^{12} cm^{-2}.

The spin–orbit scattering in a *random potential* is known to drive the system into the symplectic ensemble, resulting in an antilocalizing correction to the conductivity $\Delta\sigma_{\rm symp}=(e^2/\pi h)\ln(l_\varphi/l)$ (Ref. 11), where l is the mean free path, $l_\varphi=(D\tau_\varphi)^{1/2}$ is the phase-breaking length associated with the phase relaxation time τ_φ, and D is the diffusion coefficient. In the case of the Bychkov–Rashba term, SU(2) symmetry is broken on the level of the *regular Hamiltonian*, while the potential scattering may be considered as spin independent. From symmetry considerations one might expect that the symplectic correction $\Delta\sigma_{\rm symp}$ should be recovered in the limit of a large spin splitting. We will see, however, that the correction becomes antilocalizing at $x=(l_\varphi/l)^{1/3}\ll 1$, nearly approaches $\Delta\sigma_{\rm symp}$ for $x\leqslant 1$, but *vanishes* for $x\gg 1$. Such a peculiar behavior is due to the presence of the two chiral branches that are well separated in the limit $\Delta\gg\tau^{-1}$.

Weak localization effects in the presence of different types of spin–orbit splitting, including the Bychkov–Rashba splitting, have been studied extensively in Refs. 12. However, the authors were interested mainly in the behavior of the magnetoresistance, while the quantum correction at zero magnetic field and for $x\geqslant 1$ when H_{so} cannot be treated as a small perturbation had not been investigated.

We consider a 2D noninteracting electron gas with the Bychkov–Rashba term in the Hamiltonian:

$$H=\frac{\hat{\mathbf{p}}^2}{2m}+\alpha\hat{p}_y\sigma_x-\alpha\hat{p}_x\sigma_y+U(\mathbf{r}), \tag{5}$$

where $U(\mathbf{r})$ is a random spin-independent impurity potential, which for the sake of simplicity is assumed to be Gaussian δ-correlated: $\langle U(\mathbf{r})U(\mathbf{r}')\rangle=\delta(\mathbf{r}-\mathbf{r}')/2\pi\nu\tau$. Here $\nu=m/2\pi$ is the density of states for the free Hamiltonian $\mathbf{p}^2/2m$.

The classical conductivity can easily be shown to be independent of x and given by the Drude formula $\sigma_0=ne^2\tau/m$, provided that the random potential is δ-correlated. The first quantum correction to the conductivity[13] is given by the expression[a)]

$$\Delta\sigma=-\frac{e^2}{h}\frac{v_F^2}{2}\int\frac{d^2\mathbf{p}}{(2\pi)^2}\langle G^R(\mathbf{p})\rangle^{\rho\alpha}\langle G^R(\mathbf{p})\rangle^{\lambda\sigma}\langle G^A(\mathbf{p})\rangle^{\sigma\beta}\langle G^A(\mathbf{p})\rangle^{\mu\rho}\int_{1/l_\varphi}^{1/l}\frac{d^2\mathbf{q}}{(2\pi)^2}C^{\alpha\lambda}_{\beta\mu}(\mathbf{q}), \tag{6}$$

where $\langle G^{R,A}\rangle$ are disorder-averaged retarded (advanced) Green functions, which for our problem are nondiagonal in the spin space, and the static cooperon $C(\mathbf{q})$ is determined by the ladder equation

$$C^{\alpha\lambda}_{\beta\mu}(\mathbf{q})=\frac{\delta^{\alpha\lambda}\delta^{\beta\mu}}{2\pi\nu\tau}+\frac{1}{2\pi\nu\tau}\int\frac{d^2\mathbf{p}}{(2\pi)^2}\left\langle G^R\left(\mathbf{p}+\frac{\mathbf{q}}{2}\right)\right\rangle^{\alpha\alpha'}\left\langle G^A\left(-\mathbf{p}+\frac{\mathbf{q}}{2}\right)\right\rangle^{\beta\beta'}C^{\alpha'\lambda}_{\beta'\mu}(\mathbf{q}). \tag{7}$$

The averaged Green function obeys the Dyson equation $\langle G(\mathbf{p})\rangle^{-1}=G^{(0)}(\mathbf{p})^{-1}-\Sigma$, where $G^{(0)}(\mathbf{p})$ is the Green function of the unperturbed Hamiltonian. In the quasiclassical limit, $\epsilon_F\tau\gg 1$, only diagrams without intersections of impurity lines are important, and the self-energy function is

$$\Sigma^{\alpha\beta}_{R,A}=\frac{1}{2\pi\nu\tau}\int\frac{d^2\mathbf{p}}{(2\pi)^2}\langle G^{R,A}(\mathbf{p})\rangle^{\alpha\beta}.$$

On solving the Dyson equation we obtain the Green function that can be written near the poles as ($\mathbf{n}=\mathbf{p}/p$)

$$\langle G^{R,A}(\mathbf{p})\rangle = \frac{-\xi(p)\pm\dfrac{i}{2\tau}+\Delta(n_y\sigma_x-n_x\sigma_y)/2}{\left(-\xi(p)-\dfrac{\Delta}{2}\pm\dfrac{i}{2\tau}\right)\left(-\xi(p)+\dfrac{\Delta}{2}\pm\dfrac{i}{2\tau}\right)}. \tag{8}$$

Here we have taken advantage of the fact that $\Delta\ll\epsilon_F$ and replaced αp by $\Delta/2$. The relaxation times for the two chiral branches appear to be equal to each other and coincide with the mean free time τ. This is a consequence of the model with δ-correlated disorder. For a more realistic model with a finite correlation length the lifetimes will be different for the two chiralities, but the difference will be small in the limit $\Delta\ll\epsilon_F$.

The crucial quantity that determines the spin structure of the cooperon is the integral of the retarded and advanced Green functions,

$$I^{\alpha\alpha'}_{\beta\beta'}(\mathbf{q}) = \frac{1}{2\pi\nu\tau}\int\frac{d^2\mathbf{p}}{(2\pi)^2}\left\langle G^R\left(\mathbf{p}+\frac{\mathbf{q}}{2}\right)\right\rangle^{\alpha\alpha'}\left\langle G^A\left(-\mathbf{p}+\frac{\mathbf{q}}{2}\right)\right\rangle^{\beta\beta'}. \tag{9}$$

Calculating this integral as a function of x, expanding to second order in q, and substituting into Eq. (7), we get

$$\hat{C}(\mathbf{q}) = \frac{\hat{A}^{-1}(\mathbf{q})}{2\pi\nu\tau}, \tag{10}$$

where the operator $\hat{A}(\mathbf{q})=\hat{\mathbf{1}}-\hat{I}(\mathbf{q})$ expressed in terms of the total cooperon spin $\hat{\mathbf{S}}=\frac{1}{2}(\hat{\sigma}^R+\hat{\sigma}^A)$ reads

$$\hat{A}(\mathbf{q}) = \frac{1}{2}q^2l^2+x^2\left(\frac{1}{2(1+x^2)}-\frac{6+3x^2+x^4}{8(1+x^2)^3}q^2l^2\right)(\hat{\mathbf{S}}^2-\hat{S}_z^2)$$
$$-\frac{x^2(6+3x^2+x^4)}{4(1+x^2)^3}(\mathbf{q}\times\hat{\mathbf{S}})^2l^2-\frac{x}{(1+x^2)^2}(\mathbf{q}\times\hat{\mathbf{S}})l. \tag{11}$$

The next step is to invert the matrix $\hat{A}$ and to obtain the cooperon. According to Eq. (11), the singlet mode is gapless while the triplet sector acquires a gap proportional to x. To study the lifting of the triplet sector let us first consider the case of small $x\ll 1$. Then, for $ql\gg x$, the spin structure of $\hat{A}$ may be neglected, so that $\hat{A}^{-1}=(2/q^2l^2)\hat{\mathbf{1}}$. For $ql\leqslant x$, the triplet sector of the inverse operator $\hat{A}^{-1}$ becomes complicated, with different triplet modes having different gaps because of the low symmetry of Eq. (11), but this region does not contribute to the logarithmic integral over q. So we may write

$$\hat{A}^{-1}\simeq\frac{2}{q^2l^2}\left(1-\frac{\hat{\mathbf{S}}^2}{2}\right)+\frac{2}{q^2l^2+x^2}\frac{\hat{\mathbf{S}}^2}{2}. \tag{12}$$

Although this is not an exact formula it captures correctly the logarithmically large terms in the q integration.

Inserting (12) to Eq. (10) and performing the integration, we obtain the expression for the cooperon integral:

$$\int_{1/l_\varphi}^{1/l} \frac{d^2\mathbf{q}}{(2\pi)^2} C_{\beta\mu}^{\alpha\lambda}(\mathbf{q}) = \frac{1}{8\pi^2 \nu v_F^2 \tau^3} \left\{ \left(\ln\frac{l_\varphi}{l} + 3f \right) \delta^{\alpha\lambda}\delta^{\beta\mu} - \left(\ln\frac{l_\varphi}{l} - f \right) \sum_{i=1}^{3} \sigma_i^{\alpha\lambda}\sigma_i^{\beta\mu} \right\}, \tag{13}$$

where the contribution of the triplet sector is

$$f\left(x, \frac{l_\varphi}{l}\right) = \begin{cases} \ln\dfrac{l_\varphi}{l} & \text{for } x \ll \dfrac{l}{l_\varphi}; \\ \ln\dfrac{1}{x} & \text{for } \dfrac{l}{l_\varphi} \ll x \ll 1; \\ O(1) & \text{for } x \gg 1. \end{cases} \tag{14}$$

The last thing to do is to compute the integral of four Green functions in Eq. (6):

$$\int \frac{d^2\mathbf{p}}{(2\pi)^2} \langle G^A(\mathbf{p})\rangle^{\mu\rho} \langle G^R(\mathbf{p})\rangle^{\rho\alpha} \langle G^R(\mathbf{p})\rangle^{\lambda\sigma} \langle G^A(\mathbf{p})\rangle^{\sigma\beta}$$
$$= \frac{4\pi\nu\tau^3}{1+x^2} \left[\left(1 + \frac{x^2}{2}\right) \delta^{\mu\alpha}\delta^{\lambda\beta} + \frac{x^2}{4} (\sigma_x^{\mu\alpha}\sigma_x^{\lambda\beta} + \sigma_y^{\mu\alpha}\sigma_y^{\lambda\beta}) \right]. \tag{15}$$

This integral is diagonal in the spin space for small $x \ll 1$ but has a more complex structure for $x \gg 1$, when the chiral branches are well separated.

Finally, we combine all together. Substituting (13) and (15) into Eq. (6), after some arithmetic with the Pauli matrices we obtain the final expression

$$\Delta\sigma = \frac{2e^2}{\pi h} \frac{1}{1+x^2} \left[\frac{1}{2} \ln\frac{l_\varphi}{l} - \left(\frac{3}{2} + x^2 \right) f\left(x, \frac{l_\varphi}{l}\right) \right]. \tag{16}$$

Let us study $\Delta\sigma$ as a function of x for a given $l_\varphi \gg l$. For $x \ll l/l_\varphi$, the spin splitting can be neglected and we obtain the orthogonal universality class correction $\Delta\sigma_{\text{orth}}$ which can be interpreted as a sum of a localizing contribution from the triplet sector and an antilocalizing contribution from the singlet sector. Then for $l/l_\varphi \ll x \ll 1$ the triplet modes acquire a gap that reduces their contribution, and the total correction changes sign and becomes antilocalizing at

$$x_* = \left(\frac{l}{l_\varphi} \right)^{1/3}. \tag{17}$$

For $x_* \ll x \leqslant 1$, the antilocalization becomes more pronounced, nearly approaching $\Delta\sigma_{\text{symp}}$. However, for $x \geqslant 1$ it goes rapidly to zero (as x^{-2}). Summarizing, we present the behavior of $\Delta\sigma$ in the form

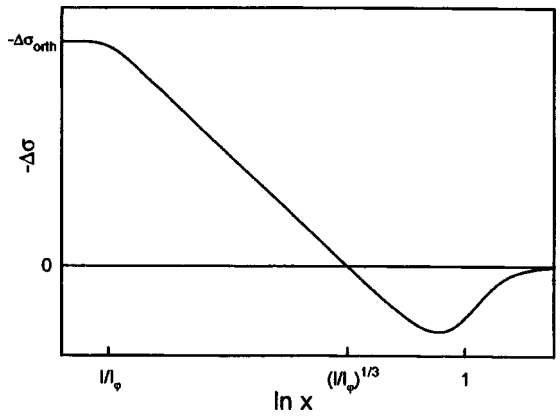

FIG. 1. A sketch of $\Delta\sigma$ versus the strength x of the chiral splitting; $\Delta\sigma_{\rm orth}=-(2e^2/\pi h)\ln(l_\varphi/l)$.

$$\Delta\sigma=\frac{2}{\pi}\frac{e^2}{h}\begin{cases}-\ln\dfrac{l_\varphi}{l} & \text{for } x\ll\dfrac{l}{l_\varphi};\\ \dfrac{1}{2}\ln\left(\dfrac{l_\varphi}{l}x^3\right) & \text{for } \dfrac{l}{l_\varphi}\ll x\ll 1;\\ \dfrac{1}{2x^2}\ln\dfrac{l_\varphi}{l} & \text{for } x\gg 1.\end{cases}\tag{18}$$

The crossover from the orthogonal to the symplectic corrections obtained for $x\ll 1$ is related to the appearance of a gap in the triplet sector of the cooperon. On the other hand, the reduction of $\Delta\sigma$ for $x\gg 1$ must be attributed to the spin structure of the integral (15), which annihilates the singlet cooperon mode in the limit of a large splitting between the chiral branches. In other words, the result obtained implies the absence of the first quantum correction to the conductivity in a system of 2D chiral fermions with only one sort of chirality. The other example where a certain type of spin–orbit coupling leads to the absence of the first interference correction was considered in Ref. 14. The behavior of $\Delta\sigma$ as a function of x is sketched in Fig. 1.

The large-x asymptotic behavior can be traced up to $x\sim\sqrt{\ln(l_\varphi/l)}$. In order to find $\Delta\sigma$ for even larger x, one has to go beyond the diffusion approximation to calculate the function $f(x)$ that competes with the vanishing term $\ln(l_\varphi/l)/x^2$.

In conclusion, we have considered the quantum interference correction to the conductivity of noninteracting fermions in the presence of the Bychkov–Rashba spin–orbit interaction. At small chiral splittings, $x<1$, the correction changes sign and becomes antilocalizing. It vanishes for $x\gg 1$, where the scattering between the different chiralities is strongly suppressed. The present theory may be regarded as a step toward understanding of the conducting phase in Si MOSFETs, which likely contain Coulomb-interacting chiral fermions. It might also explain the low-temperature log T behavior of the resistivity obtained for some samples below 300 mK deep in the metallic phase. The correction is antilocalizing, $\Delta\sigma\simeq -C(e^2/h)\ln(T/T_0)$, with very small $C\sim 10^{-2}$ (Ref. 9); this is consistent with our formula for experimental values $x\gg 1$.

It is a pleasure for me to thank V. M. Pudalov for the idea of this work and M. V. Feigel'man for helpful discussions. Support from INTAS–RFBR Grant 95-0302 and the Swiss National Science Foundation Collaboration Grant 7SUP J048531 is gratefully acknowledged. I acknowledge that this material is based upon work supported by U. S. Civilian Research and Development Foundation (CRDF) under Award #RP1-273.

[a)]In the presence of the Bychkov–Rashba term, velocity is no longer proportional to momentum: $\hat{v}_i = \hat{p}_i/m + \alpha \epsilon_{ij}\sigma_j$. This should modify Eq. (6), but the corrections are small in the limit $\Delta/\epsilon_F \ll 1$.

