JETP Letters, Vol. 81, No. 10, 2005, pp. 485–486. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 603–604.
Original Russian Text Copyright © 2005 by Dubovskiĭ.

Analog of the Migdal–Kohn Singularity in the Polariton Spectrum of a Two-Dimensional Crystal

O. A. Dubovskiĭ

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow region, 142190 Russia

e-mail: in@isan.troitsk.ru

Received April 8, 2005

It is shown that, in a two-dimensional crystal, the radiative polariton branch with an anomalously large radiative width (superradiance) exhibits an analog of the Migdal–Kohn singularity in the phase space at the point of joining with the discovered additional branch. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 71.35.Aa, 71.36.+c

It is shown in [1, 2] that, because of the retarded interaction of Coulomb-type Frenkel excitons and phonons, high-frequency branches in the polariton spectra of 1D and 2D crystals exhibit anomalously large radiative half-widths $\Gamma_1 \cong \gamma(\lambda/a)$ and $\Gamma_2 \cong \gamma(\lambda/a)^2$, where γ is the radiative half-width of the line of optical dipole radiation with the frequency ω_0 for the isolated monomer, λ is the wavelength, and $a \ll \lambda$ is the lattice constant. This effect (superradiance, $\Gamma_{1,2} \gg \gamma$) was revealed experimentally in a number of works using picosecond lasers [3]. In [1, 2], it was assumed that the radiative branches must have end points of dispersion curves at which the anomalously large broadening is already comparable with the frequency of the radiative branch. However, it has been recently shown [4] that the radiative branch in a 1D crystal at a certain point of the phase space (the end point of the spectrum) is continuously joined at an infinite group velocity (an analog of the Migdal–Kohn singularity) with the newly revealed branch (an additional solution of the corresponding dispersion equation). In this case, the anomalously large radiative width is nevertheless considerably smaller than the frequency of the polariton branch. This behavior of two branches joining with an infinite derivative is most clearly seen from the exact analytical solution of the dispersion equation for the slope of dipoles to the chain axis equal to arccos $(1/\sqrt{3})$. This solution gives an inverted representation of the dispersion dependence of the wave vector $k = k(\omega)$ on the frequency ω [4]

$$k = \frac{1}{a}\arccos\left(\cos\left(\frac{\omega a}{c}\right) + \frac{1}{2}\exp\left(\frac{\omega^2 - \omega_0^2}{(2/3)T\omega^2}\right)\right), \quad (1)$$

where the constant $T = 2|P|^2\omega_0/\hbar ac^2$ involves the matrix element P of the dipole transition in the monomer.

Of interest is the fundamental question of whether the spectral features indicated in [4] for 1D crystals also exist in 2D crystals. There is no exact analytical representation of the corresponding dipole–dipole sums for 2D crystals similar to that used in the derivation of Eq. (1); therefore, the corresponding dispersion equation is obtained in this work and its computer solution is presented. A 2D crystal with a square lattice and dipole transitions in the monomers orthogonal to the crystal plane is considered. The Fourier component of the orthogonal component of the electric field $E_\omega(\mathbf{r}_1)$ at site $\mathbf{r}_1$ of the emission from dipole $d_\omega(\mathbf{r}_2)$ at site $\mathbf{r}_2$ is represented in the classical field theory as follows:

$$E_\omega = \left(\frac{\omega^2}{Rc^2} + \frac{i\omega}{R^2c} - \frac{1}{R^3}\right)\exp\left(i\frac{\omega}{c}R\right)d_\omega, \quad R = |\mathbf{r}_1 - \mathbf{r}_2|. \quad (2)$$

In the second quantization representation of the quantum field theory with the Hamiltonian quadratic in the Bose creation and annihilation operators of photons and the Bose operators of site excitations is diagonalized using the conventional Bogolyubov–Tyablikov u, v transformation [4]; after some calculations for wave functions with the wave vector $\mathbf{k}$ (k_x, k_y) and the energy $E = \hbar\omega$, the following dispersion equation is obtained:

$$\omega^2 - \omega_0^2 + 2i\gamma\omega_0 = S(\omega, k_x, k_y),$$

$$S = 4T\left(\frac{c}{a}\right)^2 \times \sum_{n=0}^{N}\sum_{m=0}^{N}\Delta_{nm}\left[\frac{1}{r_{nm}^3} - i\left(\frac{\omega a}{c}\right)\frac{1}{r_{nm}^2} - \left(\frac{\omega a}{c}\right)^2\frac{1}{r_{nm}}\right] \times \cos(k_x r_{nm} a)\cos(k_y r_{nm} a)\exp\left(i\frac{\omega r_{nm} a}{c}\right). \quad (3)$$

In Eq. (3), $\gamma = (1/3)T(\omega_0 a/c)\omega_0$ is the intrinsic radiative half-width of the emission line of an isolated oscillator,

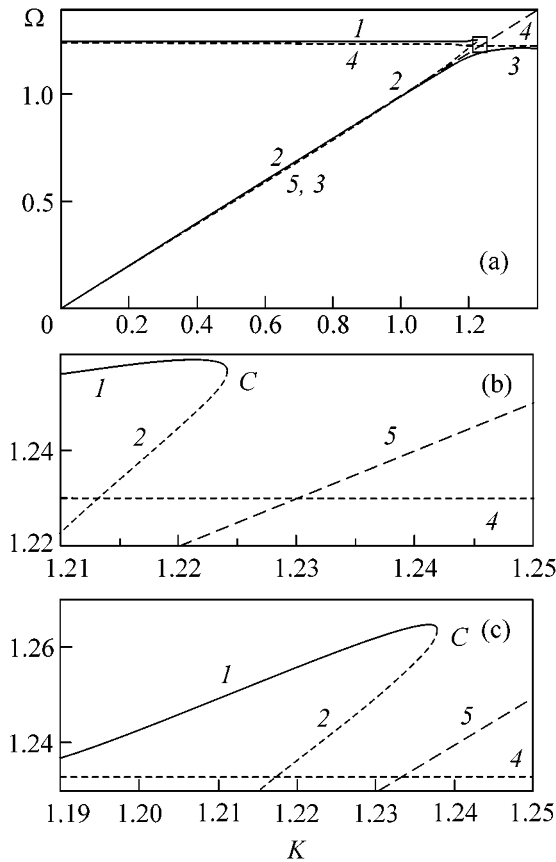

Fig. 1. Polariton dispersion curves.

$r_{nm} = (n^2 + m^2)^{1/2}$, and $\Delta_{nm} = 1 - \delta_{n0}\delta_{m0}$. In an exact analytical approach, it is necessary to assume that $N = \infty$, and, in numerical calculations, this value should be taken so large that the results would not change upon its increase. This condition was verified. The imaginary part of the corresponding Green's function $G = (\omega^2 - \omega_0^2 + 2i\gamma\omega_0 - S(\omega, k_x, k_y))^{-1}$ at $S = S' + iS''$ determines the density of states and the absorption coefficient $\chi = 2\Gamma\omega_0/[(\omega^2 - \omega_0^2 - S')^2 + (2\Gamma\omega_0)^2]$, where the radiative half-width $\Gamma(\omega, k_x, k_y)\gamma - (1/2\omega_0)S''(\omega, k_x, k_y)$. The dispersion dependence of the branches $\omega = \omega(k_x, k_y)$ determining the position of resonance peaks was found by numerically solving the equation (it is difficult to solve it analytically)

$$\omega^2 - \omega_0^2 - S'(\omega, k_x, k_y) = 0. \quad (4)$$

The calculated results for dispersion dependences $\Omega = \Omega(K)$ are displayed in Fig. 1 for the dimensionless quantities $\Omega = \omega/\omega_0$ and $K = (c\sqrt{k_x^2 + k_y^2}/\omega_0)$ at the values of the parameters $T = 4 \times 10^{-4}$, $\omega_0 a/c = 0.1$, and $N = 10^3$, which are close to real optical characteristics and provide sufficiently good convergence of the series. Figure 1 presents the general pattern of the dispersion dependences of frequencies on the wave vector for various branches in the direction $k_x \neq 0$, $k_y = 0$. Curve *1* represents the radiatively decaying high-frequency polariton branch ($\Omega_1 > K$), curve *2* is the new branch ($\Omega_2 > K$) joining with *1* in the distinguished square, and curve *3* is the nonradiative polariton branch ($\Omega_3 < K$). Curve *4* represents the Coulomb exciton band with the retarded components excluded in Eq. (3), and curve *5* is the photonic straight line ($\Omega = K$). It is evident that the linear dependence $\Omega_1(K)$ at small K, postulated in [1, 2], is not observed. This result can actually be obtained analytically because of the compensation of the corresponding terms. Because some curves are extremely close, the most interesting area in Fig. 1a distinguished by a square is represented in Fig. 1b at a larger scale. It is evident that curve *1* first rises with increasing K, then drops, and continuously joins the ascending curve *2* at point C with infinite derivatives, which is confirmed by the representation of the results at even larger scale. At this limiting point, the decay of branch *1* is anomalously high $\Gamma = 1.53\gamma(\lambda/a)^2$ (superradiance). Calculations of the frequency dependence of the absorption coefficient reveal two peaks, joining at point C with increasing K and then vanishing. Because the group velocity for branch *2* $v > c$, this branch may be considered to be nonphysical [4]. It is interesting only to the extent that it determines the limiting behavior of branch *1*. However, note that $v > c$ was experimentally maintained in [5]. Dispersion dependences for the direction $k_x = k_y$ are presented in Fig. 1c. It is evident that the dispersion of the same curves changes only qualitatively. Dependences intermediate between those presented in Figs. 1b and 1c are observed for nonsymmetrical directions.

I am grateful to V.M. Agranovich, A.M. Kamchatnov, and Yu.E. Lozovik for useful comments. This work was supported by the Russian Foundation for Basic Research, project no. 03-02-16293.