[1] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[2] S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux *et al.*, Phys. Rev. B **50**, 8039 (1994); S. V. Kravchenko, W. E. Mason, G. E. Bowker *et al.*, Phys. Rev. B **51**, 7038 (1995); S. V. Kravchenko, D. Simonian, M. P. Sarachik *et al.*, Phys. Rev. Lett. **77**, 4938 (1996).
[3] P. Phillips, Y. Wan, I. Martin, S. Knysh, and D. Dalidovich, http://xxx.lanl.gov/abs/cond-mat/9709168.
[4] D. Belitz and T. R. Kirkpatrick, http://xxx.lanl.gov/abs/cond-mat/9705023.
[5] F. C. Zhang and T. M. Rice, http://xxx.lanl.gov/abs/cond-mat/9708050.
[6] P. Phillips and Y. Wan, http://xxx.lanl.gov/abs/cond-mat/9704200.
[7] V. M. Pudalov, http://xxx.lanl.gov/abs/cond-mat/9707076; JETP Lett. **66**, 175 (1997).
[8] Yu. A. Bychkov and E. I. Rashba, JETP Lett. **39**, 78 (1984).
[9] V. M. Pudalov, private communication.
[10] V. M. Pudalov, G. Brunthaler, A. Prinz, and G. Bauer, http://xxx.lanl.gov/abs/cond-mat/9707054; JETP Lett. **65**, 932 (1997) [Erratum **66**, 131 (1997)]; D. Simonian, S. V. Kravchenko, M. P. Sarachik, and V. M. Pudalov, Phys. Rev. Lett. **79**, 2304 (1997).
[11] S. Hikami, A. I. Larkin, and Y. Nagaoka, Prog. Theor. Phys. **63**, 707 (1980); S. Hikami, Prog. Theor. Phys. **64**, 1425 (1980).
[12] W. Knap *et al.*, Phys. Rev. B **53**, 3912 (1996); S. V. Iordanskii, Yu. B. Lyanda-Geller, and G. E. Pikus, JETP Lett. **60**, 206 (1994).
[13] P. W. Anderson, E. Abrahams, and T. V. Ramakrishnan, Phys. Rev. Lett. **43**, 718 (1979); L. P. Gor'kov, A. I. Larkin, and D. E. Khmelnitskii, JETP Lett. **30**, 228 (1979).
[14] Yu. B. Lyanda-Geller and A. D. Mirlin, Phys. Rev. Lett. **72**, 1894 (1994).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Critical velocity and event horizon in pair-correlated systems with ''relativistic'' fermionic quasiparticles

N. B. Kopnin and G. E. Volovik

Helsinki University of Technology, Low Temperature Laboratory, P. O. Box 2200, FIN-02015 HUT, Finland; Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia

(Submitted 16 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 124–129 (25 January 1998)

The condition for the appearance of an event horizon is considered in pair-correlated systems (superfluids and superconductors) in which the fermionic quasiparticles obey ''relativistic'' equations. In these systems the Landau critical velocity of superflow corresponds to the speed of light. In conventional systems, such as s-wave superconductors, the superflow remains stable even above the Landau threshold. We show, however, that, in ''relativistic'' systems, the quantum vacuum becomes unstable and the superflow collapses after the ''speed of light'' is reached, so that the horizon cannot appear. Thus an equilibrium dissipationless superflow state and the horizon are incompatible on account of quantum effects. This negative result is consistent with the quantum Hawking radiation from the horizon, which would lead to a dissipation of the flow. © *1998 American Institute of Physics.*
[S0021-3640(98)00702-6]

PACS numbers: 67.20.+k, 74.20.−z, 04.60.−m

1. It is known that some aspects of the problem of black holes can be modeled in condensed matter physics.[1–5] This comes from the fact that acoustic waves propagating in a moving classical liquid[1–3] and fermions propagating in a texture of superfluid ^{3}He-A (Refs. 4, 6, and 7) obey relativistic-type equations in a curved space whose metric is produced by both the flow field and the texture (in ^{3}He-A). In both systems the corresponding velocity of light can be exceeded, which affords the possibility of investigating the event horizon problem.

^{3}He-A and other pair-correlated systems (including d-wave superconductors, which also contain relativistic fermions; see, e.g., Ref. 8) are better models for simulations of the event horizon than a classical liquid, since these are quantum systems with ground states that are in many respects similar to the quantum vacuum of high-energy physics. That is why they can be used for investigating the quantum effects related with the event horizon, such as Hawking radiation[9] and statistical entropy.[10]

Here we address the stability problem of a quantum vacuum in the presence of an event horizon: whether a nondissipative flow of superfluid is possible in the presence of a horizon, or whether the horizon always leads to a vacuum reconstruction into a state

with dissipation. This can be considered using an example of a superflow with velocity exceeding the Landau critical velocity v_L.

Let the superfluid move at $T=0$ with a superfluid velocity $\mathbf{v}_s$, while the container walls, i.e., the preferred reference frame, move with a normal velocity $\mathbf{v}_n$. The physical properties of the vacuum state depend on the relative (counterflow) velocity $\mathbf{w}=\mathbf{v}_s-\mathbf{v}_n$. In a subcritical regime, $w<v_L$, the order parameter is independent of w: the observer moving with $\mathbf{v}_s$ does not see any difference in the liquid as compared to the case when $\mathbf{v}_s=\mathbf{v}_n$. The system thus retains some kind of Galilean invariance even in the presence of the container wall. For example, the density of the superfluid component in the expression for the mass current $\mathbf{j}=\rho_s\mathbf{v}_s+\rho_n\mathbf{v}_n$, which represents the vacuum component of the liquid, does not depend on w and is equal to the total density: $\rho_s(T=0,\ w<v_L)=\rho$, while the density of the normal component, which represents the matter, is always zero: $\rho_s(T=0,\ w<v_L)=\rho-\rho_s(T=0,\ w<v_L)=0$.

The observer starts to see a dependence on the velocity relative to the reference frame if w exceeds the Landau critical velocity v_L, at which negative energy levels appear, i.e., states with negative Doppler-shifted energy $E_{\mathbf{p}}+\mathbf{p}\cdot\mathbf{w}<0$. If the system is fermionic, the typical situation in a supercritical flow regime above v_L is as follows: The negative energy levels become finally occupied and, after such vacuum reconstruction, the system comes again into a new equilibrium state with a frictionless superflow. Now, however, all the physical quantities depend on w, as seen by a comoving observer. The vacuum state becomes anisotropic and the superfluid density now depends on w and becomes smaller than the total mass density, $\rho_s(T=0,w>v_L)<\rho$. The other part of the liquid, with the so-called normal density $\rho_n=\rho-\rho_s$, comprises the normal component (matter) and consists of trapped fermions with negative energy. In equilibrium, this component is at rest with respect to the container reference frame, i.e., it moves with the velocity $\mathbf{v}_n$. The previous ''Galilean'' symmetry is thus broken by the created matter.

There is another critical velocity, v_c, at which the superfluid vacuum is exhausted, i.e., the superfluid density completely disappears, $\rho_s(w=v_c)=0$, and thus the nondissipative superflow does not exist any more. Typically, $v_c>v_L$ and thus the violation of the Galilean symmetry occurs earlier than the superfluidity collapses. This happens, for example, in conventional s-wave superconductors, where v_L and v_c are of the same order, with $v_c>v_L$, and also in superfluid ^{3}He-A, where $v_L=0$ while v_c is finite (see, e.g., Refs. 11–13).

The existence of a region with stable supercritical superflow, $v_L<w<v_c$, allows us to raise the question of the quantum effects of an event horizon. Let us consider a system in which the quasiparticles are described by effective relativistic equations, such that the Landau velocity corresponds to the ''speed of light'' c. In this case a supercritical velocity $w>v_L$ corresponds to a superluminal velocity, and thus an event horizon can be constructed (see below). If quantum effects are taken into account, there arises the problem that on the one hand the frictionless superflow is stable in the regime $v_L<w<v_c$, and this stability cannot be destroyed by the presence of the horizon, while on the other hand, the Hawking radiation from the horizon means that such a superflow is dissipative in the presence of the horizon.

Thus we have a dilemma: (i) either one should doubt the fundamentality of the Hawking radiation from the event horizon in the supercritical regime, or (ii) in the

relativistic-type systems the ''superluminal'' regime of superflow is prohibited, i.e., $v_L = v_c = c$. Here we consider a pair-correlated fermionic system with a superconducting/superfluid state of the polar type, which has ''relativistic'' Bogoliubov fermionic quasiparticles. We find that in this system the second alternative occurs: a nondissipative superflow collapses at $w = v_L = c$, which means that the horizon never appears in the stationary nondissipative superflow: it can only exist in a dissipative flow state.

2. The energy spectrum of the pair-correlated system and its vacuum state are determined by the self-consistent equations for the gap function $\Delta_{\mathbf{p}}$ which determines the quasiparticle spectrum and thus the ''speed of light'' c:

$$\Delta_{\mathbf{p}} = \sum_{\mathbf{p}'} V_{\mathbf{p},\mathbf{p}'} \frac{\Delta_{\mathbf{p}'}}{E_{\mathbf{p}'}} (1 - n_{\mathbf{p}'} - n_{-\mathbf{p}'}). \tag{1}$$

Here $V_{\mathbf{p},\mathbf{p}'}$ is the pairing potential, $E_{\mathbf{p}}$ is the energy of the quasiparticles in the pair-correlated state and $n_{\mathbf{p}}$ is their thermal distribution

$$E_{\mathbf{p}} = \sqrt{\Delta_{\mathbf{p}}^2 + \epsilon_{\mathbf{p}}^2}, \quad n_{\mathbf{p}} = \frac{1}{1 + \exp[(E_{\mathbf{p}} + \mathbf{p}\cdot\mathbf{w})/T]}, \tag{2}$$

and $\epsilon_{\mathbf{p}} = (p^2 - p_F^2)/2m$ is the fermion energy in the absence of the pair correlation. If the superflow velocity $\mathbf{v}_s$ deviates from the container reference frame velocity $\mathbf{v}_n$, the distribution function is Doppler shifted; further we assume that $\mathbf{v}_n = 0$ and thus $\mathbf{w} = \mathbf{v}_s$.

Let us consider how the vacuum state ($T = 0$) is disturbed by the counterflow w in the supercritical regime. We are interested in the case when the spectrum of quasiparticles is ''relativistic,'' so that the horizon problem can arise. For this reason we consider the two-dimensional case, spin-triplet pairing with orbital momentum $L = 1$, for which the pairing potential $V_{\mathbf{p},\mathbf{p}'} = 2(V_1/p_F^2)\mathbf{p}\cdot\mathbf{p}'$, and the gap function corresponding to the polar phase is[11]

$$\Delta_{\mathbf{p}} = c p_x, \tag{3}$$

where the factor c plays the role of the speed of light along the x axis.

Let the velocity of the superflow be along the same axis x, i.e., $\mathbf{v}_s = w\hat{x}$, so that the superflow does not break the symmetry of the polar state and Eq. (3) remains the solution even in the presence of the superflow. The Doppler shifted energy of the fermions in this pair-correlated state is

$$E(p_x, \epsilon) = E_{\mathbf{p}} + \mathbf{p}\cdot\mathbf{v}_s = \sqrt{\epsilon^2 + c^2 p_x^2} + w p_x \tag{4}$$

or

$$(E - w p_x)^2 = \epsilon^2 + c^2 p_x^2 \tag{5}$$

This corresponds to a relativistic 1D particle of mass ϵ moving in the Lorentzian metric:

$$g^{00} = -1, \quad g^{01} = w, \quad g^{11} = c^2 - w^2. \tag{6}$$

If the counterflow velocity w can exceed c, one can construct an inhomogeneous flow state with a coordinate dependence of $w(x)$ and $c(x)$ such that $w(x)$ crosses $c(x)$. In this case the metric element $g^{11}(x) = c^2 - w^2$ crosses zero at the points $x = x_h$ at which $w(x_h) = c(x_h)$, and thus the event horizon appears at these points.

From Eq. (1) one finds that there are no quasiparticles for $w<c$ and $T=0$: $n_{\mathbf{p}}=0$, i.e., the vacuum remains intact. The system is effectively Galilean invariant and the speed of light is independent of w.

The problem is whether the flow velocity w can exceed the Landau velocity v_L, which is now the ''speed of light'' c. If the answer is yes, the horizon can be constructed. Let us consider the gap equation (1) in the case when $w>c$. We assume that the speed of light is small compared to the Fermi velocity, $c \ll v_F = p_F/m$, which is typical for a weakly interacting Fermi liquid. In this case, the momentum is concentrated near the Fermi momentum, $\mathbf{p}=(p_F \sin\phi,\ p_F\cos\phi)$, and one can write

$$\sum_{\mathbf{p}} = \int \frac{d^2p}{(2\pi)^2} = \frac{m}{2\pi}\int d\epsilon \int \frac{d\phi}{2\pi}. \tag{7}$$

In principle, one can expect that at $w>c$ the speed of light becomes dependent on w for $w>c$. Thus let us introduce the bare speed of light $c_0=c(w=0)$ and the current (variable) value of the speed of light $c(w)$ if $w>c(w)$. As we have seen, $c(w<c)=c_0$, and this solution persists until w reaches the Landau velocity c_0. Thus the first branch of $c(w)$ is

$$c_1(w)=c_0, \quad w \leqslant c_0. \tag{8}$$

3. If $w>c$, the Galilean invariance becomes broken due to fermions filling the negative levels of the energy in Eq. (4). The number of particles on the negative energy levels is the fermionic step function of the energy

$$n_{\mathbf{p}}=\Theta(-E(p_x,\epsilon)). \tag{9}$$

From Eq. (1) one obtains the following equation for the factor $c(w)$ in the gap function (3):

$$\int_0^\infty d\epsilon \int_0^{2\pi} \frac{d\phi}{2\pi}\left(\frac{\sin^2\phi}{\sqrt{\epsilon^2+c^2\sin^2\phi}} - \frac{\sin^2\phi}{\sqrt{\epsilon^2+c_0^2\sin^2\phi}}\right)$$

$$= 2\int_0^{\pi} \frac{d\phi}{2\pi}\sin^2\phi \int_0^{\sin\phi\sqrt{w^2-c^2}} \frac{d\epsilon}{\sqrt{\epsilon^2+c^2\sin^2\phi}}. \tag{10}$$

Note that the details of pair interaction are concealed in the bare speed of light c_0, determined by this interaction. From Eq. (10) one has

$$\ln\frac{c_0}{c} = \sinh^{-1}\left(\sqrt{\frac{w^2}{c^2}-1}\right) \tag{11}$$

which gives the solution for the ''speed of light'' c in the superluminal regime $w>c(w)$:

$$c_2(w)=c_0\sqrt{\frac{2w}{c_0}-1}, \quad \frac{1}{2}c_0<w<c_0. \tag{12}$$

It follows that no solution exists above the Landau velocity, i.e., at $w>c_0$, which means that the Landau velocity coincides with the velocity of the superflow collapse and thus with the bare speed of light: $v_L=v_c=c_0$.

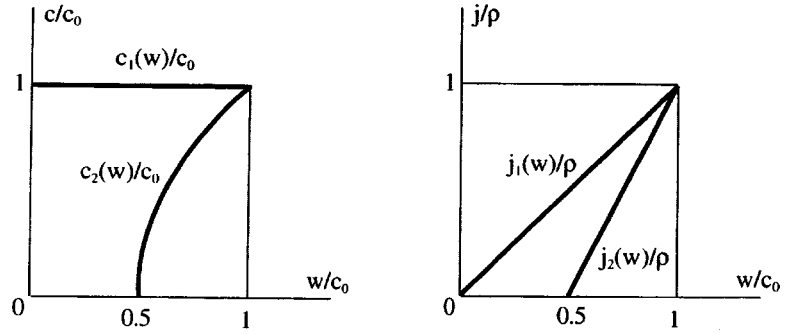

FIG. 1. Two states of the superflow. In the subluminal flow state the speed of light $c_1(w)=c_0$ does not depend on flow velocity w with respect to the container wall. The superluminal state with $c_2(w)<w$ is locally unstable. The mass current in these two states is presented on the right.

Below the Landau velocity, one has two branches, $c_1(w)=c_0$ and $c_2(w)$ (see Fig. 1). Both can be obtained as extrema of the superfluid vacuum energy in the presence of the mass current

$$\Omega_S(s,w)-\Omega_N=-jw+\frac{1}{2}\rho w^2-\frac{1}{4}\rho c^2-\frac{1}{2}\rho c^2\,\ln\frac{c_0}{c}$$
$$+\frac{1}{2}\rho c^2\left\{\ln\left[\frac{w}{c}+\sqrt{\frac{w^2}{c^2}-1}\right]-\frac{w}{c}\sqrt{\frac{w^2}{c^2}-1}\right\}\Theta(w-c). \tag{13}$$

Here Ω_N is the free energy of the normal state, i.e., at $c=0$; the mass density in this 2D model is $\rho=mp_F^2/2\pi\hbar^2$; the first term $-jw$ means that the free energy is to be extremized at the given mass current $\mathbf{j}=\rho\mathbf{v}_s+\Sigma_{\mathbf{k}}\mathbf{k}n_{\mathbf{k}}$. The current in a given state (see Fig. 1) can be obtained from the extremum of the vacuum energy with respect to w: $\partial\Omega_S/\partial w=0$. This gives the general expression for the mass current density

$$j(w,c)=\rho(w-\Theta(w-c)\sqrt{w^2-c^2}). \tag{14}$$

The second branch, corresponding to the superluminal flow $c_2(w)<w$, represents the saddle point solution of the vacuum energy and thus is unstable against the formation of the regular branch, corresponding to the subluminal flow, $c_1(w)=c_0<w$. This second branch with similar behavior has been also found for ^{3}He-B under superflow.[14]

In conclusion, we found that, in a superfluid analog of the relativistic system, a stable superflow with velocity exceeding the corresponding ''speed of light,'' $w>c$, does not exist, and thus the dissipationless state with a horizon does not appear. The collapse of the superfluid quantum vacuum in the superluminal regime is compatible with the Hawking radiation, which leads to the dissipation in the presence of a horizon and thus cannot exist in the stable superflow. A horizon can appear only if the flow state is dissipative. This can happen if the external body or the order parameter texture moves in the superfluid with a supercritical velocity, as was discussed in Ref. 4 for the case of a moving topological soliton in ^{3}He-A. The Hawking radiation gives rise to additional dissipation during the motion of the object.

GEV thanks Ted Jacobson for illuminating discussions. This work was supported by the Russian Fund for Fundamental Research under Grant 96-02-16072, by the Statistical Physics Program of the Russian Academy of Sciences, by INTAS Grant 96-0610, and by the European Science Foundation.

[1] W. G. Unruh, Phys. Rev. Lett. **46**, 1351 (1981); Phys. Rev. D **51**, 2827 (1995).
[2] T. Jacobson, Phys. Rev. D **44**, 1731 (1991).
[3] M. Visser, http://xxx.lanl.gov/abs/gr-qc/9311028, gr-qc/9712010.
[4] T. Jacobson and G. E. Volovik, ''Event horizons in ^{3}He,'' to be published.
[5] N. B. Kopnin and G. E. Volovik, ''Rotating vortex core: an instrument for detecting the core excitations,'' http://xxx.lanl.gov/abs/cond-mat/9706082.
[6] G. E. Volovik, *Exotic properties of superfluid* ^{3}*He*, World Scientific, Singapore–New Jersey–London–Hong Kong, 1992.
[7] G. E. Volovik, JETP Lett. **63**, 483 (1996)]; ''Simulation of quantum field theory and gravity in superfluid He-3,'' http://xxx.lanl.gov/abs/cond-mat/9706172.
[8] S. H. Simon and P. A. Lee, Phys. Rev. Lett. **78**, 1548 (1997).
[9] S. W. Hawking, Nature **248**, 30 (1974); Commun. Math. Phys. **43**, 199 (1975).
[10] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
[11] D. Vollhardt and P. Wölfle, *The Superfluid Phases of Helium 3*, Taylor & Francis, London, 1990.
[12] P. Muzikar, and D. Rainer, Phys. Rev. B **27**, 4243 (1983).
[13] K. Nagai, J. Low Temp. Phys. **55**, 233 (1984).
[14] D. Vollhardt, K. Maki, and N. Schopohl, J. Low Temp. Phys. **39**, 79 (1980).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Scanning tunneling spectroscopy of charge effects on semiconductor surfaces and atomic clusters

N. S. Maslova, S. I. Oreshkin, V. I. Panov, and S. V. Savinov
Chair of Quantum Radio Physics, Moscow State University, 119899 Moscow, Russia

A. Depuydt and C. Van Haesendonck
Laboratorium voor Vaste-Stoffysica en Magnetisme, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium

(Submitted 15 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 130–135 (25 January 1998)

We have used scanning tunneling microscopy and scanning tunneling spectroscopy at liquid helium temperature to study the electronic structure of *in situ* cleaved, (110) oriented surfaces of InAs single crystals. Both unperturbed, atomically flat areas and areas with an atomic-size defect cluster have been investigated. We show that the anomalous behavior of the local tunneling conductivity, which indicates a pronounced enhancement of the semiconductor band gap for the flat areas, is consistent with band bending induced by charges localized at the apex of the tip. Atomic-size defect clusters contain additional charges which modify the band bending; this explains the different behavior of the tunneling conductivity near the defect cluster. The experimentally observed oscillations of the tunneling conductivity near the band gap edges can be directly related to resonant tunneling through quantized surface states which appear because of the band bending. © *1998 American Institute of Physics.* [S0021-3640(98)00802-0]

PACS numbers: 71.20.Nr, 71.24.+q, 61.16.Ch

Measurements with scanning tunneling microscopy (STM) and scanning tunneling spectroscopy (STS) are usually interpreted in terms of standard models for electron tunneling.[1] However, the STM and STS data often also reveal anomalous features which cannot be explained in this way. These STM and STS anomalies tend to become more pronounced as the temperature is lowered.

With decrease of the tunneling contact area and the temperature, nonequilibrium processes start to play a significant role. If the relaxation rate for the electron states is not infinitely large, a nonequilibrium occupation of the electron states will appear in the presence of a nonzero tunneling voltage, even for macroscopic systems.[2–4] For nanometer-scale tunneling contacts the nonequilibrium effects can drastically change the expected tunneling conductivity, especially in the presence of localized states. In this article we show that the importance of the nonequilibrium effects is supported by our STS measurements on InAs semiconductor surfaces.

The relaxation rate for the electron states in the vicinity of the tunneling contact

usually incorporates all possible sources of inelastic scattering. At low temperatures, where inelastic scattering processes are strongly suppressed, the finite relaxation rate will induce a nonequilibrium steady-state distribution in energy space in the contact area. The important decrease of the relaxation rate at liquid helium temperatures implies that the standard expression for the tunneling current has to be modified to

$$I \propto 2\pi e \int d\varepsilon \frac{(n_t^0(\varepsilon) - n_s^0(\varepsilon)) V^2 \rho_t \rho_s \Gamma_t \Gamma_s}{V^2 \rho_t \Gamma_t + V^2 \rho_s \Gamma_s + \Gamma_t \Gamma_s}, \tag{1}$$

where Γ_t and Γ_s are the relaxation rates in the tip and in the sample, respectively, and ρ_t and ρ_s are the corresponding tip and sample density of states, V is the tunneling transfer matrix element, and $n_t^0(\varepsilon)$ and $n_s^0(\varepsilon)$ are the equilibrium electron distribution functions for the tip and sample, respectively.

When the relaxation rate for the nonequilibrium electrons in the tip or in the sample is smaller than the tunneling rate, the current will be determined by the electron relaxation processes, which become very slow at liquid helium temperatures.[5] This explains why typical values for the tunneling current in low-temperature STM measurements tend to be much smaller than at room temperature for the same tip-to-sample separation. For our measurements on InAs surfaces, which are discussed in more detail below, typical tunneling currents at room temperature are in the 1 nA range. At liquid helium temperatures a typical tunneling current is 10 pA.

Experimental and theoretical studies of electron transport properties in the mesoscopic regime have revealed that low-temperature properties of very small structures depend not only on material properties such as the atomic composition, the lattice structure, or the electron effective mass. Transport and tunneling phenomena in mesoscopic systems also depend strongly on the relevant dimensionality and on the specific geometry and configuration of the system. As was indicated in Ref. 6, transport phenomena which are related to electron tunneling in nanostructures have to be treated in terms of a total transmission probability. The latter probability has to include the tunneling process as well as all other relaxation and scattering processes in the system, in a manner similar to the Landauer approach for quantum transport phenomena.[7]

In order to test the relevance of the aforementioned nonstandard tunneling phenomena, we have performed detailed STM and STS measurements at liquid helium temperature on InAs(110) surface for atomically flat areas as well as for areas which contain an atomic-size defect cluster. The experimental results allow us to demonstrate the influence of charging effects which occur in the vicinity of the STM tunnel junction. The STM and STS data have been obtained with a home-built low-temperature microscope with an *in situ* cleavage mechanism to obtain a clean sample surface at low temperature.[8] The samples are *n*-type InAs semiconductor single crystals which have been heavily doped with Sn ($n \simeq 5 \times 10^{17}$ cm^{-3}) and are cleaved along the (110) plane after cooling down to liquid helium temperature.

It is well known that the STM topography of III–V compound semiconductor surfaces, including the InAs(110) surface, depends on the polarity of the applied bias voltage.[9] For negative sample voltages the STM image is determined by the As sublattice, while for positive sample voltages the In sublattice becomes visible. Relaxation of the surface atomic structure results in a tilt of the As atoms in the vertical direction,

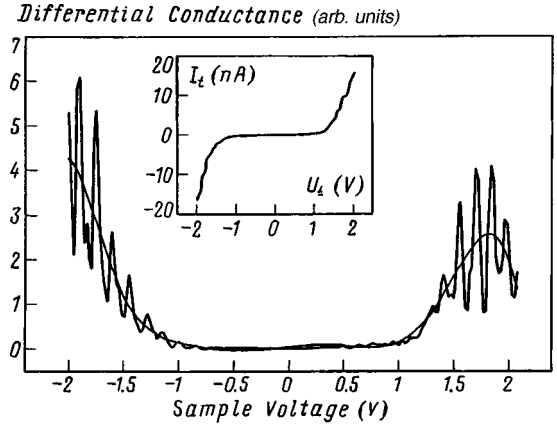

FIG. 1. Differential conductance curve measured above the atomically flat area of the InAs(110) surface; the current–voltage characteristic is shown in the inset.

which can be described in terms of a buckling angle and gives rise to a lateral shift between the In and As sublattices.[10] Our most striking observations for the STS measurements can be summarized as follows. (i) A strongly enhanced semiconductor band gap (about 1.8 eV) is observed for atomically flat surface regions. (ii) The Fermi level is shifted from the conduction band edge inside the band gap despite the high doping level. (iii) In the vicinity of an atomic-size defect cluster the band gap is considerably reduced, but the Fermi level remains pinned in the band gap. A strongly enhanced band gap has been reported previously for low temperature STS measurements on the InAs(110) surface.[11] Even at room temperature the value and position of the measured band gap can be different from the values for the bulk material.[12]

In order to explain the unusual behavior of the tunneling conductivity, we rely on the above theoretical approach and take into account the finite relaxation rate for the electrons in the tunneling contact, which induces a nonequilibrium steady-state electron distribution in the presence of an applied bias voltage. We will argue that for nanometer-scale junctions the nonequilibrium electron distribution results in charging effects and can drastically change the experimentally observed $I(V)$ characteristics.

The inset in Fig. 1 shows a typical $I(V)$ characteristic for an atomically flat area on the InAs(110) surface. The presence of a wide band gap $\simeq 1.8$ eV can be clearly observed. This measured band gap value strongly differs from the bulk value which is 0.43 eV at 4.2 K. Moreover, the Fermi level E_f is located inside the measured band gap. Another result shown in Fig. 1 is the presence of oscillations of the tunneling conductivity which appear on the $dI/dV(V)$ curves near the band gap edges. The period of the oscillations is 0.14 eV for both polarities of the applied voltage.

Our topographic STM images also reveal atomic-size defect clusters (see Fig. 2). The lateral size of the cluster image is about 1 nm and its height is less than 0.5 nm. The measured band gap value above the cluster is 0.4 eV, which is significantly smaller than the observed band gap value for the flat surface regions. We note that the gap structure in $dI/dV(V)$ is much less pronounced when compared to the atomically flat areas (see Fig. 1). On the other hand, E_f is still located inside the band gap and again we observe

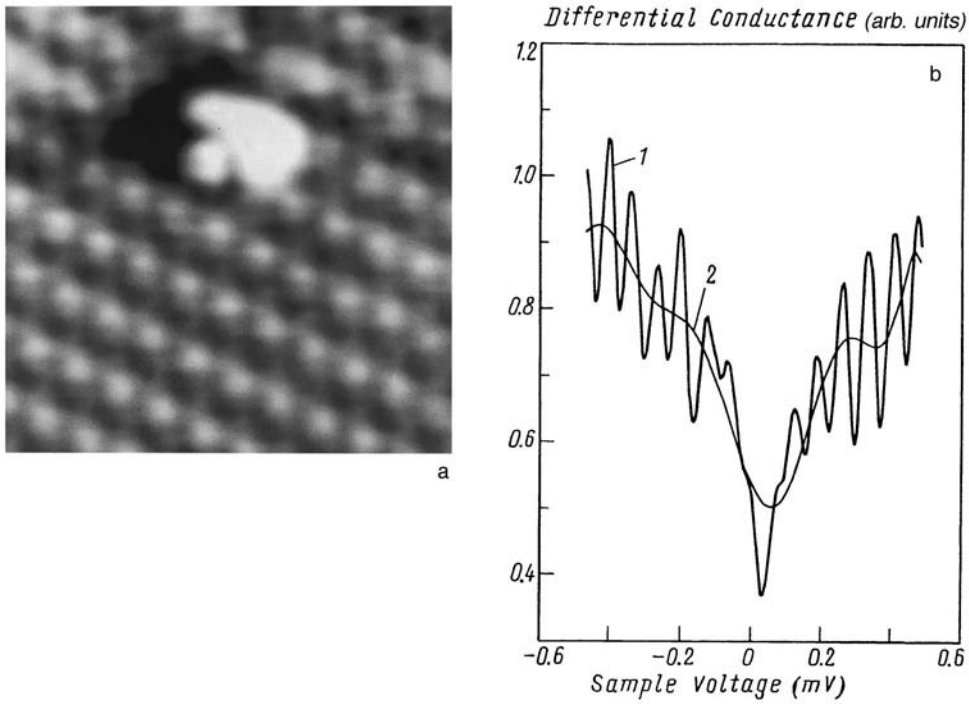

FIG. 2. a: STM image of an atomic-scale defect cluster appearing on the InAs surface. The scanned area measures 44×44 Å. The tunnel current is fixed at 20 pA, while the sample is biased at −500 mV. b: Differential conductance curve measured above the defect cluster shown in Fig. 2a. The curve marked *1* corresponds to the measured curve, while the curve marked *2* corresponds to a fitted 9th order polynomial. The latter curve permits an estimate of the width of the semiconductor band gap.

oscillations of the tunneling conductivity near the gap edges. The period of these oscillations is about 0.09 eV, which is different from the oscillations observed on the flat surface. As is discussed in more detail below, we can in both cases link the anomalous behavior of the tunneling conductivity directly to the band bending induced by charging effects.

In Fig. 3a we illustrate how the anomalously large experimental value for the InAs band gap can be explained for the atomically flat surface regions. We assume that the band bending is induced by charges which are occupying localized states present at the apex of the tip. This charge depends on the polarity and the magnitude of the applied voltage. That is why the commonly used model for tip-induced band bending[13] has to be modified.

The external charge appearing on the apex of the tip is proportional to the difference between the nonequilibrium distribution function for the electrons in the presence of the tunneling current and the distribution function at zero applied voltage. So, the sign of the extra charge at the apex of the tip follows the changes in polarity and value of applied voltage bias. As discussed above, the nonequilibrium distribution appears because of the finite relaxation rate for the electrons, implying that the electron distribution function in the tunneling contact area can be different from the Fermi–Dirac distribution which is

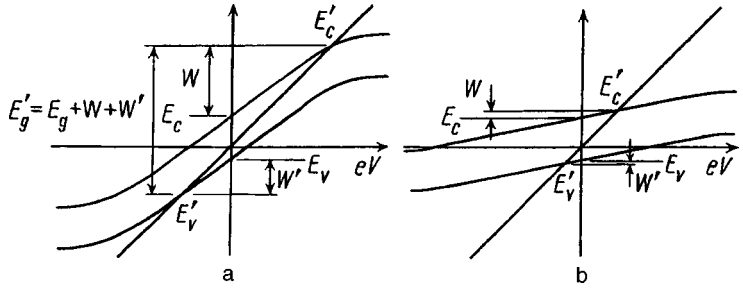

FIG. 3. Schematic view of the tip-induced band bending which occurs near the InAs surface. E_c and E_v correspond to the conduction and valence bands edges at zero bias respectively, while E_c' and E_v' correspond to the measured position of the band edges. The straight line in the first quadrant corresponds to the position of the Fermi level of the tip relative to the Fermi level of the sample. W and W' give the band bending. a: Above an atomically flat InAs surface; b: above an atomic-size cluster.

present in the macroscopic leads. Therefore, in the presence of a tunneling current, the nonequilibrium electron distribution in the contact area results in a negative charge at the apex of the tip for positive sample voltages and in a positive charge for negative sample voltages. The shape of band edges follows the dependence of additional charge versus applied bias voltage as obtained in Ref. 14. Each localized state at the apex of the tip cannot have a charge of more than $2e$, and so saturation of the charge as a function of the applied bias voltage occurs.

The oscillations of the tunneling conductivity cannot be explained in terms of Coulomb blockade or Coulomb staircase effects. If the oscillations are caused by Coulomb charging effects, the period of these oscillations can be directly related to the size a of the relevant particle which is being charged. The oscillation period should be $\Delta E \sim e^2/a$. For $a \sim 0.5$–1.0 nm we obtain $\Delta E \sim 1$–0.5 eV, which does not agree with the experimental data.

The observed oscillations likewise cannot be explained by interference effects. For a tip-to-sample separation $b \sim 0.5$ nm the period of the electron density oscillations ΔE should be much larger than the period obtained from our experimental results. Indeed, $\Delta E \sim (2\pi\hbar p_F)/(mb)$, where m is the effective mass of the electron and $p_F \sim \hbar/a_0$ is the electron wave vector at the Fermi level (a_0 is the InAs lattice constant). This implies that $\Delta E \sim \hbar^2/mba_0 \sim 1$ eV, in clear disagreement with our experimental result $\Delta E \sim 0.09$ eV. A period of 0.09 eV would imply an unrealistic tip-to-sample separation of more than 5 nm.

The band bending values W and W' indicated in Fig. 3 can be estimated to be $W \sim W' \sim e^2/b$, where, as before, b is the tip-to-sample separation. For $b \sim 0.5$–1.0 nm we obtain $W \sim W' \sim 1$–0.5 eV. These values are in line with the usual band bending values.[14] The essential feature of our model (see Fig. 3) is the fact that on account of the voltage dependence of the localized charge, the exact values of W and W' now depend on the polarity as well as on the amplitude of the applied voltage in our case. The experimentally observed band gap will be $E_g \sim W + W' + E_{gb}$, where $E_{gb} \sim 0.43$ eV is the bulk value of the InAs band gap. Therefore, the band bending induced by a charged tip apex can account for the experimentally observed increase of the band gap.