REFERENCES

1. V. M. Agranovich and O. A. Dubovskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **3**, 345 (1966) [JETP Lett. **3**, 223 (1966)].
2. A. M. Agranovich, *Theory of Excitons* (Nauka, Moscow, 1968) [in Russian].
3. B. Deveaud, F. Clerot, N. Roy, *et al.*, Phys. Rev. Lett. **67**, 2355 (1991).
4. O. A. Dubovskiĭ, Zh. Éksp. Teor. Fiz. **125**, 272 (2004) [JETP **98**, 240 (2004)].
5. N. Brunner, V. Scarani, M. Wegmüller, *et al.*, Phys. Rev. Lett. **93**, 203 902 (2004).

Translated by A. Bagatur'yants

JETP Letters, Vol. 81, No. 10, 2005, pp. 487–490. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 605–609.
Original Russian Text Copyright © 2005 by Serebryannikov, Ming-Lie Hu, Yan-Feng Li, Ching-Yue Wang, Zhuan Wang, Lu Chai, Zheltikov.

Enhanced Soliton Self-Frequency Shift of Ultrashort Light Pulses

E. E. Serebryannikov[1], Ming-Lie Hu[2], Yan-Feng Li[2], Ching-Yue Wang[2], Zhuan Wang[2], Lu Chai[2], and A. M. Zheltikov[1]

[1] *Faculty of Physics, International Laser Center, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
e-mail: zheltikov@phys.msu.ru

[2] *Ultrafast Laser Lab, School of Precision Instruments and Optoelectronics Engineering, Key Laboratory of Optoelectronic Information Technical Science, Tianjin University, 300072 Tianjin, Peoples Republic of China*

Received April 8, 2005

Photonic-crystal fibers are used to study scenarios of soliton self-frequency shift for laser pulses with initial pulse lengths much less than the Raman-mode period of the fiber material. A typical frequency shift of subnanojoule Ti: sapphire-laser pulses with an initial duration of about 30 fs transmitted through a fiber with a core diameter of about 1.6 μm and a length of about 7 cm exceeds 100 THz. The rate of soliton self-frequency shift is radically increased by reducing the initial pulse width. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 42.65.Wi, 42.81.Qb

Nonstationary spectral transformations of ultrashort laser pulses in media with noninstantaneous optical nonlinearity are often accompanied by a continuous downshift of the central frequency [1, 2]. For Raman-active media, this effect is illustratively interpreted in the frequency domain [3] as the amplification of the low-frequency part of the spectrum of an ultrashort pulse by its high-frequency wing through stimulated Raman scattering (SRS). As a result, the carrier frequency of an ultrashort pulse propagating through a Raman medium in the soliton regime undergoes a permanent red shift. This phenomenon, called soliton self-frequency shift (SSFS), allows the creation of fiber-optic frequency shifters for ultrashort laser pulses.

The initial pulse length is one of the key parameters controlling the regime of SSFS. According to the SSFS analysis based on the nonlinear Schrödinger equation with an assumption of a linear frequency dependence of the Raman gain $R(\nu)$ [4], the shift rate of the central frequency ν_0 of the soliton propagating along the z axis is highly sensitive to the soliton duration τ_0: $d\nu_0/dz \propto \tau_0^{-4}$. Deviation of $R(\nu)$ from a linear function and high-order dispersion effects can violate the scaling $d\nu_0/dz \propto \tau_0^{-4}$. It would be natural to expect that deviations from this scaling law should become quite dramatic for pulses with an initial duration τ_0 less than the period τ_R of the Raman mode ($\tau_R \approx 80$ fs for fused silica [3, 5]). Experimental studies of the SSFS in the regime where $\tau_0 \ll \tau_R$ have demonstrated [6] that spectral shifts exceeding 180 THz can be achieved for pulses with an initial duration $\tau_0 \approx 6$ fs and an appropriate profile of the spectral phase. An efficient coupling of ultrashort laser pulses into a frequency-shifted soliton becomes possible, as shown by Teisset *et al.* [6], due to a special dispersion profile provided by a photonic-crystal fiber (PCF) [7–9]. Observation of SSFS in a tapered PCF has been earlier reported by Liu *et al.* [10]. A strong confinement of the laser field in a PCF core due to a large refractive index step from the silica PCF core to the microstructured cladding enhances the optical nonlinearity and, as a consequence, reduces the energy of laser radiation and the fiber length required for soliton formation [6, 10–14].

In this work, we examine specific features of soliton self-frequency shift of subnanojoule ultrashort laser pulses in PCFs in the regime where $\tau_0 < \tau_R$. The results of our experiments show that the SSFS rate can be substantially increased by reducing the input pulse width. Soliton frequency shifts exceeding 100 THz have been achieved in our studies for subnanojoule Ti: sapphire-laser pulses with an initial pulse width of about 30 fs transmitted through a PCF with a core diameter of about 1.6 μm and a length of about 7 cm.

Our analysis of soliton dynamics in PCFs is based on the numerical solution of the generalized nonlinear Schrödinger equation [3]

$$\frac{\partial A}{\partial z} = i\sum_{k=2}^{6}\frac{(i)^k}{k!}\beta^{(k)}\frac{\partial^k A}{\partial \tau^k} + i\gamma\left(1+\frac{i}{\omega_0}\right)A(\tau, z)\Psi(z, \tau), \quad (1)$$

where A is the pulse envelope, $\beta^{(k)} = \partial^k\beta/\partial\omega^k$ are the coefficients in the Taylor-series expansion for the propagation constant β of the relevant guided mode, ω_0 is the carrier frequency, τ is the time in the retarded frame of reference, $\gamma = n_2\omega_0/cS_{\text{eff}}$ is the nonlinearity coeffi-

0021-3640/05/8110-0487$26.00 © 2005 Pleiades Publishing, Inc.

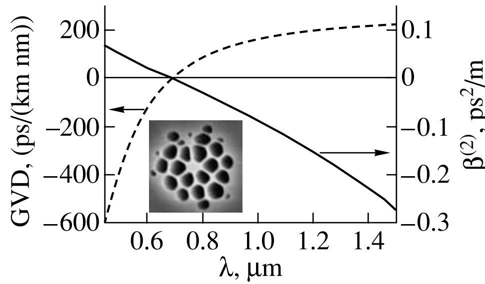

Fig. 1. (Dashed line) Group-velocity dispersion and (solid line) the coefficient $\beta^{(2)}$ vs. the radiation wavelength for the fundamental mode of the photonic-crystal fiber with a cross section structure shown in the inset.

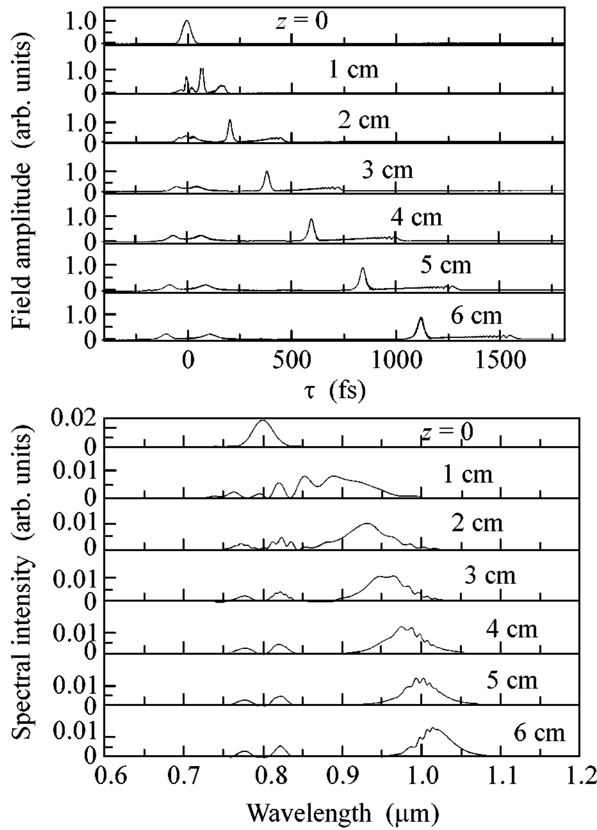

Fig. 2. Evolution of (a) the temporal envelope and (b) the spectral intensity of the laser pulse with an initial pulse width of 30 fs and an energy of 0.3 nJ in the photonic-crystal fiber.

cient, n_2 is the refractive index controlling the Kerr additive to the refractive index,

$$S_{\text{eff}} = \frac{\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(x,y)|^2 dxdy\right]^2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(x,y)|^4 dxdy} \quad (2)$$

is the effective area of the waveguide mode, $F(x, y)$ is the transverse field profile in the waveguide mode,

$$\Psi(z,\tau) = \int_{-\infty}^{\infty} R(t)|A(z,\tau-\eta)|^2 d\eta, \quad (3)$$

and $R(t)$ is the retarded nonlinear response. For fused silica, we take $n_2 \approx 3.2 \times 10^{-16}$ cm^2/W and employ the following model of the Raman response [3, 5]:

$$R(t) = (1-f_R)\delta(t) + f_R\Theta(t)\frac{\tau_1^2+\tau_2^2}{\tau_1\tau_2^2}e^{-\frac{t}{\tau_2}}\sin\left(\frac{t}{\tau_1}\right), \quad (4)$$

where $f_R = 0.18$ is the fraction of the Raman effect in the total nonlinear response, $\delta(t)$ is the delta function, $\Theta(t)$ is the Heaviside step function, and $\tau_1 = 12.5$ fs and $\tau_2 = 32$ fs are the typical times of the Raman response in fused silica.

To find $\beta^{(k)}$, we numerically solved the Maxwell equations for the transverse components of the electric and magnetic fields using the method of Hermite–Gaussian polynomial expansions [15]. Simulations were performed for a fused silica PCF with a core diameter of 1.6 μm and the structure of the cross section shown in the inset to Fig. 1. The frequency dependences of propagation constants β calculated with the use of this procedure were represented as polynomials

$$\beta(\omega) = \beta(\omega_0) + \sum_{k=1}^{6}\frac{\beta^{(k)}}{k!}(\omega-\omega_0)^k.$$

The level of tolerance for the errors of this polynomial fit in the range of wavelengths from 580 to 1220 nm was set equal to 0.1%. For the central frequency ω_0 corresponding to a wavelength of 800 nm, the required accuracy was achieved with the following set of expansion coefficients: $\beta^{(2)} \approx -0.0293$ ps^2/m, $\beta^{(3)} \approx 9.316 \times 10^{-5}$ ps^3/m, $\beta^{(4)} \approx -9.666 \times 10^{-8}$ ps^4/m, $\beta^{(5)} \approx 1.63 \times 10^{-10}$ ps^5/m, and $\beta^{(6)} \approx -3.07 \times 10^{-13}$ ps^6/m. The fact that $\beta^{(2)}$ is negative implies that the wavelength of 800 nm falls within the range of anomalous dispersion. Femtosecond pulses with such a central frequency can be therefore coupled into solitons in the considered PCF. The group-velocity dispersion (GVD) and the coefficient $\beta^{(2)}$ are shown in Fig. 1 as functions of the radiation wavelength for the fundamental mode of the PCF. The GVD vanishes, as can be seen from these plots, at the wavelength $\lambda_z \approx 690$ nm.

Figure 2 displays the spectral and temporal-envelope evolution of a pulse with an initial width of 30 fs propagating through a PCF with the above-specified dispersion coefficients $\beta^{(k)}$. Characteristic isolated features appearing in the spectrum and the temporal envelope of the field indicate formation of solitons. A finite response time of nonlinearity leads to the red shifting of solitonic features—a typical SSFS scenario. In the regime of anomalous dispersion, low-frequency com-

ponents of the field are slower than high-frequency components. As a result, Raman solitons become isolated from the nonsoliton part of the field in both time (Fig. 2a) and frequency (Fig. 2b) domains.

Numerical simulations indicate that shorter pulses noticeably enhance SSFS in PCFs. This result agrees well with the predictions of earlier studies [4] on the SSFS based on the analysis of the nonlinear Schrödinger equation. To provide qualitative physical insights into this tendency, we consider here the properties of the function $\Psi(z, \tau)$, which controls red shifting in Eq. (1), governing the evolution of an ultrashort pulse. As can be seen from Eq. (3), the function $\Psi(z, \tau)$ is given by a convolution of the temporal envelope of the field intensity and the retarded (Raman) response $R(t)$ of the nonlinear material. The Raman response $R(t)$ of fused silica is approximated [see Eq. (4)] with a damped oscillating function with a period $\tau_R = 2\pi\tau_1 \approx$ 80 fs. For pulse widths $\tau_0 \gg \tau_R$, the function $\Psi(z, \tau)$, as can be seen from Eq. (3), closely follows the temporal envelope of the field intensity (see Fig. 3a). The effect of the finite response time of nonlinearity is weak in this regime, and the red shifting of the Raman soliton is characterized by lower rates $d\nu_0/dz$.

For short laser pulses, $\tau_0 < \tau_R$, the function $\Psi(z, \tau)$, as can be seen from the results of calculations presented in Fig. 3b, can noticeably differ from the temporal profile of the field intensity. Deviation of $\Psi(z, \tau)$ from the temporal profile of the light intensity, as can be seen from calculations performed for pulse widths of 100 and 30 fs (Figs. 3a and 3b, respectively), increases for shorter input pulses, leading to higher SSFS rates $d\nu_0/dz$. Numerical analysis of soliton dynamics in a PCF shows that, as the initial pulse width is reduced from 50 to 30 fs, the soliton frequency shift in the PCF with a length of 7 cm increases for the above-specified dispersion profile by a factor of 1.5–2.

For the experimental studies, we employed an argon-laser-pumped self-starting Ti: sapphire laser with SESAM mirrors. Laser radiation was coupled into a PCF (see the inset to Fig. 1) with the use of standard micro-objectives. The spectra of radiation coming out of the fiber were measured by an Ando spectrum analyzer. The mode structure of the output radiation was analyzed with the use of a CCD camera.

The energy of the input laser pulse was chosen in such a way as to provide formation of the fundamental soliton. The energy of 30-fs laser pulses coupled into the fiber in our experiments was estimated as 0.3 nJ. In agreement with the results of our numerical simulations (Fig. 2), the initial stage of the pulse propagation involves rather complicated spectral and temporal dynamics of laser pulses. Nonlinear-optical spectral transformation of laser pulses under these conditions is dominated by self-phase modulation, stimulated Raman scattering, and parametric four-wave mixing [9, 16].

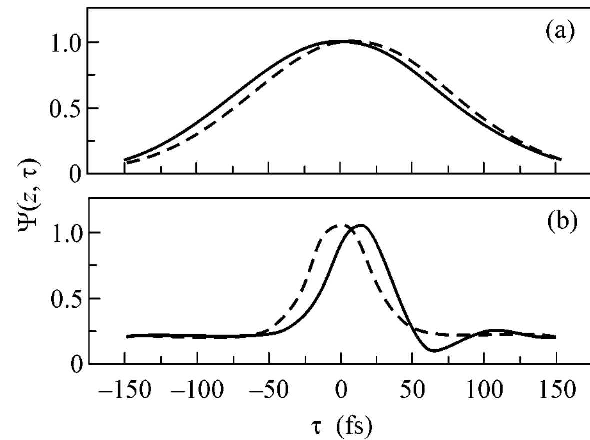

Fig. 3. Function $\Psi(z, \tau)$ representing a convolution of the temporal envelope of the field intensity and the function $R(t)$ governing the retarded (Raman) response of a nonlinear material for a light pulse with a duration of (a) 100 and (b) 30 fs.

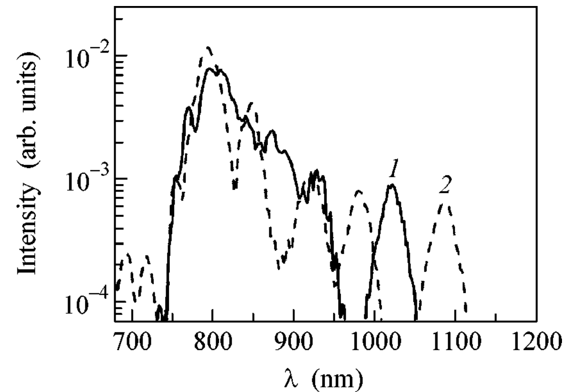

Fig. 4. Spectrum of a Ti: sapphire-laser pulse transmitted through a 7-cm photonic-crystal fiber. The initial pulse width is 30 fs. The input pulse energy is (*1*) 340 and (*2*) 540 pJ.

Within the propagation length of a few centimeters, we observe formation of red-shifted isolated frequency components (Fig. 4). Pulses with an initial duration of 30 fs transmitted through a 7-cm PCF section give rise to a red-shifted component with a central wavelength of about 1090 nm (line *2* in Fig. 4). The frequency shift of this component with respect to the central frequency of the pump field exceeds 100 THz. Higher input radiation energies in these experiments do not necessarily result in a noticeable increase in the maximum soliton frequency shift, which can be attributed to the excitation of high-order guided modes and formation of high-order solitons (cf. lines *1* and *2* in Fig. 4).

The experimental studies and numerical simulations presented in this paper reveal interesting features of the soliton self-frequency shift of laser pulses with initial

pulse widths less than the period of the Raman mode of the nonlinear material. We have shown that the rate of soliton self-frequency shift can be substantially increased by reducing the initial pulse width. A typical frequency shift of subnanojoule Ti: sapphire-laser pulses with an initial duration of about 30 fs transmitted through a PCF with a core diameter of about 1.6 μm and a length of about 7 cm exceeds 100 THz.

We are grateful to Yu.N. Kondrat'ev, V.S. Shevandin, K.V. Dukel'skii, and A.V. Khokhlov for the fabrication of fiber samples. This study was supported in part by the Council of the President of Russian Federation for Support of Young Russian Scientists and Leading Scientific Schools (project no. MD-42.2003.02), the Russian Foundation for Basic Research (project nos. 03-02-16929, 03-02-20002-BNTS, 04-02-39002-GFEN2004, and 04-02-81036-Bel2004), INTAS (project nos. 03-51-5037 and 03-51-5288), the U.S. Civilian Research and Development Foundation (award no. RP2-2558), the National Key Basic Research Special Foundation of China (grant no. 2003CB314904), the National Nature Science Foundation of China (grant no. 60278003), and the National High-Technology Program of China (grant no. 2003AA311010).