In the presence of an atomic defect cluster, the band bending will be modified in the vicinity of the defect cluster.[15] This occurs because of the additional charging of the localized states which are associated with the defect cluster. For negative sample voltages a negative charge is expected to appear near the cluster, while for positive sample voltages the defect cluster should acquire a positive charge. As in to the case of a charged tip apex, the external charge appearing on the cluster is proportional to the difference between the electron distribution function in the presence of a tunneling current and the distribution function at zero applied voltage. The charge on the defect cluster causes a band bending which partially compensates the band bending induced by the charged tip. The gap edges position now weakly depends on applied bias voltage as is shown in Fig. 3b. Consequently, the experimentally observed band gap above the defect cluster will be reduced in comparison to the band gap which is observed for atomically flat surface areas. However, E_f remains in the band gap. This can be explained by the fact that the band bending is asymmetric: the bending of the conduction band differs from the valence band bending. The decrease of the band gap above a defect cluster has also been observed[16] for a GaAs(110) surface, but the effect was less pronounced than for our InAs(110) surfaces.

In general additional localized states can appear in the band gap due to tip–sample interaction.[17] In this case the common model for tunneling processes has to be modified considerably,[18] and the tunneling current can be completely dominated by these additional localized states for bias voltages less than the band gap value, provided the finite relaxation time for the nonequilibrium electrons is taken into account.

In conclusion, our scanning tunneling spectroscopy measurements of the InAs(110) surface confirm the importance of the band bending induced by localized charges present near the tunneling contact. We have presented theoretical arguments which rely on the finite relaxation time for the electrons at the apex of the tip and on the on the surface of the sample and also on the presence of a nonequilibrium electron distribution at finite voltages. This model consistently explains our main experimental observations, including an enhancement of the semiconductor band gap and the appearance of oscillations in the tunneling conductivity near the band gap edges. The presence of an additional charge on atomic-size defect clusters accounts for the different behavior when compared to atomically flat areas.

The work in Moscow has been supported by the Russian Ministry of Research (Surface Atomic Structures, Grant 95-1.22; Nanostructures, Grant 1-032) and the Russian Fund for Fundamental Research (RFBR, Grants 96-0219640a and 96-15-96420). The work at the KULeuven has been supported by the Fund for Scientific Research–Flanders (FWO) as well as by the Flemish Concerted Action (GOA) and the Belgian Inter-University Attraction Poles (IUAP) research programs. The collaboration between Moscow and Leuven has been funded by the European Commission (INTAS Project 94-3562).

[1] J. Tersoff and D. R. Haman, Phys. Rev. B **31**, 805 (1985).
[2] E. L. Wolf, *Principles of Electron Tunneling Spectroscopy*, Oxford Universitary Press, 1985.
[3] P. N. Trofimenkoff, H. J. Kreuzer, W. J. Wattamaniuk, and J. G. Adler, Phys. Rev. Lett. **29**, 597 (1972).
[4] J. G. Aoller, H. J. Kreuzer, and J. Straus, Phys. Rev. **13**, 2812 (1975).
[5] *Quantum Transport in Nano-Structured Semiconductors*, Kluwer Academic Publishers, 1996.
[6] O. Agam, N. S. Wingreen, and B. L. Altshuler, Phys. Rev. Lett. **78**, 1956 (1997).

[7] R. Landauer, in *Localization, Interaction, and Transport Phenomena*, Proceedings of the International Conference, August 23–28, 1984, Braunschweig, FDR, edited by B. Kramer, G. Bergmann, and Y. Bruynseraede, Springer Series in Solid-State Science, Vol. 61, Springer Verlag, Berlin (1985), p. 38.
[8] S. I. Oreshkin, V. I. Panov, S. V. Savinov *et al.*, Prib. Tekh. Eksp. , No. 4, 145 (1997).
[9] R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987).
[10] N. S. Maslova, S. I. Oreshkin, V. I. Panov *et al.*, ''Charge effects on InAs(110) surfaces by low temperature scanning tunneling spectroscopy,'' *STM-97*, July 20–25, Hamburg, 1997.
[11] J. Wildöer, *Low Temperature Scanning Tunneling Microscopy on Mesoscopic Systems*, Ph. D. thesis, Delft University of Technology, 1996.
[12] R. M. Feenstra and J. A. Stroscio, J. Vac. Sci. Technol. B **5**, 923 (1987).
[13] H. W. M. Salemink, M. B. Johnson, and O. Albrektsen, J. Vac. Sci. Technol. B **12**, 362 (1994); H. W. M. Salemink and O. Albrektsen, Phys. Rev. B **47**, 16044 (1993).
[14] P. I. Arseev, N. S. Maslova, V. I. Panov, and S. V. Savinov, ''Role of charge effects in tunneling investigations of semiconductors,'' *Semiconductors-97*, December 1–5, Moscow, 1997.
[15] J. F. Zheng, X. Liu, N. Newman, E. R. Weber *et al.*, Phys. Rev. Lett. **72**, 1490 (1994).
[16] R. S. Goldman, R. M. Feenstra, B. G. Briner *et al.*, Appl. Phys. Lett. **69**, 3698 (1996).
[17] N. S. Maslova, Yu. N. Moiseev, V. I. Panov *et al.*, Zh. Eksp. Teor. Fiz. **102**, 925 (1992) [Sov. Phys. JETP **75**, 505 (1992)].
[18] P. I. Arseyev and N. S. Maslova, Zh. Éksp. Teor. Fiz. **102**, 1056 (1992) [Sov. Phys. JETP **75**, 575 (1992)].

Published in English in the original Russian journal. Edited by Steve Torstveit.

Oscillations of the magnetoresistance in an inclined magnetic field for MIS structures on (100) silicon with a high electron density

S. G. Semenchinskiĭ[a)]
All-Russia Scientific-Research Institute of the Metrological Service, 119361 Moscow, Russia

L. Smrčka and P. Vašek
Institute of Physics, Academy of Sciences of the Czech Republic, 162 00 Prague 6, Czech Republic

L. Jansen
Grenoble High Magnetic Field Laboratory, BP 166, 38042 Grenoble Cedex 09, France

(Submitted 8 December 1997; resubmitted 19 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 136–140 (25 January 1998)

At electron densities $N_S > 6\times10^{12}$ cm^{-2} a second series of oscillations, which are tentatively attributed to population of the second energy subband, is observed in addition to the main series of Shubnikov–de Haas oscillations. A change in phase of the oscillations of the second series is observed at some angle of inclination α_e of the field. The measured value of α_e is used to calculate the ratio of the cyclotron mass to the effective g factor. The maximum possible cyclotron mass is also determined as $m_H < 0.32m_e$. On this basis it is concluded that the second series of oscillations is due to electrons which have an in-plane effective mass $m^* \approx 0.2m_e$ and which belong to the same valleys of the Fermi surface as in the case of the main oscillations. © *1998 American Institute of Physics.* [S0021-3640(98)00902-5]

PACS numbers: 73.40.Qv, 73.50.Jt

Two-dimensional electron transport on the (100) surface of silicon has been intensively investigated for many years. Nonetheless, its physical picture is still not sufficiently well understood. Specifically, the case of high electron densities, where the theory predicts that more than one energy subband is populated, has not been adequately studied.

The electronic Fermi surface in the bulk of silicon consists of six equal ellipsoids with major axes lying in pairs on the crystallographic axes (100), (010), and (001), i.e., the electrons are evenly distributed between six equivalent valleys. The valleys in the inversion layer at the surface are, generally speaking, no longer equivalent because the effective electron masses m_z in the direction of the normal z to the surface are different. For the (100) surface, electrons of the two valleys with the maximum mass $m_z = 0.91m_e$ and in-plane effective mass $m^* \sim 0.2m_e$ have the lowest potential energy. The

energy E_0 of the ground state of the electrons of these valleys is the energy of the bottom of the first two-dimensional subband. The electronic states in this subband are fourfold degenerate — twofold degenerate with respect to spin and twofold degenerate with respect to the valleys.

The bottom of the next (in energy) subbands should be the energy E_0 of the ground state for four ellipsoids with $m_z = 0.19m_e$ and the energy E_1 of the next level above the ground state for the two valleys with $m_z = 0.91m_e$. The theory predicts close values for these energies. Therefore it is very important to obtain experimental data on the nature of the second subband. However, only several experimental attempts to do so have been published thus far.

In Ref. 1 it was observed that the linear curve of the number of occupied Landau levels of the first subband versus the gate voltage V_g has a kink at some value $V_g = V_B$. The authors explained the kink by the change induced in the electronic density of states in the layer by the onset of population of the second subband. The value of V_B was varied by applying to the sample a uniaxial compression along the (110) axis. On this basis it was concluded that it is the subband $0'$ that starts to be populated when $V_g > V_B$. In Ref. 2 the threshold for population of the second subband was investigated as a function of the density of acceptor impurities in silicon. As a result, a conclusion in favor of subband 1 was drawn. In Ref. 3 the electron mobility in the second subband and its temperature dependence were analyzed and the second subband was identified as $0'$. The values of the mobility were determined indirectly from measurements of the Hall and longitudinal resistances. The validity of such a determination was questioned in Ref. 4.

In Ref. 5 we reported what we thought was the first observation of the series of Shubnikov–de Hass (SdH) oscillations associated with the electrons in the second subband (this series is referred to below as series II). After our work was published, we learned that similar oscillations had been observed earlier in Ref. 6. On the basis of their measurements of the periods of the oscillations in both series and the total electron density N_S in the layer (calculated from the electrical capacitance of the sample and V_g), the authors of that paper[6] determined the degree of degeneracy in the second subband to be 8,[b)] which pointed to the subband $0'$. In Ref. 5 it was observed that the amplitudes of the series-II oscillations are much more strongly temperature-dependent than the amplitudes of the series-I oscillations. For this reason, the series-II oscillations were also attributed to the subband $0'$, where the cyclotron mass is larger. Now, however, an alternative possible reason for the intensification of the temperature dependence of the amplitudes is known. It could be caused by an increase in the fraction of the current transported by the electrons of the second subband as the temperature decreases,[4] so that the conclusion that the cyclotron masses are different for the series-I and series-II oscillations may be incorrect.

The electron density N_{II} in the second subband, determined in Ref. 5 from the period of the oscillations with the use of the proposed degeneracy factor of 8 was found to be much higher than the difference of N_S and the electron density in the first subband N_{I}, determined from the period of the oscillations in the main series (series I). Approximate equality between N_{II} and $N_S - N_{\mathrm{I}}$ can be obtained only by assuming that the series-II degeneracy factor equals 1, i.e., there is no spin and valley degeneracy. This makes it

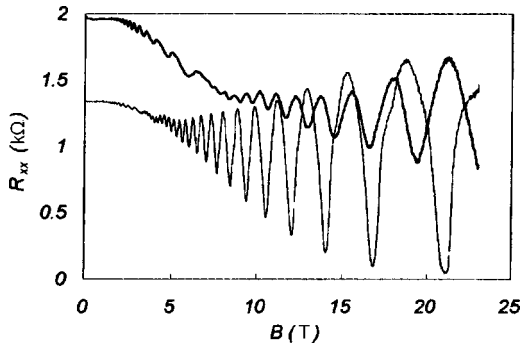

FIG. 1. Longitudinal resistance of a channel versus the magnetic induction for $V_g = 100$ V (top curve) and $V_g = 70$ V (bottom curve).

fundamentally important to obtain direct experimental information about the value of the degeneracy factor.

The objective of the experiment described in the present letter was to make a direct determination of the presence of spin splitting in the spectrum of the electrons responsible for the series-II oscillations and to determine the cyclotron mass of these electrons. We employed the fact that the ratio between the cyclotron and spin splittings of the energy levels of the two-dimensional electron layer depends on the angle of inclination α of the field relative to the direction perpendicular to the plane of the layer, since the spin splitting is determined by the total magnetic field H and the cyclotron splitting is determined only by its normal component H_z. This makes it possible to change the ratio between the values of these splittings in the course of the experiment, making the first harmonic of the SdH oscillations vanish by inclining the field. This should happen at an angle of inclination α_e for which the spin splitting equals half the cyclotron splitting. The phases of the oscillations for $\alpha < \alpha_e$ and $\alpha > \alpha_e$ will differ by 180°. This method was used in Ref. 7 to determine the effective g factor of the electrons in the ground subband of the two-dimensional electron layer on the (100) Si surface.

The experimental sample was a silicon MIS structure fabricated by the standard technology on the surface of silicon with acceptor density 10^{15} cm^{-3}. The sample possessed a rectangular Hall geometry with a 2.5×0.25 mm channel. The subgate oxide layer was 200 nm thick; the ratio $N_S/V_g = 1.15 \times 10^{11}$ cm^{-2}V^{-1}. The peak mobility at 0.4 K was equal to 26000 cm^{2}V^{-1}s^{-1} at $V_g = 3.5$ V. The sample was taken from a new substrate (wafer) which had not provided any of the samples investigated in Refs. 4 and 5. Preliminary measurements showed that the effects described in Refs. 5 and 4 are completely reproduced in this sample also.

During the measurements the sample was placed in liquid ^{3}He at temperature $T = 0.4$ K. A magnetic induction of up to $B = 23$ T was produced with a Bitter magnet. Such a strong maximum magnetic field made it possible to perform measurements over a wide range of angles of inclination while maintaining a maximum value $B_z = 5$ T sufficient for observing series-II oscillations at angles up to $\alpha > 70°$.

Examples of the original experimental traces of the longitudinal resistance of the sample are presented in Fig. 1. The SdH oscillations of both series are clearly visible in

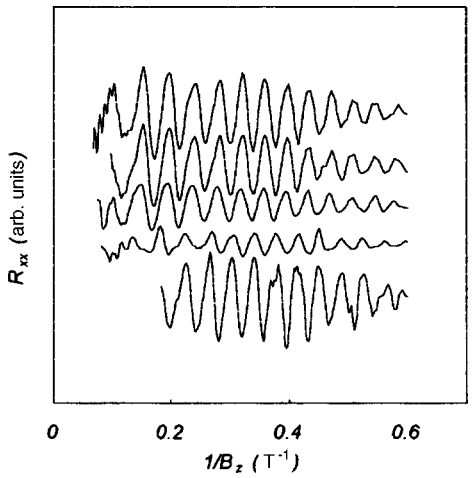

FIG. 2. Series-II oscillations in an inclined magnetic field. The inclination angles α of the field with respect to the z axis for the different curves (top to bottom) are: 0°, 32°, 47.5°, 60.7°, and 69.5°.

them. The series-II oscillations start in weaker magnetic fields and have a longer period than the series-I oscillations.

The effect of the inclination of a magnetic field on the series-II oscillations is illustrated in Fig. 2. For greater clarity, the experimental data were prefiltered with a high-frequency digital filter in order to eliminate the monotonic component of the dependence $R_{xx}(B)$. It is evident from the figure that the positions of the oscillations remain unchanged relative to B_z, while their phase changes sign after some angle α_e in the range $47.5° < \alpha_e < 60.7°$. Interpolation of the α dependence of the amplitude of the first harmonic of the oscillations gives $\alpha_e = 58° \pm 2°$ at $V_g = 100$ V and $\alpha_e = 52° \pm 4°$ at $V_g = 70$ V. A similar change in phase of the oscillations is also observed for series I. In this case $\alpha_e = 54° \pm 2°$ in the entire range 70 V $< V_g <$ 100 V (this is somewhat lower than the value obtained in Ref. 7).

The observed change in phase is in complete agreement with that expected for a two-dimensional electronic system with two different spin states. On the basis of the fact that the moment of the change in phase of the first harmonic of the oscillations corresponds to $\Delta_H - \Delta_s = \Delta_s$ (here Δ_H and Δ_s are the cyclotron and spin energy splittings, respectively), it is possible to determine the ratio between the cyclotron mass m_H and the effective g factor g^*. Rewriting this equation in the form $(\hbar e/m_H c)H \cos \alpha_e = 2g^* \mu_B H$, where $\hbar$ is Planck's constant, $\mu_B = e\hbar/2m_e c$ is the Bohr magneton, c is the speed of light, and m_e is the mass of a free electron, we obtain $g^* = (m_e/m_H)\cos \alpha_e$. Substituting the value of α_e, we obtain for series-II oscillations $g^* = (m_e/m_H)(0.54 \pm 0.02)$ for $V_g = 100$ V and $g^* = (m_e/m_H)(0.60 \pm 0.04)$ for $V_g = 70$ V. For series-I oscillations $g^* = (m_e/m_H)(0.59 \pm 0.02)$ in the range of V_g from 70 V to 100 V.