REFERENCES

1. E. M. Dianov, A. Ya. Karasik, P. V. Mamyshev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **41**, 242 (1985) [JETP Lett. **41**, 294 (1985)].
2. F. M. Mitschke and L. F. Mollenauer, Opt. Lett. **11**, 659 (1986).
3. G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. (Academic, Boston, 2001; Mir, Moscow, 1996).
4. J. P. Gordon, Opt. Lett. **11**, 662 (1986).
5. K. J. Blow and D. Wood, IEEE J. Quantum Electron. **25**, 2665 (1989).
6. C. Y. Teisset, N. Ishii, T. Fuji, *et al.*, in *Proceedings of Conference on Lasers and Electro-Optics (CLEO2005)* (Baltimore, ML, 2005).
7. P. St. J. Russell, Science **299**, 358 (2003).
8. J. C. Knight, Nature **424**, 847 (2003).
9. A. M. Zheltikov, *Optics of Microstructure Fibers* (Nauka, Moscow, 2004) [in Russian].
10. X. Liu, C. Xu, W. H. Knox, *et al.*, Opt. Lett. **26**, 358 (2001).
11. I. G. Corrmack, D. T. Reid, W. J. Wadsworth, *et al.*, Electron. Lett. **38**, 167 (2002).
12. D. T. Reid, I. G. Corrmack, W. J. Wadsworth, *et al.*, J. Mod. Opt. **49**, 757 (2002).
13. F. G. Omenetto, A. J. Taylor, M. D. Moores, *et al.*, Opt. Lett. **26**, 1158 (2001).
14. B. R. Washburn, S. E. Ralph, P. A. Lacourt, *et al.*, Electron. Lett. **37**, 1510 (2001).
15. T. M. Monro, D. J. Richardson, N. G. R. Broderick, and P. J. Bennet, J. Lightwave Technol. **18**, 50 (2000).
16. W. J. Wadsworth, A. Ortigosa-Blanch, J. C. Knight, *et al.*, J. Opt. Soc. Am. B **19**, 2148 (2002).

Translated by A. Zheltikov

JETP Letters, Vol. 81, No. 10, 2005, pp. 491–493. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 610–613.
Original Russian Text Copyright © 2005 by Touriyanskiĭ, Pirshin.

Tunneling of X-ray Photons through a Thin Film under Total Internal Reflection Conditions

A. G. Touriyanskiĭ and I. V. Pirshin

Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 119991 Russia

e-mail: tour@mail1.lebedev.ru

Received April 22, 2005

Tunneling of 0.154- and 0.139-nm x-ray photons through a thin film under total internal reflection conditions has been experimentally demonstrated. The $NiSi_2$ film 13 nm thick is deposited by magnetron sputtering on a polished Si substrate. A beam with an angular spread of 20″ is directed to the Si/$NiSi_2$ interface from the inside through the lateral surface of a sample. A peak associated with tunneling of photons from Si to air through the $NiSi_2$ film is observed at grazing angles of $\theta_1 > 0.4\theta_c$, where θ_c is the critical angle of total internal reflection at the Si/$NiSi_2$ interface. The integral intensity of tunneling peaks that is measured for various θ_1 angles agrees with the calculations. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 07.60.Hv, 07.85.Fv, 41.50.+h

The phenomenon of frustrated total internal reflection is well known in optics [1, 2]. The passage of radiation through a thin gap with an optically less dense medium under the formal conditions of total internal reflection is usually treated as tunneling of photons through an optical barrier. The effective wave-penetration depth z_e for grazing angles θ_1 smaller than the critical angle θ_c of total internal reflection depends on the incident radiation wavelength λ and the environment parameters. For this reason, frustrated total internal reflection in optical and IR spectral ranges is widely used at present to analyze the composition and structure of thin films and interfaces [3–5]. For hard x-ray radiation with $\lambda \sim 0.1$ nm, the refractive index decrement $\delta = 1 - n$ (n is the refractive index) for most materials lies in the range 10^{-6}–10^{-5}. For these δ values, total internal reflection can be observed only in a very narrow angular range. According to calculations and experimental data [6], the typical z_e values are 2–3 nm, which are less than one hundredth of the typical z_e values for the optical range.

Therefore, the direct observation of frustrated total internal reflection in the x-ray range is of fundamental and practical interest. In this work, we show that the effect can be detected under laboratory conditions and present the results of measurements of the angular distribution of the intensity of x-ray photon tunneling from silicon to air through the $NiSi_2$ thin film.

An optically polished Si single crystal disk was used as a substrate. The $NiSi_2$ film was deposited by the magnetotron sputtering. The film was partially oxidized due to interaction with air. According to x-ray reflectometry, the thickness of the nonoxidized $NiSi_2$ film was equal to 12.8 nm. A plate 1-cm wide was made from the central part of the disk by scribing from the side opposite to the film and subsequent splitting.

Measurements were carried out by means of an x-ray reflectometer whose x-ray optical scheme was described in [7]. A sharp focus x-ray tube with a copper anode was used as a radiation source. Figure 1 shows the position of a sample with respect to the incident radiation. A beam with an angular spread of 20″ was directed through the lateral surface of the sample. The edge of the lateral face of the sample was aligned with the reflectometer rotation axis. Thus, the flux density of the x-ray radiation incident on the lateral face remained constant under small rotations. The angular distribution of the intensity $I(\psi)$ of radiation passed through the sample at fixed grazing angles θ_1 of the primary beam was measured by scanning by a receiving slit that had a width s of 30 or 100 µm and was situated at a distance

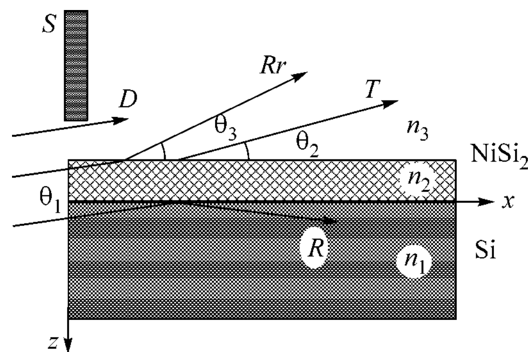

Fig. 1. Radiation passage scheme: (S) protection screen, (D) direct beam at grazing angle θ_1, (Rr) refracted beam at grazing angle θ_3, (R) reflected beam, and (T) beam tunneling through $NiSi_2$ at grazing angle θ_2.

0021-3640/05/8110-0491$26.00 © 2005 Pleiades Publishing, Inc.

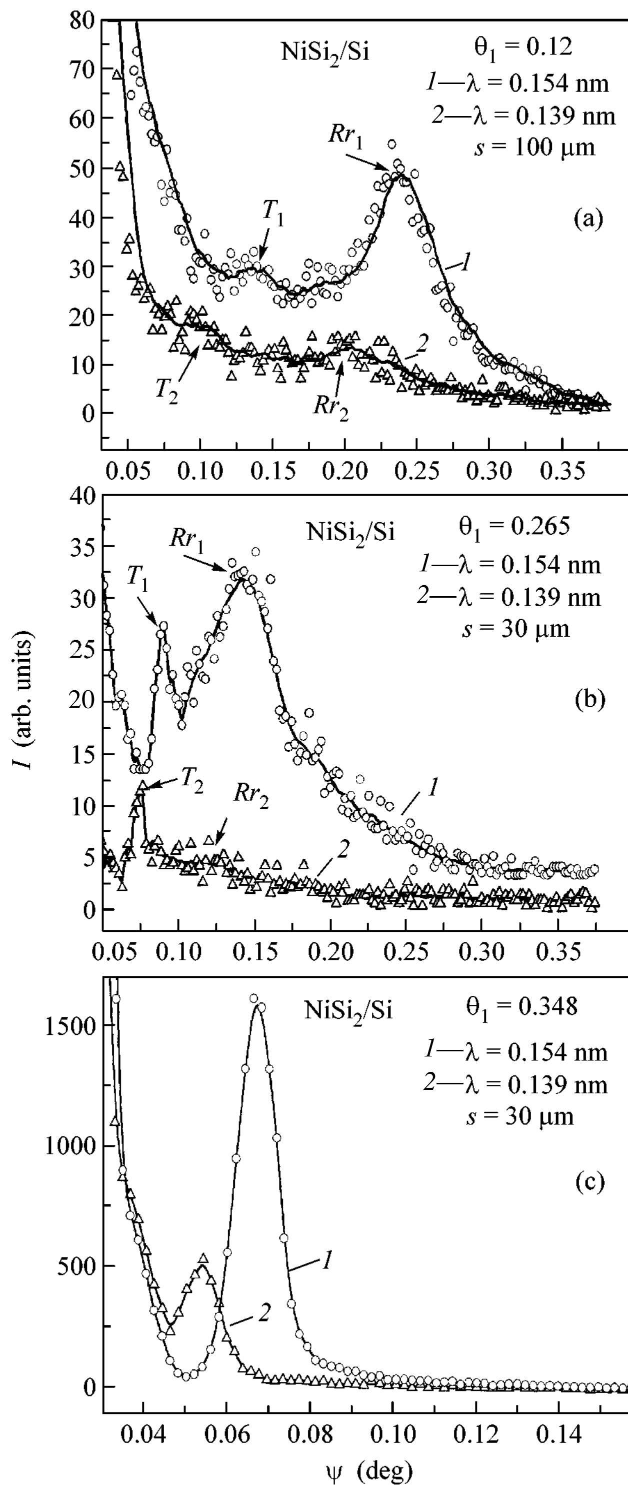

Fig. 2. Angular distributions of the intensity of the radiation passed through the sample for the grazing angles of the primary beam θ_1 = (a) 0.12°, (b) 0.265°, and (c) 0.348° and for (*1*) CuK_α 0.154-nm and (*2*) CuK_β 0.139-nm lines; panels (a) and (b) correspond to the tunneling regions and panel (c), to the reflection region.

of 225 mm from the reflectometer axis. The deviation angle ψ of radiation passed through the sample was measured from the primary beam axis. The spectral lines CuK_α (0.154 nm) and CuK_β (0.139 nm) were separated by semitransparent monochromators from radiation passed through the receiving slit. The angular scanning step and the direct radiation flux at the CuK_α line were equal to 3.6″ and 10^6 photons per second, respectively.