The ratios obtained between g^* and m_H make it possible to estimate the upper limit of the possible values of m_H for series-II oscillations. It is known that the g factor of bulk conduction electrons in silicon equals 2. The electron–electron interaction in the two-dimensional electron layer on the surface can increase the effective g factor,[8] but there

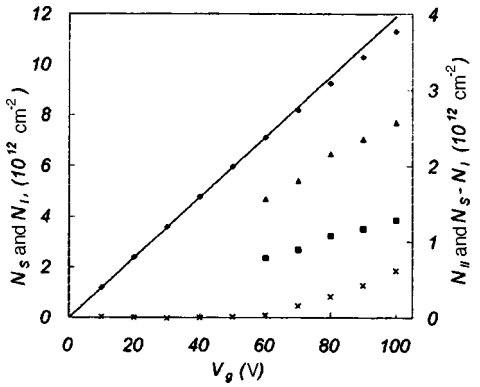

FIG. 3. Electron density determined from the periods of the SdH oscillations for different series of oscillations: $\diamond$ — N_{I}, $\square$ — N_{II} for degeneracy factor 2, Δ — N_{I} for degeneracy factor 4, $\times$ — $N_S - N_{\mathrm{I}}$. The straight line is the total electron density N_S.

are no grounds for assuming that it decreases. Taking for estimates $g^* > 2$, we obtain $m_H/m_e < 0.28$ for $V_g = 100$ V and $m_H/m_e < 0.32$ for $V_g = 70$ V. Since this is much less than the expected value $m_H/m_e = 0.46$ for the subband $0'$, the oscillations must be attributed to electrons belonging to the same valleys as the series-I oscillations and having an in-plane effective mass $m^* \approx 0.2m_e$.

It follows on this basis that at least a twofold spin degeneracy is present in the series-II oscillations. One can expect also a twofold valley degeneracy for these oscillations, as for the series-I oscillations. Figure 3 shows the corresponding electron densities calculated from the positions of the series-I and series-II SdH oscillations as well as the total in-layer electron density as a function of V_g. One can see that the difference $N_S - N_{\mathrm{I}}$ is appreciably less than N_{II} even when the possible valley degeneracy of the series-II oscillations is neglected.

At present we cannot explain this discrepancy. It is possible that for the actually achievable density N_S we are dealing with population of the ''tails'' of the second subband (it was predicted in Ref. 9 that these tails are long), and the discrepancy would vanish if higher densities could be achieved. Possible reasons could be electron–electron interaction and intervalley scattering of the electrons. According to Ref. 10, such an interaction for the (110) orientation can lead to the appearance of a domain structure with population of only one valley in each domain and, in consequence, to a decrease of the observed degeneracy factor. In principle, similar effects can also be conjectured to occur in other cases where several subbands are filled, including also for the (100) orientation, but at present we do not know of any calculations for this case.

The measurements were performed at the High Magnetic Field Laboratory in Grenoble, France. This work was supported by Grant 96-02-16838 from the Russian Fund for Fundamental Research and Grant 202/96/0036 from the Grant Agency of the Czech Republic.

[a)]e-mail: semench@ssemen.msk.su
[b)]Apparently, an error was made in these calculations. A recalculation using the experimental data presented in that same paper[6] gives a much lower value for the degeneracy factor.

[1]D. C. Tsui and G. Kaminsky, Phys. Rev. Lett. **35**, 1468 (1975).
[2]W. E. Howard and F. F. Fang, Phys. Rev. B **13**, 2519 (1976).
[3]G. H. Kruithof and T. M. Klapwijk, Phys. Rev. B **42**, 11412 (1990).
[4]S. G. Semenchinsky, L. Smrcka, J. Stehno, and V. Borzenets, Phys. Lett. A **217**, 335 (1996).
[5]S. G. Semenchinsky, L. Smrcka, and J. Stehno, Phys. Lett. A **209**, 218 (1995).
[6]J. Herfort, J. Kerner, and W. Braune, Semicond. Sci. Technol. **7**, 1260 (1992).
[7]F. F. Fang and P. J. Stiles, Phys. Rev. **174**, 823 (1968).
[8]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[9]S. Mori and T. Ando, Phys. Rev. B **19**, 6433 (1980).
[10]M. J. Kelly, L. M. Falicov, Phys. Rev. B **15**, 1974 (1977).

Translated by M. E. Alferieff

A simple approach to the evaluation of the Hall conductivity in impure metals within the Green function formalism

V. M. Edelstein[a)]

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 22 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 141–145 (25 January 1998)

A simple method of dealing with the Hall effect in metals with short-ranged impurities in a weak magnetic field is proposed. The method is based on a Schwinger representation for the electron Green function in the magnetic field. The efficacy of the method is demonstrated on a calculation of the antisymmetric components of the conductivity tensor at finite wave vector. © *1998 American Institute of Physics.* [S0021-3640(98)01002-0]

PACS numbers: 72.10.Bg, 72.15.Gd

Despite the discovery of the quantum Hall effect, the theoretical analysis of the Hall effect in metals in *weak* magnetic fields has been attracting a lot of attention[1–6] because of its practical as well as scientific interest. The quantum-mechanical treatments of the topic published to date have been based on a method proposed in Ref. 2. In this method, an electron system subjected to a uniform external magnetic field is considered to be a limiting case of the system placed in a fictitious nonuniform magnetic field with the vector potential $\mathbf{A}(\mathbf{r})=\mathbf{A}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}}$. One finally lets the wave vector $\mathbf{q}\to 0$ to recover the case of uniform magnetic field. The nonuniform field results in inhomogeneity of the system which induces the carrier diffusion. In the Feynman-diagram language this implies the appearance of diffusion poles. The necessity of eliminating the poles to obtain a divergent-free expression for the Hall conductivity $\sigma_{ij}^{(H)}$ makes the method somewhat cumbersome (in spite of some improvements published later[5]). Furthermore, the method has only been formulated for evaluation of the Hall conductivity at zero wave vector.

The purpose of the present paper is to propose another quantum approach which is convenient in the particular case of p-independent impurity scattering. The distinctive feature of our approach is that the external magnetic field is considered to be uniform from the very beginning. Therefore, no carrier diffusion takes place and we only have the Feynman diagrams that do not contain the diffusion poles. In addition, the method can be extended to finite wave vectors without any difficulties, yielding an expression for $\sigma_{ij}^{(H)}\times(\omega,\mathbf{q})$. (Here we consider only macroscopic systems.)

Let $\mathbf{a}(\mathbf{x},t)$ and $\mathbf{A}(\mathbf{x})$ be the vector potentials of the driving electric field $\mathbf{E}=-\partial\mathbf{a}/c\partial t$ and external static magnetic field $\mathbf{H}=\nabla\times\mathbf{A}$, respectively. The linear response to $\mathbf{a}$ is known to be determined by the current–current correlation function

$$Q_{ij}(\mathbf{x},\mathbf{y};i\omega_n)=\int_0^{1/T} d\tau\langle\hat{T}\hat{J}_i(\mathbf{x},\tau)\hat{J}_j(\mathbf{x},0)\rangle, \tag{1}$$

where T in the upper limit of the integral is the temperature and $\hat{T}$ in the angle brackets is the time-ordering operator. In what follows we assume that the so-called ''diamagnetic part'' of the response, $(ne^2/m)\,\delta(\mathbf{x}-\mathbf{y})$, is canceled by the appropriate part of Q_{ij} (analytically continued to the real frequency axis) in the usual manner. The electron charge is taken to be $-e$. Then the current operator $\hat{\mathbf{J}}$ is the sum of the kinetic part,

$$\hat{\mathbf{j}}_{\rm kin}=-\frac{ie}{2m}[\psi^+(\mathbf{r})\nabla\psi(\mathbf{r})-\nabla\psi^+(\mathbf{r})\psi(\mathbf{r})],$$

and the diamagnetic part,

$$\hat{\mathbf{J}}_{\rm dia}=-\frac{e^2}{mc}\mathbf{A}(\mathbf{r})\psi^+(\mathbf{r})\psi(\mathbf{r}).$$

For the sake of brevity, we drop the spinor indices on the field operators and Green functions. The field $\mathbf{H}$ is assumed to be small so that $\omega_c\tau<1$ (here $\omega_c=eH/mc$ is the cyclotron frequency and τ is the mean free time) and, hence, one can neglect the Landau quantization and use the representation for the electron Green function originally due to Schwinger[7]

$$G_\epsilon(\mathbf{r},\mathbf{r}';\mathbf{H})=\exp\left(\frac{ie}{c}\int_{\mathbf{r}}^{\mathbf{r}'}d\mathbf{r}''\cdot\mathbf{A}(\mathbf{r}'')\right)G_\epsilon(\mathbf{r}-\mathbf{r}';0), \tag{2}$$

where $G_\epsilon(\mathbf{r}-\mathbf{r}';0)$ is the Green function in zero magnetic field. In this paper the symmetric gauge, $\mathbf{A}=\frac{1}{2}\mathbf{H}\times\mathbf{r}$, is chosen, and we put $\hbar=1$ throughout. It should be noticed that such a representation has been successfully employed in many branches of physics, e.g., in superconductivity[8] and plasma physics,[9] but never (to our knowledge) in the theory of the Hall effect in weak magnetic fields. The current operator can be written as

$$\hat{\mathbf{J}}(\mathbf{x})=\lim_{\mathbf{x}'\to\mathbf{x}}\hat{\mathbf{J}}(\mathbf{x},\mathbf{x}'),\quad \hat{\mathbf{J}}(\mathbf{x},\mathbf{x}')=\psi^+(\mathbf{x})\mathbf{J}(\mathbf{x},\mathbf{x}')\psi(\mathbf{x}'), \tag{3}$$

$$\mathbf{J}(\mathbf{x},\mathbf{x}')=-e\left[\frac{\nabla_x-\nabla_{x'}}{2im}+\frac{e}{2mc}\mathbf{A}(\mathbf{x}+\mathbf{x}')\right]. \tag{4}$$

The thermal average in (1) gives rise to a set of diagrams[10] which have the form of a two-vertex electron loop with various impurity-line insertions and with the current operator (in the Schrödinger representation) at both these vertices. The impurity potential $U(\mathbf{r})$ will be assumed short-ranged so that $\langle U(\mathbf{r})U(\mathbf{r}')\rangle\sim\delta(\mathbf{r}-\mathbf{r}')$. If, as usual, one includes the impurity self-energy in the electron Green function, the remaining impurity lines necessarily connect the upper and lower electron lines forming a given loop and can be considered as impurity vertex corrections. Without an external magnetic field, the evaluation of such diagrams is greatly simplified by making use of the Fourier transform. In the presence of the magnetic field the translational invariance of the Green function is broken on account of the Schwinger factors

$$\Phi(\mathbf{x},\mathbf{y})=\exp\left(\frac{ie}{c}\int_{\mathbf{x}}^{\mathbf{y}}d\mathbf{r}\cdot\mathbf{A}(\mathbf{r})\right)=\exp\left(\frac{ie}{2c}\mathbf{H}\cdot(\mathbf{x}\times\mathbf{y})\right), \tag{5}$$

making the immediate application of the Fourier method impossible. However, the general theorem[11] stating that any electron loop taken as a whole has to be invariant under translations means that the translational invariance must be recovered by explicit evaluation of the loop. Let us show how this happens in the case of ''empty'' loop (without impurity insertions). The analytical expression for the loop contains

$$\lim_{\substack{\mathbf{x}'\to\mathbf{x}\\ \mathbf{y}'\to\mathbf{y}}} J_i(\mathbf{x},\mathbf{x}')J_j(\mathbf{y},\mathbf{y}')\Phi(\mathbf{x},\mathbf{y}')\Phi(\mathbf{y},\mathbf{x}')G(\mathbf{x}-\mathbf{y}')_{i\epsilon+i\omega}G(\mathbf{y}-\mathbf{x}')_{i\epsilon}, \tag{6}$$

which, in view of the relations

$$\frac{\nabla_x}{2i}\Phi(\mathbf{x},\mathbf{y})=\Phi(\mathbf{x},\mathbf{y})\left[\frac{\nabla_x}{2i}-\frac{e}{2c}\mathbf{A}(\mathbf{y})\right],\quad \frac{\nabla_y}{2i}\Phi(\mathbf{x},\mathbf{y})=\Phi(\mathbf{x},\mathbf{y})\left[\frac{\nabla_y}{2i}+\frac{e}{2c}\mathbf{A}(\mathbf{x})\right] \tag{7}$$

can be recast as

$$\lim_{\substack{\mathbf{x}'\to\mathbf{x}\\ \mathbf{y}'\to\mathbf{y}}} \Phi(\mathbf{x},\mathbf{y}')\Phi(\mathbf{y},\mathbf{x}')\left[\frac{\nabla_x-\nabla_{x'}}{2im}+\frac{e}{2mc}\mathbf{A}(\mathbf{x}+\mathbf{x}'-\mathbf{y}-\mathbf{y}')\right]_i$$
$$\times\left[\frac{\nabla_y-\nabla_{y'}}{2im}-\frac{e}{2mc}\mathbf{A}(\mathbf{x}+\mathbf{x}'-\mathbf{y}-\mathbf{y}')\right]_j G(\mathbf{x}-\mathbf{y}')_{i\epsilon+i\omega}G(\mathbf{y}-\mathbf{x}')_{i\epsilon}. \tag{8}$$

In Eq. (8) the derivatives do not act on the phase factors; therefore, one can set $\mathbf{y}'=\mathbf{y}$ and $\mathbf{x}'=\mathbf{x}$ in *the factors*. As a result, the product of the factors reduces to unity. Now one should substitute into (8) the Fourier representation for the free-electron Green functions and perform the coordinate differentiations. After that one can immediately set $\mathbf{y}'=\mathbf{y}$ and $\mathbf{x}'=\mathbf{x}$ in remaining functions. Then the total expression for the loop takes the explicitly translationally invariant form

$$-e^2\sum_{\epsilon,\mathbf{p},\mathbf{q}_1}\left[\frac{\mathbf{p}}{m}+\frac{e}{mc}\mathbf{A}(\mathbf{x}-\mathbf{y})\right]_i e^{i\mathbf{q}_1\cdot(\mathbf{x}-\mathbf{y})}G_{i\epsilon+i\omega}$$
$$\times\left(\mathbf{p}+\frac{\mathbf{q_1}}{2}\right)G_{i\epsilon}\left(\mathbf{p}-\frac{\mathbf{q_1}}{2}\right)\left[\frac{\mathbf{p}}{m}-\frac{e}{mc}\mathbf{A}(\mathbf{x}-\mathbf{y})\right]_j. \tag{9}$$

The part of the response $Q^R_{ij}(\omega,\mathbf{q})=\int d^3(\mathbf{x}-\mathbf{y})e^{-i\mathbf{q}_1\cdot(\mathbf{x}-\mathbf{y})}Q_{ij}(\mathbf{x}-\mathbf{y};\omega+i0)$ linear in $\mathbf{H}$ can be written [with the help of the identity $\int d^3re^{i\mathbf{q}\cdot\mathbf{r}}\mathbf{r}=-i(2\pi)^3\nabla_{\mathbf{q}}\delta(\mathbf{q})$] as

$$\sigma^{(H)}_{ij}(\omega,\mathbf{q})=\frac{e^2}{\pi}\left(\frac{ie}{4mc}\right)\left[(\mathbf{H}\times\nabla^{R-A}_{\mathbf{p}})_i G^R_\omega\left(\mathbf{p}+\frac{\mathbf{q}}{2}\right)G^A_0\left(\mathbf{p}-\frac{\mathbf{q}}{2}\right)\frac{p_j}{m}\right.$$
$$\left.-\frac{p_i}{m}G^R_\omega\left(\mathbf{p}+\frac{\mathbf{q}}{2}\right)G^A_0\left(\mathbf{p}-\frac{\mathbf{q}}{2}\right)(\mathbf{H}\times\nabla^{R-A}_{\mathbf{p}})_j\right], \tag{10}$$

where the superscript $R(A)$ stands for the retarded (advanced) part of the function and the operator $\nabla^{R-A}_{\mathbf{p}}$ in the first term acts only on the pair of the Green functions according to the rule $\nabla^{R-A}_{\mathbf{p}}G^RG^A=(\nabla_{\mathbf{p}}G^R)G^A-G^R(\nabla_{\mathbf{p}}G^A)$, but *not* on the velocity vertex $\mathbf{p}/m$. The operator $\nabla^{R-A}_{\mathbf{p}}$ in the second term acts to the left in the same way. Proceeding in the same manner, one can see that the contribution of all impurity-ladder diagrams has the form depicted in Fig. 1. The evaluation of $\sigma^{(H)}_{ij}$ is now reduced to the level of, say, the Drude

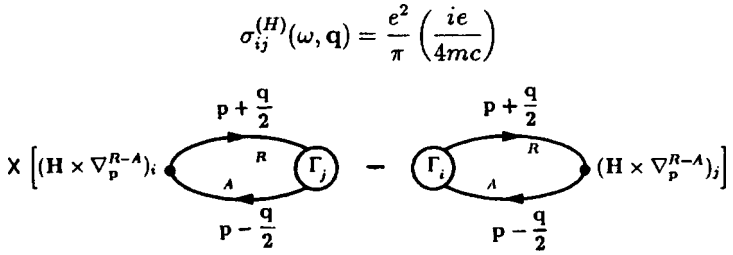

FIG. 1. The Feynman diagrams for the Hall conductivity. The gradient operator $\nabla_{\mathbf{p}}^{R-A}$ in the second term acts to the left on the pair of the Green functions $G^R G^A$. $\mathbf{\Gamma}$ is the usual impurity-renormalized velocity vertex.

conductivity. It should be stressed that the δ-functional form of the impurity-field correlator is essential in deriving this result. Straightforward calculations yield

$$\sigma_{ij}^{(H)}(\omega,\mathbf{q}) = \sigma_3 \frac{3\omega_c\tau}{2(1-i\Omega)^2} e_{ijn}\left\{ h_n\left(\frac{\beta-\gamma}{t^2}\right)\frac{i\Omega}{1-\beta-i\Omega} + \hat{q}_n(\hat{\mathbf{q}}\cdot\mathbf{h})\left[\left(\frac{\gamma-\beta}{t^2}\right)\frac{i\Omega}{1-\beta-i\Omega} - 2\frac{1-\beta}{t^2}\right]\right\}, \tag{11}$$

where $\sigma_3 = n_3 e^2\tau/m$, $n_3 = p_F^3/3\pi^2$, $\Omega = \omega\tau$, $t = qv_F\tau/(1-i\Omega)$, $\mathbf{h} = \mathbf{H}/H$, $\beta = t^{-1}\times\arctan t$, and $\gamma = (1+t^2)^{-1}$. The known symmetric (Drude) components of the conductivity tensor, $\sigma_{ij}^{(D)}$, in the same notation have the form[12]

$$\sigma_{ij}^{(D)}(\omega,\mathbf{q}) = \sigma_l^{(3)}\hat{q}_i\hat{q}_i + \sigma_{tr}^{(3)}(\delta_{ij}-\hat{q}_i\hat{q}_i), \tag{12}$$

$$\sigma_l^{(3)} = \frac{3\sigma_3}{1-i\Omega}\left(\frac{1-\beta}{t^2}\right)\frac{-i\Omega}{1-\beta-i\Omega}, \quad \sigma_{tr}^{(3)} = \frac{3\sigma_3}{2(1-i\Omega)}\left(\beta-\frac{1-\beta}{t^2}\right).$$

Applied to a 2D electron system, our method yields

$$\sigma_{ij}^{(H)}(\omega,\mathbf{q}) = e_{ijn}h_n\sigma_2^H, \quad \sigma_2^H = \sigma_2\frac{\omega_c\tau}{(1-i\Omega)^2}\left(\frac{i\Omega}{1+t^2}\right)\frac{1}{\sqrt{1+t^2}(1-i\Omega)-1}, \tag{13}$$

$$\sigma_{ij}^{(D)}(\omega,\mathbf{q}) = \sigma_l^{(2)}\hat{q}_i\hat{q}_i + \sigma_{tr}^{(2)}(\delta_{ij}-\hat{q}_i\hat{q}_i), \tag{14}$$

$$\sigma_{tr}^{(2)} = \frac{2\sigma_2}{1-i\Omega}\left(\frac{\sqrt{1+t^2}-1}{t^2}\right), \quad \sigma_l^{(2)} = \frac{2\sigma_2}{1-i\Omega}\left(\frac{i\Omega}{t^2}\right)\frac{\sqrt{1+t^2}-1}{1-(1-i\Omega)\sqrt{1+t^2}}$$

($\sigma_2 = n_2 e^2\tau/m$, $n_2 = p_F^2/2\pi$) which agrees with the result recently obtained[6] by means of the 2D classical Boltzmann equation with a modified collision integral.

In conclusion, we have proposed a method for microscopic calculation of the Hall conductivity in weak magnetic fields. In essence, the method is based solely on the Schwinger formula, i.e., on the gauge transformation rules in quantum mechanics. The diagrammatic expression obtained for σ_{ij}^H appears to be topologically similar to that for the Drude conductivity. Our approach complements the previous one[2] in that it gives $\sigma(\omega,\mathbf{q})$ at finite $\mathbf{q}$ but only for the p-independent scattering, while the method of Ref. 2

gives $\sigma(\omega,0)$ but for arbitrary impurity scattering. In addition, it is hoped that the method reported here will provide a better handle in studying the effects of weak localization and (short-ranged) interparticle interaction on the Hall conductivity.

This work was supported, in part, by Grant 96-02-19568 from the Russian Fund for Fundamental Research.

[a)]e-mail: edelsh@issp.ac.ru

---

[1] M. N. Cohen, M. J. Harrison, and W. A. Harrison, Phys. Rev. **117**, 937 (1960).
[2] H. Fukuyama, H. Ebisawa, and Y. Wada, Prog. Theor. Phys. **42**, 494 (1969).
[3] G. J. Morgan and M. A. Howson, J. Phys. C **18**, 4327 (1985).
[4] H. Kohno and K. Yamada, Prog. Theor. Phys. **80**, 623 (1988).
[5] See M. Itoh, Phys. Rev. B **45**, 4241 (1992), and references cited therein.
[6] A. D. Mirlin and P. Wölfle, Phys. Rev. Lett. **78**, 3717 (1997).
[7] J. Schwinger, Phys. Rev. **82**, 664 (1957).
[8] L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **36**, 1918 (1959) [Sov. Phys. JETP **9**, 1364 (1959)]; P. A. Lee and M. G. Payne, Phys. Rev. **5**, 923 (1972).
[9] N. J. Noring, Ann. Phys. (N.Y.) **31**, 1 (1969).
[10] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood Cliffs, N.J., 1963 [Russian original: Fiz. Mat. Giz., Moscow, 1962].
[11] H. Scher and T. Holstein, Phys. Rev. **148**, 598 (1966).
[12] See, e.g., L. R. Ram Mohan, E. Kartheuser, and S. Rodrigues, Phys. Rev. B **20**, 3233 (1979).

Published in English in the original Russian journal. Edited by Steve Torstveit.

Characteristic features of microcontact spectra of contacts between a metal and a quasi-1D conductor with a charge density wave

A. A. Sinchenko
Moscow State Engineering Physics Institute, 115409 Moscow, Russia

Yu. I. Latyshev, S. G. Zybtsev, and I. G. Gorlova[a)]
Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia

(Submitted 23 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 146–151 (25 January 1998)

The current–voltage characteristics (IVCs) of Cu–$K_{0.3}MoO_3$ point contacts are investigated. The character of the nonlinearity and the observed asymmetric features of the IVCs indicate that a substantial shift of the chemical potential occurs near the boundary with the normal metal. Deformation of the charge density wave by an applied electric field leads to strong bending of the energy bands and to the formation of a potential well, whereupon the Fermi level falls within the region of allowed single-electron states. © *1998 American Institute of Physics.* [S0021-3640(98)01102-5]

PACS numbers: 71.45.Lr, 73.40.Ns

In quasi-1D conductors, a charge density wave (CDW) or a spin-density wave (SDW) arises as the temperature decreases below the Peierls transition point T_P. Many properties of materials with a CDW and SDW have been studied in detail; a review can be found in Ref. 1. In the Peierls state single-electron excitations coexist with a deformable mobile CDW. In subthreshold electric fields a CDW cannot move as a whole, and in Peierls semiconductors (PSs) the single-electron excitations determine the conductivity. However, a deformation of the CDW (for example, by an electric field) changes the electron and hole densities, as a result of which local disturbances of the CDW can make a substantial contribution to the conductivity in subthreshold fields. In this sense a PS can be viewed as a conventional semiconductor but with a variable degree of doping that depends on external perturbations. It has been shown theoretically in recent works[2–5] that nonuniform perturbations of the phase, which can be caused by nonuniformity of the sample, by the presence of the contacts, or by variations induced in the phase of the CDW by pinning centers, play an important role in the transport properties of PSs, and screening by single-electron excitations is effective right down to the lowest temperatures. Defects of an n-type crystal (phase or amplitude solitons, dislocations) can produce strong bending of the energy bands, so that the Fermi level can even be brought into the region of allowed single-electron states.

A point contact of a normal metal (M) with a PS is a very convenient object for

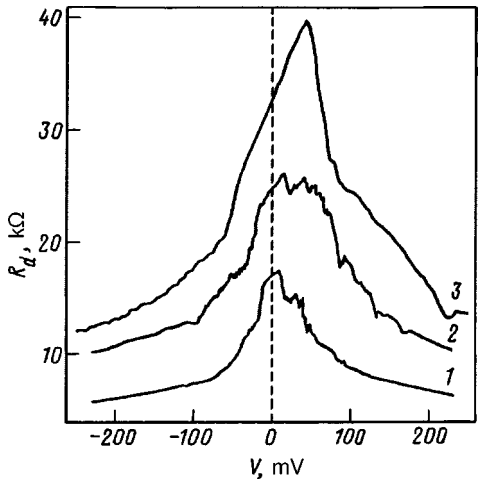

FIG. 1. $R_d(V)$ of three different Cu–$K_{0.3}MoO_3$ contacts at $T=77$ K.

investigating the effect of local perturbations on the properties of materials with a CDW because of the fact that the electric field is localized near the contact in a small region with a characteristic size of the order of the contact diameter d in the case of an isotropic metal,[6] and with $d^*=d\sqrt{n}$, where n is an anisotropy factor, in the case of a PS.

In the present work the characteristics of Cu–$K_{0.3}MoO_3$ point contacts oriented along chains were investigated in the temperature range 77–300 K. The sampes were $K_{0.3}MoO_3$ single crystals with a length of 0.5–1.0 mm and a cross section of $\sim 2-5 \times 10^3$ μm^2. The Peierls transition temperature was equal to $T_P=183$ K. The single crystals were provided by the Center for Low Temperature Research (CRTBT–CNRS, Grenoble). Ten $K_{0.3}MoO_3$ single crystals were investigated. At least 10 point contacts were made for each single crystal, and their characteristics were measured. The experimental apparatus described in detail in Ref. 7 was used for making point contacts directly at low temperature and for measuring the current–voltage characteristics and their first derivatives dV/dI. Thin copper wires (40 and 71 μm in diameter) with electrochemically sharpened tips were used as the normal counterelectrodes. The radius of curvature of a tip did not exceed 1 μm.

Typical values of the resistance of the contacts with the tip of the normal electrode touching the surface of the experimental crystals were equal to 100–200 kΩ at $T=77$ K. These contacts were extremely unstable and were characterized by nonreproducibility. As the clamping force increased, the resistance decreased to 15–80 kΩ and the stability of the contacts improved. In this range of resistances the characteristics of the contacts did not show hysteresis and were completely reversible. Figure 1 shows the curves of the differential resistance $R_d=dV/dI$, measured at $T=77$ K, versus the voltage V for point contacts of this type with three different samples. The point-contact spectra are quite complicated. However, most of the contacts investigated exhibit a number of characteristic features. First, the curves have a pronounced asymmetry and the maximum of $R_d(V)$ is shifted from zero to positive voltages. Second, a section with a linear variation of $R_d(V)$ is present near zero bias. Third, a sharp decrease of the differential resistance,

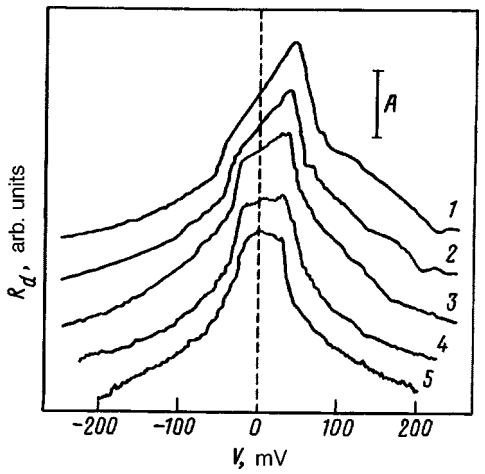

FIG. 2. $R_d(V)$ of contact No. 3 (Fig. 1) with T, $R_d(0)$, and A equal to: 77 K, 32.6 kΩ, and 10 kΩ (*1*); 83.5 K, 19.3 kΩ, and 6 kΩ (*2*); 91 K, 11.4 kΩ, and 3 kΩ (*3*); 96.8 K, 8.3 kΩ, and 2.5 kΩ (*4*); 108.4 K, 3.7 kΩ, and 1 kΩ (*5*).

often looking like a discontinuity (curves *1* and *2* in Fig. 1), accompanied by an increase in noise was observed for most contacts when the voltage reached the value $V=V_0$ corresponding to the maximum of R_d. In several cases the resistance dropped sharply at voltages somewhat above V_0 (curve *3* in Fig. 1). For some contacts it was possible to observe the evolution of the point-contact characteristics as a function of temperature. Figure 2 shows the curves $R_d(V)$ for contact No. 3 (Fig. 1) at different temperatures. As one can see, the slope of the linear section of $R_d(V)$ decreases monotonically with increasing temperature and vanishes completely at temperatures $T>100$ K. In this temperature range the maximum of the differential resistance corresponds to zero bias voltage, and the dependence $R_d(V)$ is approximately symmetric.

For the two contacts whose curves $R_d(V)$ are shown in Figs. 3 and 4 a sharp, deep minimum of width $\sim$15–20 mV was observed in the differential resistance at $T=77$ K with a positive bias voltage ranging from 150 to 200 mV. This feature was reproduced well and did not look like a discontinuity. An attempt to trace the evolution of the minimum while changing the resistance of the contact by continuously increasing the pressure of the needle resulted in rapid broadening and complete vanishing of the feature. For these contacts the behavior of the differential resistance near the maximum of $R_d(V)$ is interesting (see the insets in Figs. 3 and 4). A sharp decrease in the resistance occurs after $R_d(V)$ reaches its maximum value at a voltage above V_0. The linearity in $R_d(V)$ also remained after the maximum value of the differential resistance was reached, and the sign of the slope changed.

To explain the results we shall proceed from the fact that the entire sample is in a pinning state, i.e., coherent sliding of the CDW in the entire volume of the sample is impossible in the experimental voltage range. Indeed, the resistance of a point contact is formed in a small region near the contact and equals $\sim 10^4$–10^5 Ω, which is much higher than the resistance of the thinnest samples in the pinning state ($\sim$1 kΩ). An estimate of the diameters of the contacts using the Sharvin formula[8] gives $d=5$–15 nm, whence we

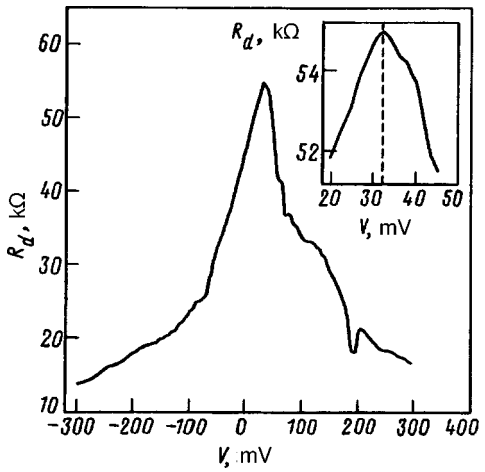

FIG. 3. $R_d(V)$ of contact No. 3 at $T=77$ K.

estimate the region of variation of the electric field as $d^* \sim 50-100$ nm. In the absence of sliding, this field will lead to a deformation of the CDW and to a change in the quasiparticle density and therefore also to a shift of the chemical potential μ. As is well known, the quasiparticle conductivity of a PS satisfies $\sigma \propto |\Delta q|$ (see, for example, Ref. 9), where Δq is the change in the wave number of the CDW corresponding to a shift of μ from the center of the band gap. It is logical to associate the section of linear variation of $R_d(V)$ to elastic deformation of the CDW. Then the maximum of the differential resistance for curve *3* in Fig. 1 and also for the curves in Figs. 3 and 4 corresponds to $\Delta q = 0$. As the temperature increases, the screening of the electric field by the single-electron excitations becomes more effective,[4] which could be the reason why the slope of the linear section decreases with increasing T. The chemical potential on the M–PS

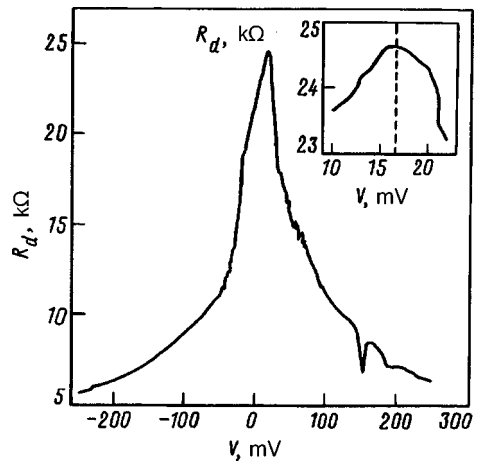

FIG. 4. $R_d(V)$ of contact No. 4 at $T=77$ K.

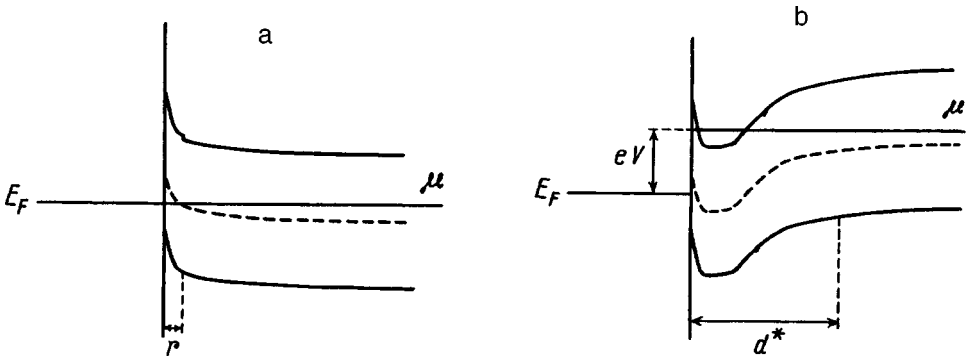

FIG. 5. Schematic diagram of the metal–Peierls-semiconductor interface: a) $V=0$; b) $V>0$.

interface for $V=0$ is shifted downwards relative to the center of the gap and therefore the semiconductor is *p*-type at the interface with the normal metal $K_{0.3}MoO_3$ and *n*-type in the bulk, where the material is a blue bronze.[10,11] That is, a p–n junction forms in a subsurface layer of the $K_{0.3}MoO_3$ as a result of the contact with the metal. Figure 5a shows a schematic band diagram illustrating this phenomenon. The bands bend upwards, since the work function in $K_{0.3}MoO_3$ is $\Phi\approx 3.5$ eV,[12] while in copper $\Phi=4.4$ eV. As was shown in Ref. 5, in the nonuniform case, such as we are dealing with here, two characteristic screening lengths for the electric field exist along the chains: a short length, r, of the order of the correlation length (~ 6 nm), and a macroscopic length, which is determined by the long decay lengths of the perturbations of the potential. The greatest change in the potential occurs over the distance r (see Fig. 5).