For the measurement geometry shown in Fig. 1, a certain part of the beam passed between the absorbing screen *S* and sample surface and was used as an angular reference. Another part of the beam passed through the lateral surface of the film and was refracted at the $NiSi_2$/air interface at the grazing angle θ_3. The main part of the beam entering the substrate was absorbed in Si. Photons whose path in Si is smaller than the mean free path l_s of x-ray photons predominantly reach the Si/$NiSi_2$ interface. The mean free path l_s for the CuK_α and CuK_β lines is equal to 71 and 99 μm, respectively. According to the table data [8], $\delta(\mathrm{Si}) = 7.57 \times 10^{-6}$ and $\delta(\mathrm{NiSi_2}) = 21.5 \times 10^{-6}$ for $\lambda = 0.154$ nm; i.e., the necessary condition of the total internal reflection $1 - \delta(\mathrm{Si}) > 1 - \delta(\mathrm{NiSi_2})$ is satisfied at the Si/$NiSi_2$ interface. Thus, this part of the x-ray flux can tunnel from Si to air through the $NiSi_2$ film.

The condition of the passage of the wave front through the lateral surface of the substrate without significant distortion can be written in the form

$$\Delta L(n_3 - n_1)/\lambda = \Delta L\delta(\mathrm{Si})/\lambda \ll 1, \qquad (1)$$

where ΔL is the mean height of ribs on the chip surface. Substituting the table δ value for Si at $\lambda = 0.154$ nm and, for definiteness, assuming that the left-hand side of the inequality is no more than 0.01, we obtain $\Delta L <$ 0.2 μm. The above condition is easily satisfied on the chips of the single crystal along the cleavage planes within sections that are much larger than the diameter of the first Fresnel zone.

Figure 2 shows the measured angular distributions of the intensity $I(\psi)$ of radiation passed through the sample at fixed grazing angles θ_1 of the primary beam. As was mentioned above, a sharp increase in the intensity for angles $\psi \longrightarrow 0$ is associated with the passage of a part of the direct beam over the sample edge. As follows from the geometry of the radiation, the broad peak on the right side of Figs. 2a and 2b arises due to the refraction of x-ray radiation passed through the end of the $NiSi_2$ film. The corresponding refraction maxima are denoted as Rr_1 and Rr_2. In the diagram measured on the CuK_β line, the Rr_2 peak is lower than the background signal, which is caused by the jump in the photoabsorption in Ni for $\lambda = 0.139$ nm. The tunneling flux can be measured beginning with the grazing angle θ_1 = 0.12° (Fig. 2a). The tunneling peaks for the two wavelengths in Figs. 2a and 2b are denoted as T_1 and T_2. For grazing angles $\theta_1 > \theta_c$ (Fig. 2c), only the intense peak of refracted radiation passing from silicon to air through the Si/film and film/air interfaces is observed.

Figure 3 shows (*1*) experimental and (*2*) theoretical dependences of the integral intensity Φ of x-ray radiation passed through the film at the grazing angle θ_1 for $\lambda = 0.154$ nm. The function Φ values are obtained by numerically integrating the $I(\psi)$ curve near the tunneling peak and subtracting the background signal. The angular range $\theta_1 < \theta_c \approx 0.31°$ is the tunneling region.

The angular dependence $\Phi(\theta_1)$ is calculated under the following assumptions. We assume that the angular spread of the primary beam is negligible and the linear flux density p_0 of x-ray photons incident on the end of the sample at the grazing angle θ_1 is constant. The z axis is perpendicular to the sample surface (see Fig. 1). In this case,

$$\Phi(\theta_1) = |t_1|^2 \int_0^{\infty} p_0 \exp\left(-\frac{\mu(z)z}{\sin\theta_1}\right) dz, \tag{2}$$

where p_0 is the linear density of the x-ray flux incident on the sample and μ is the linear coefficient of the absorption of radiation in the substrate for the given wavelength. The energy transmission coefficient $T = |t_1|^2$ of the film structure is calculated from the recurrence relations [9]

$$t_j = \frac{t_{j+1} t_j^{\mathrm{F}} e^{i\kappa_2 d_j}}{1 + r_{j+1} r_j^{\mathrm{F}} e^{2i\kappa_2 d_j}}, \tag{3}$$

$$r_j = \frac{r_j^{\mathrm{F}} r_{j+1} e^{2i\kappa_{j+1} d_j}}{1 + r_{j+1} r_j^{\mathrm{F}} e^{2i\kappa_{j+1} d_j}}. \tag{4}$$

Here, r_j and t_j are the complex amplitude coefficients of reflection and transmission of the structure, respectively; r_j^{F} and t_j^{F} are the Fresnel amplitude coefficients of reflection and transmission at the interfaces, respectively; $\kappa = 2\pi/\lambda(\varepsilon_j - \sin^2\varphi)^{1/2}$, where $\varphi = \pi/2 - \theta_1$ is the angle of incidence on the first interface and ε_j is the dielectric constant of the medium in the jth layer; and d_j is the layer thickness.

Comparison shows that the experimental and theoretical curves are in fairly good agreement. The result shown in Fig. 3 is obtained with the effective film thickness $D_{\mathrm{f}} = 18.5$ nm, which testifies to the presence of an oxide layer. This conclusion is also corroborated by x-ray reflectometry. For grazing angles $\theta_1 < 0.7\theta_c$, the tunneling photon flux is lower than 1% of the reflected flux for $\theta_1 \approx 1.1\theta_c$. The conditional boundary between the regions of tunneling and refraction is shown in Fig. 3 by the vertical dashed straight line. The presence of the extended transient section on the curve $\Phi(\theta_1)$ in the region $\theta_1 \approx \theta_c$ is explained by the effect of the imaginary part of ε.

In conclusion, we note that, for such materials as GaAs, Al_2O_3, Si, and fused silica that are most often used as substrates, the mean free path of photons for

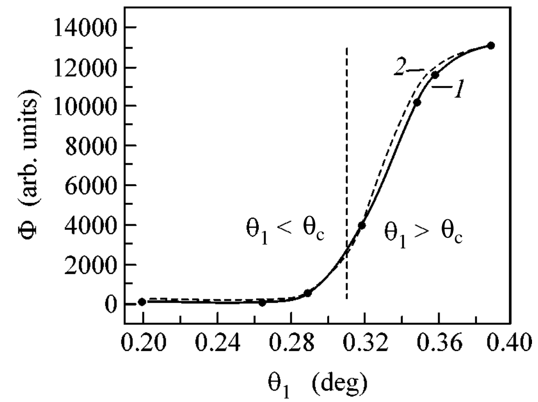

Fig. 3. (*1*) Experimental and (*2*) theoretical angular dependences of the integral flux intensity of x-ray photons passed through the outer interface: $\theta_1 < \theta_c$ is the tunneling region and $\theta_1 > \theta_c$ is the refraction region.

$\lambda = 0.154$ nm and, correspondingly, the effective width of the tunneling region are in the range 30–140 μm. These values are approximately two orders of magnitude smaller than the characteristic size of the irradiated region in x-ray reflectometry measurements. For this reason, the transmission measurement scheme proposed in this work can be used to analyze edge effects in the technology of producing film nanostructures and to control coatings on a nonflat surface with a curvature radius to 0.3–0.5 m.

We are grateful to N.N. Gerasimenko for samples placed at our disposal. This work was supported in part by the Russian Foundation for Basic Research, project no. 03-02-16976a.

REFERENCES

1. C. Hirlimann, B. Thomas, and D. Boosé, Europhys. Lett. **69**, 48 (2005).
2. J. J. Carey, J. Zawadzka, D. A. Jaroszynski, *et al.*, Phys. Rev. Lett. **84**, 1431 (2000).
3. D. C. Beddows, B. C. Griffiths, O. Samek, *et al.*, Appl. Opt. **42**, 6006 (2003).
4. A. A. Golubtsov, N. F. Pilipetskiĭ, A. N. Sudarkin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **43**, 219 (1986) [JETP Lett. **43**, 277 (1986)].
5. Yu. I. Malakhov, A. L. Kalabekov, and Yu. N. Korolev, Izmer. Tekh., No. 8, 40 (2002).
6. K. Sakurai, *X-Ray Spectrometry: Recent Technological Advances*, Ed. by K. Tsuji, J. Injuk, and R. Van Grieken (Wiley, Chichester, 2004).
7. A. G. Tur'yanskiĭ, A. V. Vinogradov, and I. V. Pirshin, Prib. Tekh. Éksp., No. 1, 105 (1999).
8. B. L. Henke, E. M. Gullikson, and J. C. Davis, At. Data Nucl. Data Tables **54**, 181 (1993).
9. P. H. Berning, in *Physics of Thin Films*, Ed. by G. Haas (Academic, New York, 1963), Vol. 1, p. 69 [P. H. Berning, *Theory and Calculations of Optical Thin Films* (Mir, Moscow, 1967), Vol. 1].

Translated by R. Tyapaev

JETP Letters, Vol. 81, No. 10, 2005, pp. 494–497. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 614–617.
Original Russian Text Copyright © 2005 by Stepikhova, Zhigunov, Shengurov, Timoshenko, Krasil'nikova, Chalkov, Svetlov, Shalygina, Kashkarov, Krasil'nik.

Population Inversion of the Energy Levels of Erbium Ions Induced by Excitation Transfer from the Semiconductor Matrix in Si–Ge Based Structures

M. V. Stepikhova[1], D. M. Zhigunov[2], V. G. Shengurov[3], V. Yu. Timoshenko[2], L. V. Krasil'nikova[1], V. Yu. Chalkov[3], S. P. Svetlov[3], O. A. Shalygina[2], P. K. Kashkarov[2], and Z. F. Krasil'nik[1]

[1] *Institute for Physics of Microstructures, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia*

[2] *Faculty of Physics, Moscow State University, Moscow, 119992 Russia*

e-mail: zhigunov@vega.phys.msu.ru

[3] *Research Physicotechnical Institute, Nizhni Novgorod State University, Nizhni Novgorod, 603950 Russia*

Received April 28, 2005

Population inversion of the energy levels of Er^{3+} ions in $Si/Si_{1-x}Ge_x$:Er/Si ($x = 0.28$) structures has been achieved due to electron excitation transfer from the semiconductor matrix. An analysis of the photoluminescence kinetics at a wavelength of 1.54 μm shows that up to 80% of the Er^{3+} ions are converted into excited states. This effect, together with the high photoluminescence intensity observed in the structures studied, shows good prospects for obtaining lasers compatible with planar silicon technology. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 42.70.Hj, 78.47.+p, 78.55.–m

A currently important task for the modern physics of semiconductors and semiconductor technology is the search for new silicon-based materials possessing high efficiency of luminescence under optical or electric excitation conditions. As is known, crystalline silicon (c-Si), which is a base material in modern semiconductor electronics, is characterized by an indirect bandgap and, hence, by a low efficiency of the interband radiative recombination. There are several possible ways to increase the luminescence yield, which employ, in particular, the features of luminescence in low-dimensional silicon structures and the introduction of luminescence activators, such as rare earth ions, into the matrix structure [1, 2]. Among such luminescence activators, the most interesting is erbium ion (Er^{3+}), which exhibits a radiative electron transition ($^4I_{13/2} \longrightarrow {}^4I_{15/2}$) in the inner 4*f* shell. This is accompanied by the emission of light with a wavelength of 1.54 μm, which corresponds to minimum losses in quartz fiber-optic communication lines [1]. An important advantage of erbium-doped silicon structures is the possibility of providing effective nonresonant excitation of Er^{3+} ions by light in a broad spectral range, as well as by electric current, which is related to the mechanisms of indirect excitation of rare earth ions by means of electron excitation energy transfer from the semiconductor matrix [2–4].

A necessary condition for the optical amplification in erbium-doped silicon structures is the creation of a population inversion of the energy levels of Er^{3+} ions. Another important requirement is the effective localization of erbium luminescence in the active layer, which implies the formation of waveguide structures. One promising variant of a system of this kind is offered by Si/SiGe:Er/Si heterostructures with an active waveguide layer of erbium-doped silicon–germanium solid solution (SiGe:Er) [5, 6]. This paper presents the results of our investigation into the photoluminescence (PL) of Er^{3+} ions in $Si/Si_{1-x}Ge_x$:Er/Si heterostructures at high levels of optical pumping. We demonstrate the possibility of obtaining population inversion of the energy levels of rare earth ions in this system due to the electron excitation energy transfer from the semiconductor matrix.

The $Si/Si_{1-x}Ge_x$:Er/Si heterostructures were grown by sublimation MBE [6] from the gas phase on (100)-oriented c-Si substrates of the KEF-4.5 grade. The substrate temperature during deposition was 500°C. The growth of an $Si_{1-x}Ge_x$ solid solution was preceded by deposition of a buffer c-Si layer ~0.2 μm thick. The flow of Er atoms was generated by evaporating a sample of polycrystalline silicon doped with this impurity (for more details, see [6]). In most samples, the thickness of the active $Si_{1-x}Ge_x$:Er layer was about 1 μm. The elemental composition of the as-grown structures was studied by secondary ion mass spectrometry (SIMS). The SIMS profiles showed a virtually homogeneous (free of segregation effects) distribution of Er over the $Si_{1-x}Ge_x$ solid solution layer thickness with an impurity concentration of $N_{Er} \sim (2–3) \times 10^{18}$ cm^{-3} at a germanium content corresponding to $x = 0.28$.

The PL was excited by quasi-continuous pulsed radiation of a copper vapor laser (working wavelengths,

511 and 578 nm; pulse duration, 20 ns; repetition rate, 12 kHz). The laser spot size on the sample was approximately 2 × 2 mm, and the average excitation intensity I_{ex} did not exceed 5 W/cm^2. The exciting radiation was modulated by means of a mechanical chopper operating at a frequency of ~30 Hz. The PL measurements were performed at 10 K in an ARS DE-204N closed-loop helium cryostat. The sample temperature increase (monitored by a temperature sensor fixed on a sample holder) at the maximum optical pumping level did not exceed 2–3 K. The PL signal was measured using an InGaAs photodiode with a time resolution of ~0.2 ms.

The samples exhibited the characteristic PL due to radiative electron transitions in erbium with a wavelength of 1.54 μm. The emission intensity was comparable with the response of c-Si:Er samples possessing a low-temperature inner quantum efficiency of 10% [5, 7]. Figure 1 shows the typical PL spectra of Si/Si$_{1-x}$Ge$_x$:Er/Si ($x = 0.28$) samples measured at various excitation intensities I_{ex}. The spectra display a series of lines due to electron transitions in Er^{3+} ions between the $^4I_{13/2}$ and $^4I_{15/2}$ levels split in the crystal field. The presence of both relatively narrow peaks (with a full width at half height FWHM ~ 7 nm) and broad overlapping bands in the spectra is indicative of the existence of emitting centers of several types. The absence of noticeable changes in the shape of the PL spectra measured at maximum I_{ex} values is evidence of insignificant laser-induced heating of the samples.

The inset in Fig. 1 shows a plot of the PL intensity versus I_{ex} and its approximation by the function

$$I_{PL}(I_{ex}) = \frac{abI_{ex}}{1 + bI_{ex} + c\sqrt{I_{ex}}}, \quad (1)$$

where a, b, and c are constant coefficients. The shape of this function is indicative of an indirect mechanism of Er^{3+} ion excitation via the electron subsystem of the semiconductor matrix in the presence of competitive processes such as the nonradiative Auger recombination [8, 9]. The tendency to attaining a stationary PL intensity level, which is described by formula (1), is related to a finite number of emitting centers (Er^{3+} ions) in the Si/SiGe:Er/Si structures studied.

In order to determine the fraction of excited ions, we have studied the PL intensity kinetics at a wavelength of 1.54 μm. Figure 2 shows the shape of the exciting radiation pulse and the kinetic curves measured for two values of I_{ex}. An analysis of the PL decay kinetics (see the inset in Fig. 2) showed that these curves could be well described by a sum of two exponents:

$$I_{PL}(t) = A_1\exp(-t/\tau_1) + A_2\exp(-t/\tau_2), \quad (2)$$

where A_1 and A_2 are the coefficients determining the contributions of processes with the time constants $\tau_1 \approx 0.9$ ms and $\tau_2 \approx 2.5$ ms, respectively. At low excitation levels, the weighting coefficients are related as $A_2 \approx 0.2A_1$. The presence of two components in the PL kinetics probably reflects the presence of optically active centers of the two types that contain Er^{3+} ions and contribute to the emission at 1.54 μm. An analogous behavior was previously reported in [10]. The observed ratio of the amplitudes of the components indicates that a dominating role (~80% in terms of the concentration) is played by the "fast" emitting centers possessing shorter lifetimes and, hence, more tightly bound to the semi-

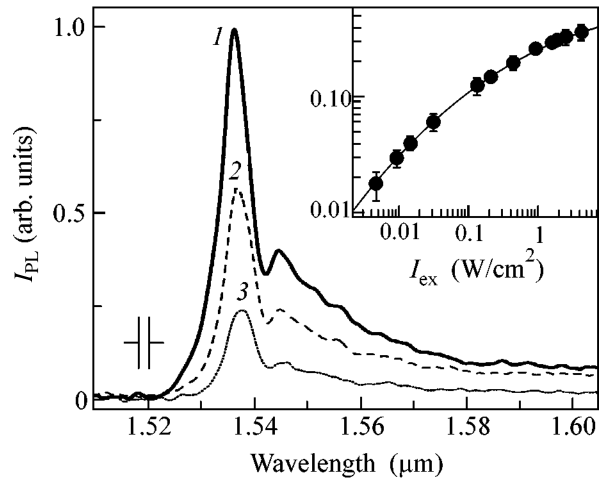

Fig. 1. Low-temperature photoluminescence spectra of the Si/Si$_{0.72}$Ge$_{0.28}$:Er/Si structure measured with the excitation intensity I_{ex} = (*1*) 1.65, (*2*) 0.2, and (*3*) 0.003 W/cm^2. The inset shows (points) the experimental photoluminescence signal intensities at 1.54 μm versus I_{ex} and (line) the approximation by formula (1).

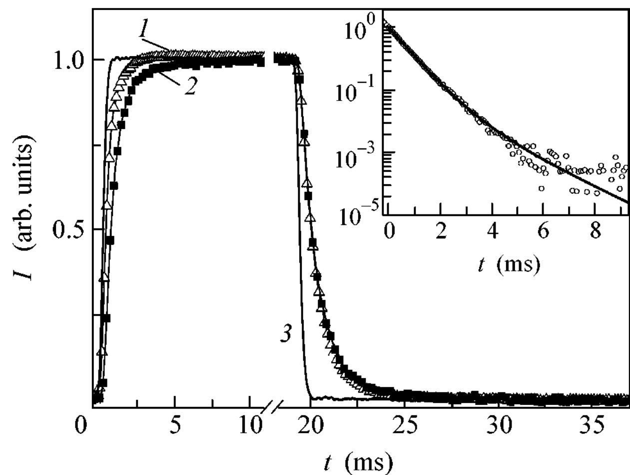

Fig. 2. Kinetics of the photoluminescence signal at 1.54 μm measured for I_{ex} = (*1*) 0.2 and (*2*) 0.005 W/cm^2 in comparison to (*3*) the exciting laser pulse profile. The inset shows (points) the experimental photoluminescence signal decay kinetics and (line) its approximation by formula (2).