A linear dependence $R_d(V)$ will exist up to the moment when the deformation of the CDW reaches a critical value, with which we associate the observed sharp decrease of the resistance. It should be noted that the contacts *1* and *2* in Fig. 1 are distinguished by the nonuniformity of the critical deformation over the cross section of the contact. One can see that after the first break in the resistance a linear section is observed once again. Such a situation is possible if the magnitude of the critical deformation is different for different groups of chains included in a contact, which in the general case is a multiple microshort.

There are two known mechanisms for removing the accumulated deformation. The first one corresponds to the onset of phase slip during a dynamic process of sliding of the CDW (see, for example, Ref. 13 and references cited therein), which up to now has been regarded as a collective process occurring in the entire sample. Since the spreading in the transverse direction is negligibly small and the electric field in the bulk of the sample is known to be subthreshold, this mechanism requires assuming in our case the possibility of local sliding of a CDW over a quite short distance along the group of chains in the contact. The second mechanism presumes the formation of a defect of the electronic crystal — a phase soliton. However, as shown in Ref. 4, solitons have a low mobility and their drift does not make a large contribution to the conductivity. At present, we cannot suggest an exhaustive explanation of this phenomenon.

We can imagine the following band picture of the process. When a positive voltage is applied to the contact, the electrical potential along the chains varies with increasing distance x from the boundary as[6]

$$\varphi \propto \frac{x}{\sqrt{x^2 + d^{*2}}}.$$

As a result of the elastic deformation, the chemical potential at the M–PS interface shifts to the center of the band gap, and in a number of cases it crosses the center line. This process occurs over the characteristic length d^* for variation of the field, which is greater than r. In the bulk of the crystal the electric field is negligibly small, there is no deformation of the CDW, and the chemical potential shifts upwards together with the bands. The bending of the conduction band forms a potential well, whose dimension, on the basis of the quasi-one-dimensionality of the material under study, can be close to zero. Near its bottom the well can be quite narrow. As a result, by the moment that the critical deformation is reached, the band picture shown schematically in Fig. 5b obtains. As one can see from the figure, the region corresponding to the maximum shift of the chemical potential from the center of the band gap is removed from the M–PS interface and lies near the bottom of the potential well. The values of eV corresponding to the critical deformation for the contacts whose characteristics are shown in Figs. 3 and 4 equal 39 meV and 21 meV, respectively. An estimate, according to Ref. 9, of the shift of the chemical potential at $T = 77$ K for $K_{0.3}MoO_3$ gives $\Delta\mu \sim 15-20$ meV. The energy gap for $K_{0.3}MoO_3$ equals $2\Delta = 100$ meV.[14–16] Therefore the chemical potential certainly crosses the well for the contact shown in Fig. 3 and can cross the well for the contact shown in Fig. 4, provided that near the contact the energy gap is less than the equilibrium value. In view of the known pressure dependence of Δ, such a situation certainly happens for both contacts.[17] That is, applying pressure to the needle produces a local decrease of the energy gap, the needle pressure for the contact in Fig. 4 being higher (the resistance almost two times smaller) and hence Δ smaller than for the contact in Fig. 3. As the voltage increases further, no further deformation of the CDW occurs, but nonstationary processes can arise. However, it can be assumed that on average the band picture is preserved and is shifted upwards as a whole until the value of eV equals the difference of the work functions of the counterelectrodes and the potential well starts to flatten out. As the voltage increases, the depth of the potential well will change, which will inevitably affect the positions of the size-quantization energy levels arising in the well. We associate the observation of a sharp unipolar minimum in the curves in Fig. 3 and 4 with the moment when the position of the low-lying size-quantization energy level coincides with the position of the chemical potential with increasing voltage across the contact (with a change in the depth of the potential well). An indication of the fact that we are dealing with an energy level is that the width of the minimum corresponds to the temperature broadening kT. The voltages corresponding to the minima are different, which could be due to the different width of the potential well (the resistances of the contacts in Figs. 3 and 4 differ by almost a factor of 2). The vanishing of the feature as the contact resistance changes can also be easily explained, since as the diameter of the contact changes, the width of the potential well and therefore the positions of the levels in it change.

In summary, in this work it has been shown experimentally that a substantial shift of the chemical potential occurs near the boundary with the normal metal, while the deformation of the CDW by an applied electric field leads to strong bending of the energy bands and to the formation of a potential well, bringing the Fermi level into the region of

allowed single-electron states. The observed features of the IVCs of differential contacts could be due to to size-quantization.

We thank S. N. Artemenko, S. V. Zaĭtsev-Zotov, and V. Ya. Pokrovskiĭ for helpful discussions of the experimental results and P. Monceau for a discussion and for encouragement. This work was supported by the Russian Fund for Fundamental Research under Grants 97-02-17108 and 95-02-04456.

[a)]e-mail: gorl@web.cplire.ru

---

[1] G. Grüner, *Density Waves in Solids*, Addison-Wesley, Reading, Massachusetts 1994.
[2] S. N. Artemenko and F. Gleisberg, JETP Lett. **61**, 779 (1995).
[3] S. N. Artemenko and F. Gleisberg, Phys. Rev. Lett. **75**, 497 (1995).
[4] S. N. Artemenko, Zh. Éksp. Teor. Fiz. **111**, 1494 (1997) [JETP **84**, 823 (1997)].
[5] S. N. Artemenko, JETP Lett. **63**, 56 (1996).
[6] I. O. Kulik, A. N. Omel'yanchuk, and R. I. Shekhter, Fiz. Nizk. Temp. **3**, 1543 (1977) [Sov. J. Low Temp. Phys. **3**, 740 (1977)].
[7] A. A. Sinchenko, Fiz. Nizk. Temp. **15**, 438 (1989) [Sov. J. Low Temp. Phys. **15**, 247 (1989)].
[8] Yu. V. Sharvin, Zh. Éksp. Teor. Fiz. **48**, 984 (1965) [Sov. Phys. JETP **21**, 655 (1965)].
[9] S. N. Artemenko, V. Ya. Pokrovskiĭ, and S. V. Zaĭtsev-Zotov, Zh. Éksp. Teor. Fiz. **110**, 1069 (1996) [JETP **83**, 590 (1996)].
[10] C. Schlenker, J. Dumas, C. Escribe-Filippini *et al.*, Philos. Mag. B **52**, 643 (1985).
[11] L. Forró, J. R. Cooper, A. Jánossy, and K. Kamarás, Phys. Rev. B **34**, 9047 (1986).
[12] H. Matsuoka, K. Ohtake, R. Yamamoto *et al.*, Physica B **143**, 189 (1986).
[13] F. Ya. Nad', in *Charge Density Waves in Solids*, edited by L. Gor'kov and G. Grüner, Elsevier Science, Amsterdam 1989, Chapter 5.
[14] B. Dardel, D. Malterre, M. Grioni *et al.*, Europhys. Lett. **19**, 525 (1992).
[15] G. Travaglini, P. Wachter, J. Marcus, and C. Schlenker, Solid State Commun. **37**, 559 (1981).
[16] A. A. Sinchenko, Yu. I. Latyshev, S. G. Zybtsev *et al.*, JETP Lett. **64**, 285 (1996).
[17] G. Mihaly and P. Canfield, Phys. Rev. Lett. **64**, 459 (1990).

Translated by M. E. Alferieff

Nuclear fusion in the mesic molecule $d\mu^3$He

L. N. Bogdanova
Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

S. S. Gershtein
Institute of High Energy Physics, 142284 Protvino, Russia

L. I. Ponomarev
Kurchatov Institute Russian Science Center, 123182 Moscow, Russia; Research Coordinative Center on the Problem of Muon Catalyzed Fusion, 123098 Moscow, Russia

(Submitted 17 December 1997)
Pis'ma Zh. Éksp. Teor. Fiz. **67**, No. 2, 89–94 (25 January 1998)

A new scheme is presented for the physical processes leading to the nuclear fusion reaction $d(^3\text{He},^4\text{He})p$ catalyzed by a negatively charged muon μ^-. It is shown that the observable rate and yield of the nuclear reaction depend on a chain of ion–molecular reactions involving the participation of the $d\mu^3$He molecule. New calculations of the nuclear fusion rates in the $d\mu^3$He molecule are presented. © *1998 American Institute of Physics.* [S0021-3640(98)00102-9]

PACS numbers: 36.10.Gv, 25.55.−e

1. The nuclear fusion reaction

$$d+{}^3\text{He}\rightarrow{}^4\text{He}+p \tag{1}$$

is of interest for many reasons: it is involved in the primordial nucleosynthesis of the light elements in the early universe,[1] it has been discussed as a prospective source of thermonuclear energy,[2] it is a mirror reaction of the important process $d(t,^4\text{He})n$ (Ref. 3), etc. In all these cases it is especially important to know the cross sections of reaction (1) at low collision energies $E\lesssim 10$ keV, i.e., in the region where direct measurements in beam experiments are complicated. For these reasons, any alternative way to measure this value is interesting.

The phenomenon of muon catalysis affords the possibility of studying this reaction (as well as many other fusion reactions[4]) at practically zero collision energy from the mesic molecular state $(d\mu^3\text{He})^{++}$. (We will henceforth use the simplified notation $(d\mu^3\text{He})^{++}\equiv d\mu^3\text{He}$.) In recent years the rates λ_f^J of nuclear reactions

$$(d\mu^3\text{He})_J \xrightarrow[\searrow]{\lambda_f^J} \begin{array}{l}\mu^4\text{He}+p\\ \mu+{}^4\text{He}+p\end{array} \tag{2}$$

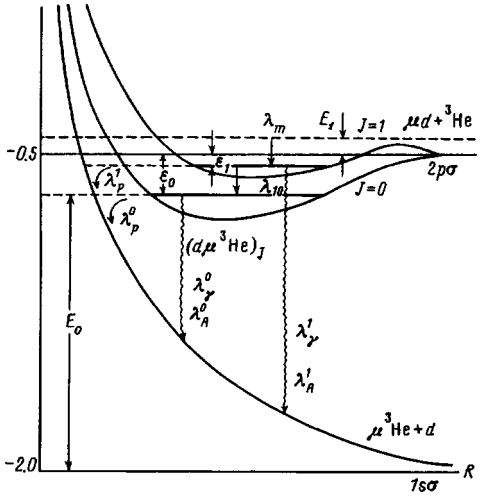

FIG. 1. Scheme of formation and decay of the $d\mu^3$He molecule. The muonic molecule $d\mu^3$He is formed in $d\mu+{}^3$He collisions in the bound state $(d\mu^3\mathrm{He})_{J=1}$ which decays to the system $\mu^3\mathrm{He}+d$ at the rates λ_γ^1, λ_p^1, and λ_A^1; the transition $(d\mu^3\mathrm{He})_{J=1}\to(d\mu^3\mathrm{He})_{J=0}$ with the rate λ_{10} competes with these decays; the binding energies of the states $(d\mu^3\mathrm{He})_J$ equal ε_J; the collision energies in the states $d\mu+{}^3$He and $\mu^3\mathrm{He}+d$ equal E_1 and E_0, respectively; the fusion rates from the states $J=0$ and $J=1$ are λ_f^0 and λ_f^1.

from the states $(d\mu^3\mathrm{He})_J$ with total angular momentum J have been calculated many times;[5–9] however, results of these calculations differ by several orders of magnitude. The most recent experimental upper limit for this rate is[10]

$$\lambda_f < 1.3\times 10^6\ \mathrm{s}^{-1}. \tag{3}$$

In this paper new results of λ_f^J calculations[11,12] and a new scheme for the kinetics of ion–molecular reactions preceding fusion in the $d\mu^3$He mesic molecule[13] are used.

2. The scheme of the processes of formation and decay of the $(d\mu^3\mathrm{He})_J$ mesic molecule is presented in Fig. 1: mesic molecules $(d\mu^3\mathrm{He})_J$ are formed in collisions of slow $d\mu$-atoms with ^{3}He atoms;[14] the states $(d\mu^3\mathrm{He})_J$ are quasistationary on account of decay to $(\mu^3\mathrm{He})_{1s}+d$ by γ-emission (4), Auger transitions (5), and predissociation (6), with rates λ_γ^J, λ_A^J and λ_p^J, respectively:

$$(d\mu^3\mathrm{He})_J\to(\mu^3\mathrm{He})_{1s}+d+\gamma, \tag{4}$$

$$(d\mu^3\mathrm{He})_J\to(\mu^3\mathrm{He})_{1s}+d+e, \tag{5}$$

$$(d\mu^3\mathrm{He})_J\to(\mu^3\mathrm{He})_{1s}+d. \tag{6}$$

Besides that, in collisions of $[(d\mu\mathrm{He})e]^+$ with D_2 and He in the chain of ion–molecular reactions,[13] transitions $(J=1)\to(J=0)$ with a rate λ_{10} can occur:

$$(d\mu^3\mathrm{He})_{J=1}\to(d\mu^3\mathrm{He})_{J=0}. \tag{7}$$

The yield N_f of nuclear fusion per stopped μ^- is determined by the fusion rates λ_f^J from the states J and by the populations w_J of these states, which depend on the rates of processes (4)–(7) and on the kinetics of the ion–molecular reactions in which the $d\mu^3$He molecule participates.

3. The rates λ_f^J are determined by the relation[4]

$$\lambda_f^J = \sum_L A_L G_L^J . \tag{8}$$

Here the A_L are reaction constants for the nuclear states with orbital angular momentum L, determined by extrapolation of the cross sections of reaction (1) to zero collision energy, and the quantities G_0^J and G_1^J are calculated with the formulas:

$$G_0^J = \int d\mathbf{r} |\Psi^J(\mathbf{r},\mathbf{R}=0)|^2, \tag{9}$$

$$G_1^J = \int d\mathbf{r} |\nabla_{\mathbf{R}} \Psi^J(\mathbf{r},\mathbf{R})|^2_{\mathbf{R}=0}, \tag{10}$$

where $\Psi^J(\mathbf{r},\mathbf{R})$ is the wave function of the $(d\mu^3\mathrm{He})_J$ mesic molecule ($\mathbf{R}$ is the internuclear distance and $\mathbf{r}$ is the muon coordinate with respect to the center of mass of the nuclei). Since $A_0 \gg A_1$, in the following we are interested only in the values of G_0^J.

To calculate the values of G_0^J two independent methods have been used and two high-accuracy numerical algorithms have been developed,[11,15] which give results that are in reasonable agreement. The muon catalysis reaction (2), as compared to the mirror reaction $d\mu t \to \mu^4\mathrm{He} + n$, $\mu + {}^4\mathrm{He} + n$ (Ref. 16), has some peculiarities due to the presence of an open channel in the system $(d\mu^3\mathrm{He})_J$. Specifically, unlike the $(d\mu t)_J$ mesic molecule, where the bound states are predominantly localized in the potential $W_{1s\sigma}(R)$ formed by muon motion in the state with quantum numbers $(Nlm)=(100)$ of the system $(t\mu)_{1s}+d$, the states of the $(d\mu^3\mathrm{He})_J$ molecule are localized in the potential $W_{2p\sigma}(R)$ with quantum numbers $(Nlm)=(210)$ of the system $(d\mu)_{1s}+{}^3\mathrm{He}$.

Unlike the case of $d\mu t$, however, due to the strong coupling between the channels $1s\sigma$ and $2p\sigma$, the wave function $\Psi^J(\mathbf{r},\mathbf{R})$ contains[11,15] all the components $\psi^L(\mathbf{R})$ representing the relative motion of nuclei with different L in the $d\mu^3\mathrm{He}$ molecule.