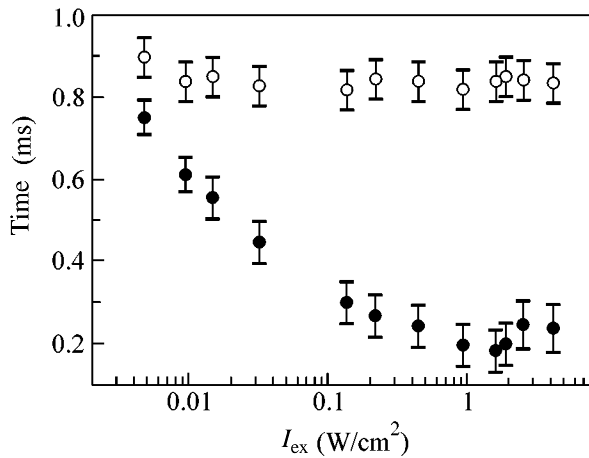

Fig. 3. Photoluminescence signal (○) decay and (●) rise times versus the excitation intensity for the fast component of photoluminescence from Er^{3+} centers.

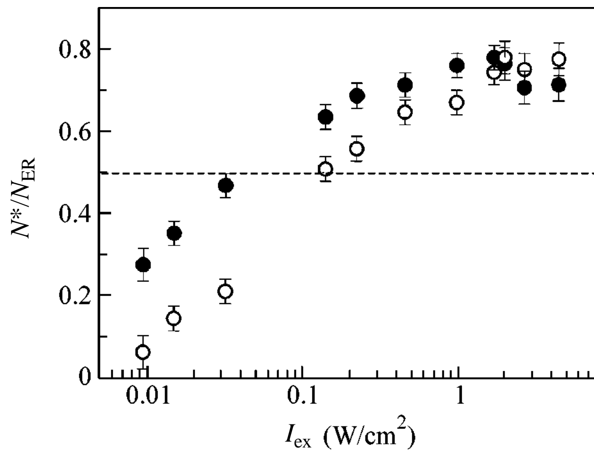

Fig. 4. Ratio of the concentration of Er^{3+} ions in the excited state $^4I_{13/2}$ to their concentration in the ground state $^4I_{15/2}$ calculated using formula (4) for (●) fast and (○) slow emitting centers. The dashed line indicates the level of N^*/N_{Er} = 0.5, above which the population inversion of the excited erbium-containing centers is achieved.

conductor matrix. The "slow" emitting centers may represent Er^{3+} ions in dielectric inclusions such as SiO_2 precipitates in silicon [1].

It was established that the PL rise kinetics for the erbium-containing emission centers could be well described by the approximate expression $1 - I_{PL}(t)$, where $I_{PL}(t)$ is the function given by formula (2). For large I_{ex}, we observed a sharp decrease in the PL rise time. An analysis of the signal intensity rise and decay for the PL from Er^{3+} ions as dependent on the excitation intensity I_{ex} allows the relative concentration of excited ions to be estimated [11]. Indeed, a kinetic equation for the excitation of Er^{3+} ions can be written as

$$\frac{dN^*}{dt} = g(N_{Er} - N^*) - \frac{N^*}{\tau^{decay}}, \quad (3)$$

where N^* is the concentration of ions in the excited state, τ^{decay} is the PL signal decay time (i.e., the lifetime of the excited state), and g is the rate of ion excitation by means of energy transfer from the semiconductor matrix. The latter characteristic, in the general case, depends on a number of parameters, in particular, on the erbium-containing center type, the sample temperature, and I_{ex}.

Equation (3) correctly describes the kinetics of variation of the number of emitting centers of various types in the absence of their interaction with each other. For the dopant concentrations studied, this assumption is likely to be valid. Then, the stationary values of the relative content of excited Er^{3+} ions is given by the relation

$$\frac{N^*}{N_{Er}} = 1 - \frac{\tau^{rise}}{\tau^{decay}}, \quad (4)$$

where $\tau^{rise} = (g + 1/\tau^{decay})^{-1}$ is the PL signal rise time (dependent on I_{ex}).

Figure 3 shows experimental plots of the PL signal rise and decay time versus optical excitation intensity I_{ex} for the fast PL component. As can be seen, the growth of I_{ex} is accompanied by a sharp decrease in τ_1^{rise}, while the τ_1^{decay} value is virtually independent of the excitation level: $\tau_1^{decay} \approx \tau_1$ = const. It should be noted that qualitatively similar dependences of the PL signal rise and decay times on I_{ex} were observed for the slow PL signal component, which is in good agreement with the above simple two-exponent kinetic model.

An analysis of the observed PL kinetics in terms of Eq. (4) gave the relative concentrations of Er^{3+} ions in the excited state for the optically active centers of two types with the life times τ_1 and τ_2. Figure 4 presents the plots of N^*/N_{Er} versus the optical excitation intensity. At relatively low I_{ex} values, the fraction of excited Er^{3+} ions entering into the fast centers is several times greater than the content of slow centers. This result is consistent with the above assumption concerning a more effective interaction between the semiconductor matrix and the erbium-containing centers possessing the shorter lifetime τ_1. For $I_{ex} > 0.05$ W/cm^2, the fraction of these centers is $N^*/N_{Er} > 1/2$, which implies the presence of a population inversion of the corresponding energy levels. For the slow centers, the state of a population inversion is achieved at a two to three times higher excitation level. At $I_{ex} \approx 1$ W/cm^2, up to about 80% of the centers of both types occur in the excited state. Further growth of I_{ex} led to some decrease in the N^*/N_{Er} ratio for the fast centers, which was probably

caused by heating of the matrix under the conditions of intense optical pumping. At the same time, the slow centers were less sensitive with respect to the thermal effects of laser radiation, which is consistent with the above assumption concerning a less effective interaction (or stronger isolation) of such centers with the semiconductor matrix.

It should be noted that the results of our investigation does not allow us to judge unambiguously on the presence or absence of the phenomena of stimulated emission and optical amplification in the $Si/Si_{1-x}Ge_x$:Er/Si structures studied. These questions can only be elucidated in special experiments using some standard techniques, such as PL measurements for a variable size of the excitation region in the sample or double-beam diagnostic methods.

To summarize, we have demonstrated that population inversion of the energy levels in both fast and slow centers containing Er^{3+} ions can be created in $Si/Si_{1-x}Ge_x$:Er/Si ($x = 0.28$) structures due to effective electron excitation energy transfer from the semiconductor matrix. The obtained results show good prospects for realization of the optical amplification regime at a wavelength of 1.54 μm in such structures provided that the necessary conditions of optical pumping and radiation localization in the waveguide channel are satisfied and are of interest for the development of silicon optoelectronics.

This study was supported by the Russian Foundation for Basic Research (project nos. 03-02-16647, 04-02-17120, 05-02-16735, 04-02-08240, and 04-02-08083) and INTAS (grant no. 03-51-6486) and was conducted at the Center of Shared Equipment, Faculty of Physics, Moscow State University.

REFERENCES

1. A. Polman, J. Appl. Phys. **82**, 1 (1997).
2. S. Coffa, G. Franzo, and F. Priolo, MRS Bull. **4**, 25 (1998).
3. W. Fuhs, I. Ulber, G. Weiser, *et al.*, Phys. Rev. B **56**, 9545 (1997).
4. P. K. Kashkarov, M. G. Lisachenko, O. A. Shalygina, *et al.*, Zh. Éksp. Teor. Fiz. **124**, 1255 (2003) [JETP **97**, 1123 (2003)].
5. Z. F. Krasilnik, V. Ya. Aleshkin, B. A. Andreev, *et al.*, in *Towards the First Silicon Laser*, Ed. by L. Pavesi, S. Gaponenko, and L. Dal Negro (Kluwer Academic, Dordrecht, 2003), p. 445.
6. S. P. Svetlov, V. G. Shengurov, V. Yu. Chalkov, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **65**, 203 (2001).
7. B. A. Andreev, T. Gregorkievich, Z. F. Krasil'nik, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **67**, 273 (2003).
8. R. Serna, J. H. Shin, M. Lohmeier, *et al.*, J. Appl. Phys. **79**, 2658 (1996).
9. J. Palm, F. Gan, B. Zheng, *et al.*, Phys. Rev. B **54**, 17603 (1996).
10. S. Coffa, G. Franzo, F. Priolo, *et al.*, Phys. Rev. B **49**, 16313 (1994).
11. V. Yu. Timoshenko, O. A. Shalygina, M. G. Lisachenko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **47**, 116 (2005) [Phys. Solid State **47**, 121 (2005)].

Translated by P. Pozdeev

JETP Letters, Vol. 81, No. 10, 2005, pp. 498–502. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 618–622.
Original English Text Copyright © 2005 by Brener, Metzner.

Semiclassical Theory of Electron Drag in Strong Magnetic Fields¶

S. Brener and W. Metzner
Max-Planck-Institut für Festkörperforschung, D-70569 Stuttgart, Germany
Received March 21, 2005

We present a semiclassical theory for electron drag between two parallel two-dimensional electron systems in a strong magnetic field, which provides a transparent picture of the most salient qualitative features of anomalous drag phenomena observed in recent experiments, especially the striking sign reversal of drag at mismatched densities. The sign of the drag is determined by the curvature of the effective dispersion relation obeyed by the drift motion of the electrons in a smooth disorder potential. Localization plays a role in explaining the activated low temperature behavior but is not crucial for anomalous drag per se. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 73.21.–b, 73.43.–f, 73.63.–b

Electron drag in double-layer two-dimensional electron systems has been established as a valuable probe of the electronic state within each layer and also of interlayer interactions [1]. In a drag experiment, a current is driven through one of the layers, which, via interlayer scattering, produces a drag voltage in the other layer. Usually, the drag is positive in the sense that the drag and drive currents have the same direction, which leads to a compensating voltage opposite to the drive current. In the absence of a magnetic field, drag phenomena in 2D electron systems are well understood [2]. In that case, the drag signal can be expressed in terms of the density response functions of each single layer, as long as the interlayer coupling is weak [3].