In the limit $R \to 0$ it has the form[17]

$$\Psi_m^J(\mathbf{r},\mathbf{R}) \approx_{\mathbf{R}\to 0} \sum_j \phi_j(\mathbf{r};R)\psi_j^L(\mathbf{R}) = \sum_{Nl} \phi_{Nlm}(\mathbf{r};R)\psi_{Nlm}^L(\mathbf{R}), \tag{11}$$

where $\phi_j(\mathbf{r};R)$ are orthonormalized adiabatic basis functions, and the functions $\psi_j^L(\mathbf{R})$ represent the relative motion of nuclei with angular momentum $L=|J-l|, \ldots, |J+l|$ in the potential $W_j(\mathbf{R})$ formed by the muon motion in the quantum state $j=(Nlm)$ at a fixed distance R between nuclei.[18] Thus for states $J=0$ both combinations $(l=0, L=0)$ and $(l=1, L=1)$ are essential

$$\Psi^{J=0}(\mathbf{r};\mathbf{R}) \approx_{\mathbf{R}\to 0} \sum_N \phi_{N00}(\mathbf{r};R)\psi_{N00}^{L=0}(\mathbf{R}) + \sum_N \phi_{N10}(\mathbf{r};R)\psi_{N10}^{L=1}(\mathbf{R}). \tag{12}$$

For $J=1$ the analogous expansion has the form:[17]

$$\Psi_m^{J=1}(\mathbf{r},\mathbf{R}) \approx_{\mathbf{R}\to 0} \sum_N \delta_{0m}\phi_{N00}(\mathbf{r};R)\psi_{N00}^{L=1}(\mathbf{R}) + \sum_N \phi_{N1m}(\mathbf{r};R)\psi_{N1m}^{L=0}(\mathbf{R}). \tag{13}$$

It follows from definitions (9), (12), and (13) that

$$G_0^0=\sum_N |\psi_{N00}^{L=0}(0)|^2, \tag{14}$$

$$G_0^1=\sum_{N,m} |\psi_{N1m}^{L=0}(0)|^2. \tag{15}$$

The functions $\psi_{Nlm}^L(\mathbf{R})$ have recently been calculated in Ref. 11 by the complex coordinate rotation method, by expanding a variational function $\Psi^J(\mathbf{r};\mathbf{R})$ over the adiabatic basis.

The calculated G_0^J factors have the values

$$G_0^0=0.63\times 10^{-12}a_\mu^{-3}=3.8\times 10^{19}\ \mathrm{cm}^{-3}, \tag{16}$$

$$G_0^1=0.86\times 10^{-15}a_\mu^{-3}=5.1\times 10^{16}\ \mathrm{cm}^{-3} \tag{17}$$

($a_\mu=2.56\times 10^{-11}$ cm is a mesic atomic length unit). In the other approach[12,15] which employs an expansion of the function $\Psi^J(\mathbf{r},\mathbf{R})$ over the adiabatic hyperspherical basis,[19] the finite width of the quasistationary $d\mu^3$He state was explicitly taken into account. The result obtained,

$$G_0^0=0.75\times 10^{-12}a_\mu^{-3}=4.4\times 10^{19}\ \mathrm{cm}^{-3}, \tag{18}$$

is in reasonable agreement with (16).

The reaction constant for unpolarized nuclei equals $A_0=0.34\times 10^{-14}\ \mathrm{cm}^3\cdot\mathrm{s}^{-1}$ (Ref. 3).

The low energy cross section of reaction (1) is dominated by the $J^P=3/2^+$ near-threshold resonance. For this reason the nuclear fusion rates λ_f^J have the values[20]

$$\lambda_f^0=3/2\cdot A_0G_0^0=1.9\times 10^5\ \mathrm{s}^{-1}, \tag{19}$$

$$\lambda_f^1=0.65\times 10^3\ \mathrm{s}^{-1}. \tag{20}$$

With G_0^0 from (18) the rate is

$$\lambda_f^0=2.3\times 10^5\ \mathrm{s}^{-1}. \tag{21}$$

4. Mesic molecules $d\mu^3$He are formed in the reaction

$$(d\mu)_{1s}+{}^3\mathrm{He}\rightarrow(d\mu^3\mathrm{He})_J^+ + e, \tag{22}$$

predominantly in the state with total angular momentum $J=1$ by an $E1$ dipole transition.[21] Since $\lambda_f^1\ll\lambda_f^0$ the fusion reaction in the $d\mu^3$He mesic molecule can be observed only in the state $J=0$. It can occur if the rate λ_{10} of reaction (7) is comparable to the decay rates $\lambda_\gamma^1,\lambda_p^1$, and λ_A^1 of the state $J=1$. The internal Auger transition

$$(d\mu^3\mathrm{He})_{J=1}\rightarrow(d\mu^3\mathrm{He})_{J=0}^+ + e \tag{23}$$

is forbidden, since the difference between energies of states $J=1$ and $J=0$, viz., $\epsilon_1-\epsilon_0=22.8$ eV, is less than the ionization energy of a helium atom (24.6 eV). Hence transition (7) can occur only in collisions of $(d\mu^3\mathrm{He})_J$ with atoms of the medium. The whole set of ion–molecular reactions leading to transition (7) has been considered in Ref.

TABLE I. Rates of the main processes in the $(d\mu^3\mathrm{He})$ muonic molecule.

Reaction	λ	J	λ, s^{-1}	Ref.
$d\mu+{}^3\mathrm{He}\rightarrow(d\mu^3\mathrm{He})_{J\upsilon}$	λ_m^J	1	1.4×10^8	21
		0	$\sim2\times10^6$	13
$(d\mu^3\mathrm{He})_J\rightarrow\mu^4\mathrm{He}+p$	λ_f^J	1	0.65×10^3	11
		0	1.9×10^5	
$(d\mu^3\mathrm{He})_J\rightarrow\mu^3\mathrm{He}+d+\gamma$	λ_γ^J	1	1.6×10^{11}	24
		0	1.8×10^{11}	
$(d\mu^3\mathrm{He})_J\rightarrow\mu^3\mathrm{He}+d$	λ_p^J	1	0.8×10^{12}	25
		0	0.7×10^{12}	
$(d\mu^3\mathrm{He})_J\rightarrow\mu^3\mathrm{He}+d+e$	λ_A^J	1	0.41×10^{11}	26
		0	0.47×10^{11}	
$(d\mu^3\mathrm{He})_{J=1}\rightarrow(d\mu^3\mathrm{He})_{J=0}$	λ_{10}	1→0	0.5×10^{12}	13

13. These reactions ($i=1$–9) and their rates λ_i are listed below, where the notation $M_J=(d\mu^3\mathrm{He})_J$ is introduced (φ is the mixture density, C_{He} is the helium concentration and $\mathrm{X}=\mathrm{D}_2,\mathrm{He}$).

$$
\begin{aligned}
&(1)\ (M_1e)^+ + \mathrm{D}_2 + \mathrm{X}\rightarrow(M_1e\,\mathrm{D}_2)^+ + \mathrm{X}, && \lambda_1\approx3\times10^{13}\varphi\ \mathrm{s}^{-1};\\
&(2)\ (M_1e)^+ + \mathrm{He} + \mathrm{X}\rightarrow(M_1e\mathrm{He})^+ + \mathrm{X}, && \lambda_2\approx3\times10^{13}\varphi C_{\mathrm{He}}\ \mathrm{s}^{-1};\\
&(3)\ (M_1e\,\mathrm{D}_2)^+ + \mathrm{He}\rightarrow(M_1e\mathrm{He})^+ + \mathrm{D}_2, && \lambda_3\approx3\times10^{13}\varphi C_{\mathrm{He}}\ \mathrm{s}^{-1};\\
&(4)\ (M_1e\mathrm{He})^+ + \mathrm{He}\rightarrow(M_1ee)+\mathrm{He}_2^+, && \lambda_4\approx10^{13}\varphi C_{\mathrm{He}}\ \mathrm{s}^{-1};\\
&(5)\ (M_1e\,\mathrm{D}_2)^+\rightarrow(M_0e)^+ + \mathrm{D}_2^+ + e, && \lambda_5\approx5\times10^{11}\ \mathrm{s}^{-1};\\
&(6)\ (M_1e\mathrm{He})^+ + \mathrm{D}_2\rightarrow(M_0e\mathrm{He})^+ + \mathrm{D}_2^+ + e, && \lambda_6\lesssim10^9\varphi\ \mathrm{s}^{-1};\\
&(7)\ (M_1ee) + \mathrm{D}_2\rightarrow(M_0ee) + \mathrm{D}_2^+ + e, && \lambda_7\approx4\times10^9\varphi\ \mathrm{s}^{-1};\\
&(8)\ (M_1e)^+ + \mathrm{He}\rightarrow(M_1ee)+\mathrm{He}^+, && \lambda_8\approx3\times10^{12}\varphi C_{\mathrm{He}}\ \mathrm{s}^{-1};\\
&(9)\ (M_1e)^+ + \mathrm{D}_2\rightarrow(M_0e)^+ + \mathrm{D}_2^+ + e, && \lambda_9\approx10^7\varphi\ \mathrm{s}^{-1}.
\end{aligned}
\tag{24}
$$

Accordingly, the total rates $\lambda_{\mathrm{dec}}^J=\lambda_\gamma^J+\lambda_p^J+\lambda_A^J$ of decay of the quasistationary states $(d\mu^3\mathrm{He})_J$ in channels (4)–(6) have the values (see Table I):

$$\lambda_{\mathrm{dec}}^0\simeq0.9\times10^{12}\ \mathrm{s}^{-1}, \tag{25}$$

$$\lambda_{\mathrm{dec}}^1\simeq1\times10^{12}\ \mathrm{s}^{-1}. \tag{26}$$

It follows from a comparison of the rates λ_i that at $\varphi\sim0.1$ and $C_{\mathrm{He}}\leqslant0.1$ the dominant channel in the chain of ion–molecular reactions leading to a change of the angular momentum of the mesic molecule, $(J=1)\rightarrow(J=0)$, involves the formation of the complex $[(d\mu^3\mathrm{He})_{J=1}e\,\mathrm{D}_2]^+$ and its subsequent decay with conversion of an electron of the D_2 molecule, namely:

$$[(d\mu^3\mathrm{He})_{J=1}e]^+\overset{\lambda_1}{\rightarrow}[(d\mu^3\mathrm{He})_{J=1}e\,\mathrm{D}_2]^+\overset{\lambda_5}{\rightarrow}[(d\mu^3\mathrm{He})_{J=0}e]^+ + \mathrm{D}_2^+ + e. \tag{27}$$

In comparison with this process the rate $\lambda_7 \lesssim 10^9 \varphi\ \mathrm{s}^{-1}$ of the reaction $(J=1)\to(J=0)$ by the external Auger transition (reaction (*7*) of (24)) is negligibly small, in contrast to the statement of Ref. 22.

At $\varphi \simeq 0.1$ and $C_{\mathrm{He}} \lesssim 0.1$ one has $\lambda_1 \simeq 3\lambda_{\mathrm{dec}}$ and $\lambda_5 \sim \lambda_3$, i.e., a noticeable fraction (~ 0.2) of the mesic molecules $(d\mu^3\mathrm{He})_{J=1}$ reach the state $J=0$, where the fusion (2) can be observed.

5. The detailed analysis of the kinetics of processes in $\mathrm{D}_2+{}^3\mathrm{He}$ mixtures is yet to be done, but even simple estimates can give rather reliable information about the expected yield N_f of fusion reactions (2) per muon stop. According to these estimates

$$N_f \simeq \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_8+\lambda_{\mathrm{dec}}^1}\,\frac{\lambda_5}{\lambda_5+\lambda_3+\lambda_{\mathrm{dec}}^1}\,\frac{\lambda_f^0}{\lambda_f^0+\lambda_{\mathrm{dec}}^0}\,n_f$$

$$\approx n_f \frac{\lambda_f^0}{\lambda_{\mathrm{dec}}^0}(1+0.03\varphi^{-1})^{-1}(2.8+60\varphi C_{\mathrm{He}})^{-1}, \tag{28}$$

where the coefficient $n_f = n_f(C_d)$ is the statistical weight of the states of the mesic molecule $(d\mu^3\mathrm{He})_J$ with total nuclear spin $S=3/2$, calculated by taking into account the kinetics of its formation in collisions $(d\mu)_F+\mathrm{He}$ and the spin-flip processes $(d\mu)_F+d \to (d\mu)_{F'}+d$ (Refs. 20 and 23). At $\varphi=0.075$, $C_{\mathrm{He}}=0.05$, and $n_f \simeq 0.5$ one has

$$N_f \simeq 0.12\cdot\lambda_f^0/\lambda_{\mathrm{dec}}^0 \approx 3\times 10^{-8}/\mu^-. \tag{29}$$

6. The understanding and quantitative description of nuclear fusion catalysis in the $d\mu^3\mathrm{He}$ mesic molecule has required the development of new theoretical methods and the consideration of new physical processes. Experiment R-94-05.1 planned at PSI will permit a check of the correctness and self-consistency of these methods and the adequacy of the processes considered. In particular, observation of the φ dependence (28) of the yield N_f would be confirmation of the ion–molecular mechanism of the transition $(d\mu^3\mathrm{He})_{J=1}\to(d\mu^3\mathrm{He})_{J=0}$ via formation of clusters $[(d\mu^3\mathrm{He})e\ \mathrm{D}_2]^+$. A comparison of the fusion rate λ_f^0 extracted from measurements of N_f with its theoretical values will test the validity of the sophisticated calculations of the three-body Coulomb problem performed recently.

The methods and details of the theoretical calculations of the rates λ_f^J and λ_i will be published elsewhere.[11–13,15,20] A preliminary version of this paper was published in Ref. 27.

The authors are grateful to D. I. Abramov, M. P. Faifman, V. V. Gusev, V. I. Korobov, and L. I. Menshikov for multifaceted help and for information about their results prior to publication, and to C. Petitjean and A. A. Vorobyov for some stimulating and challenging discussions.

This work was supported in part (L. N. B. and L. I. P.) by Grant 96-02-17279 of the RFBR.

[1] A. M. Boesgaard and G. Steigman, Annu. Rev. Astron. Astrophys. **3**, 313 (1985).
[2] I. N. Golovin and B. B. Kadomtsev, Atomnaya Énergiya **81**, 384 (1996).
[3] H. S. Bosch and G. M. Hale, Nucl. Fusion **32**, 611 (1992).
[4] L. N. Bogdanova, Muon Catal. Fusion **2**, 359 (1988).

[5] A. V. Kravtsov, N. P. Popov, and G. E. Solyakin, JETP Lett. **40**, 875 (1984).
[6] D. Harley, B. Muller, and J. Rafelsky, J. Phys. **616**, 281 (1990).
[7] Y. Kino and M. Kamimura, Hyperfine Interact. **82**, 195 (1993).
[8] W. Czaplinski, A. Kravtsov, A. Mikhailov, and N. Popov, Phys. Lett. A **219**, 86 (1996).
[9] F. M. Penkov, Yad. Fiz. **60**, 1003 (1997) [Phys. Atomic Nucl. **60**, 897 (1997)].
[10] D. V. Balin, V. A. Ganza, E. M. Maev *et al.*, Phys. Lett. (submitted).
[11] L. N. Bogdanova, V. I. Korobov, and L. I. Ponomarev, Few Body Syst. (submitted).
[12] S. S. Gershtein and V. V. Gusev, Yad. Fiz. (submitted).
[13] L. I. Menshikov and M. P. Faifman, Z. Phys. (submitted).
[14] Yu. A. Aristov, A. V. Kravtsov, N. P. Popov *et al.*, Yad. Fiz. **33**, 1066 (1981) [Sov. J. Nucl. Phys. **33**, 564 (1981)].
[15] D. I. Abramov and V. V. Gusev, Few Body Syst. (submitted).
[16] L. N. Bogdanova, V. E. Markushin, V. S. Melezhik, and L. I. Ponomarev, Yad. Fiz. **34**, 1191 (1981) [Sov. J. Nucl. Phys. **34**, 662 (1981)].
[17] S. I. Vinitskiĭ and L. I. Ponomarev, Fiz. Elem. Chast. At. Yadra **13**, 1336 (1982) [Sov. J. Part. Nucl. **13**, 557 (1982)].
[18] I. V. Komarov, L. I. Ponomarev, and S. Yu. Slavyanov, *Spheroidal and Coulomb Spheroidal Functions* [in Russian], Nauka, Moscow, 1976.
[19] D. I. Abramov, V. V. Gusev, and L. I. Ponomarev, Yad. Fiz. **60**, 1259 (1997) [Phys. Atomic Nucl. **60**, 1133 (1997)].
[20] L. N. Bogdanova, S. S. Gershtein, and L. I. Ponomarev, Z. Phys. (submitted).
[21] V. K. Ivanov, A. V. Kravtsov, A. I. Mikhailov *et al.*, Zh. Éksp. Teor. Fiz. **91**, 358 (1986) [Sov. Phys. JETP **64**, 210 (1986)].
[22] W. Czaplinski, M. Filipowicz, E. Gula *et al.*, Z. Phys. D **37**, 283 (1996).
[23] L. N. Bogdanova, V. E. Markushin, V. S. Melezhik, and L. I. Ponomarev, Zh. Éksp. Teor. Fiz. **83**, 1615 (1982) [Sov. Phys. JETP **56**, 931 (1982)].
[24] S. Hara and T. Ishihara, Phys. Rev. A **39**, 5633 (1989).
[25] S. S. Gershtein and V. V. Gusev, Hyperfine Interact. **82**, 205 (1993).
[26] A. V. Kravtsov, A. I. Mikhailov, and V. I. Savichev, Z. Phys. D **29**, 49 (1994).
[27] L. N. Bogdanova, S. S. Gerstein, and L. I. Ponomarev, Preprint PSI PR-97-33, October 1997.

Published in English in the original Russian journal. Edited by Steve Torstveit.