Drag experiments have also been performed in the presence of strong magnetic fields, where the formation of Landau levels plays a crucial role. Pronounced minima in the drag signal were observed at odd integer filling factors for magnetic fields far below the regime where spin splitting of Landau levels is seen in the intralayer resistivity [4, 5]. Completely unexpected was the discovery of a negative drag signal in the case of a suitably chosen density mismatch between the two layers [6, 7]. Previous theoretical descriptions [2, 8] would predict positive drag at any filling factor. In a more recent work, a mechanism for a sign change of the drag in strong magnetic fields was found [9], which, however, yielded negative drag for equal densities in the two layers, unlike in the experimental situation.

New hints and constraints were observed for a theory resulting from a recent detailed experimental analysis of the temperature dependence of the drag resistivity by Muraki *et al.* [10]. At high temperatures, $k_BT \gg \hbar\omega_c$, where ω_c is the cyclotron frequency, the drag resistivity ρ_D follows the conventional T^2 behavior. It decreases with decreasing T down to scales well below $\hbar\omega_c$, but, then, in the data shown for $\nu = 2n + 3/2$ with $n = 2, 3, 4, 5$, it rises again to form a pronounced peak at temperatures one order of magnitude below $\hbar\omega_c$. At still lower T, the drag resistivity decreases rapidly, which is consistent with the exponentially activated behavior, $\rho_D \propto e^{-\Delta/k_BT}$. The ν dependence of the activation energy Δ oscillates with minima near half-integer filling factors. The longitudinal intralayer resistivity ρ_{xx} also exhibits activated behavior at the lowest temperature with an activation energy matching with Δ for those ν where both gaps could be extracted from the data. For a density mismatch $\delta\nu = 1$ between the two layers, a low temperature peak with the same shape as for equal densities but with opposite (that is, negative) signs was observed.

Muraki *et al.* [10] interpreted their data in terms of particle-hole asymmetry and disorder induced localization. Landau levels are broadened by disorder, thus, leading to a band of energies for each level. Anomalous drag is observed for parameters where the Landau level broadening is much smaller than $\hbar\omega_c$. The longitudinal resistivity ρ_{xx} of each layer, as well as the drag resistivity ρ_D, is due to electrons in extended states of the highest occupied Landau level. If the states with energies near the Fermi level ϵ_F are localized, charge carriers in extended states can be created only by thermal activation or scattering across the energy gap between ϵ_F and the nearest extended state energy level. This explains the observation of the activated behavior of ρ_{xx} and ρ_D at low T. Muraki *et al.* [10] also realized that odd integer filling factors lead to a particle-hole symmetric occupation of the partially filled Landau level. Since particle-hole asymmetry is known to be crucial for drag in the absence of a magnetic field [11], it is indeed natural to relate the minima in ρ_D observed at an odd integer ν to this symmetry. Judging from the experiments, a necessary requirement for negative drag seems to be that the

¶ This article was submitted by the authors in English.

partially occupied Landau level in one layer be more than half filled and the other less than half filled. This was already noticed by Feng *et al.* [6] and led them to speculate that electrons in a more than half-filled Landau level should acquire a holelike dispersion relation due to disorder.

A comprehensive transport theory for the intralayer and drag resistivity in a strong magnetic field, which captures localization and extended states, is not available. In this work, we present a simple semiclassical picture that explains why Landau quantized electrons moving in a smooth disorder potential behave effectively as band electrons with an electron-like dispersion in the lower and a holelike dispersion in the upper half of a disorder broadened Landau level and how this leads to the observed anomalous drag.

Let us first consider the relevant length and energy scales in the double-layer system. The distance between the layers varies from 30 to 120 nm. The disorder is due to remote donors, which leads to a smooth disorder potential with a correlation length ξ of order 50–100 nm. The Landau level broadening, which is related to the amplitude of the disorder potential, has been estimated from low-field Shubnikov-de-Haas oscillations to lie in the range 0.5–2 K [6, 12]. The magnetic fields at which drag minima at odd integer filling and/or negative drag at mismatched densities are observed vary over a relatively wide range from 0.1 T to 1 T for the cleanest samples and, up to 5 T, for samples with a slightly reduced mobility. This corresponds to magnetic lengths l_B between 15 and 80 nm. Hence, ξ and l_B are generally of the same order of magnitude, which makes a quantitative theoretical analysis rather difficult. Since the anomalous drag phenomena are observed over a wide range of fields, one may hope that qualitative insight can also be gained by analyzing limiting cases. For $\xi \ll l_B$, a treatment of disorder within the self-consistent Born approximation is possible [13]. This route was taken most recently by Gornyi *et al.* [14], who obtained negative contributions to the drag resistivity at mismatched densities in agreement with experiment. Localization is not captured by the Born approximation; consequently, the resistivities obey power-laws rather than activated behavior for $T \longrightarrow 0$. For stronger magnetic fields, on the other hand, localization will become important, and a semiclassical approximation is a better starting point, which is the route we take here. Applying a criterion derived by Fogler *et al.* [15], we estimate that (classical) localization may set in already at 0.1 T.

We now discuss the semiclassical picture of electron states and drag on a qualitative level, ignoring spin splitting for simplicity. If the disorder potential varies smoothly, electron states lie essentially on contours of equal energy. These contours form closed loops, which correspond to localized states, except at a single energy level ϵ_0 in the center of each Landau level, for which there is a percolating contour through the whole system [16]. If the Landau level broadening induced by the disorder potential and also k_BT is much smaller than $\hbar\omega_c$, as is the case in the anomalous drag regime, all the Landau levels except one are either fully occupied in the bulk of the whole sample or completely empty. At zero temperature and for $\epsilon_F < \epsilon_0$, the sample then consists of islands where the highest (partially) occupied Landau level is locally full, while it is locally empty in the rest of the sample; for $\epsilon_F > \epsilon_0$, the empty regions form islands. In reality, the percolating contour at the center of the Landau level is broadened to a percolation region for two reasons. First, electrons near the percolating contour can hop across saddle points from one closed loop to another by using, for a moment, some of their cyclotron energy [15]. Second, electrons near the center of the Landau level can screen the disorder potential, such that the percolating equipotential line at ϵ_0 broadens to form equipotential terraces [17]. Hence, there is a region around the percolating path where the states are extended.

The percolating region forms a two-dimensional random network consisting of links and crossing points [18]. For simplicity we assume that the links are straight lines. The region near the link (except near the end points) can be parameterized by a cartesian coordinate system with a variable x following the equipotential lines parallel to the link and y for the transverse direction. The disorder potential varies essentially only in the transverse direction and can thus be represented by the function $U(y)$. In the Landau gauge $\mathbf{A} = (-By, 0)$, the Hamiltonian for electrons on the link is then translation invariant in the x direction and the Landau states ($\hbar = 1$ from now on)

$$\psi_{nk}(x, y) = C_n e^{-(y - l_B^2 k)^2/2l_B^2} H_n[(y - l_B^2 k)/l_B] e^{ikx} \quad (1)$$

are accurate solutions of the stationary Schrödinger equation. Here, H_n is the nth Hermite polynomial and C_n a normalization constant. The corresponding energy is simply

$$\epsilon_{nk} = (n + 1/2)\omega_c + U(l_B^2 k). \quad (2)$$

In the following, we drop the Landau level index n, because only the highest occupied level is relevant. The quantum number k is the momentum associated with the translation invariance in x direction. It is proportional to a transverse shift $y_0 = l_B^2 k$ of the wave function. The potential U lifts the degeneracy of the Landau levels and makes the energies depend on the momentum along the link. The group velocity

$$v_k = \frac{d\epsilon_k}{dk} = l_B^2 U'(y_0) \quad (3)$$

corresponds to the classical drift velocity of an electron in crossed electric and magnetic fields. In our simplified straight link approximation, $U'(y_0)$ does not depend on x. In general, it will depend slowly on the longitudi-

nal coordinate, but, near the percolating paths away from the crossing points, it has a fixed sign; hence, the group velocity has a fixed direction; that is, the motion on the links is chiral.

The drag between two parallel layers is dominated by electrons in the percolating region and corresponds to states near the center of the Landau level, since electrons in deeper localized states are not easily dragged along. In the network picture, macroscopic drag arises as a sum of the contributions from interlayer scattering processes between electrons in different links. If no current is imposed in the drive layer, both layers are in thermal equilibrium and the currents on the various links cancel each other on average. Now, assume that a small finite current is switched on in the drive layer such that the electrons move predominantly in the direction of the positive x axis (the electric current moving in the opposite direction). This means that the current on the links oriented in the positive x direction is typically larger than the current on the links oriented in the negative x direction. Interlayer scattering processes lead to momentum transfers between electrons in the drive and drag layer. The preferred direction of the momentum transfers is such that the scattering processes tend to reduce the current in the drive layer; that is, the interlayer interaction leads to friction. Now, the crucial point is that electrons moving in the disorder potential are not necessarily accelerated by gaining momentum in the direction of their motion or slowed down by loosing momentum (as for free electrons). The fastest electrons are those near the center of the Landau level: they have the highest group velocity on the links of the percolation network and they most easily get across the saddle points. Electrons in states below the Landau level center are thus accelerated by gaining some extra momentum in the direction of their motion, but electrons with an energy above ϵ_0 are pushed to still higher energy by adding momentum and are thus slowed down. To understand how negative drag arises, consider the situation with the highest occupied Landau level of the drive layer less than half filled and the one in the drag layer more than half filled. In that case, the electrons in the drag layer receive momentum transfers with a predominantly positive x component. As a consequence, electrons near the Fermi level of the drag layer are mostly slowed down if they move on links toward the positive x direction and accelerated if they move on links in the negative x direction, which is the opposite of what free electrons would do. If no net current is allowed to flow in the drag layer, an electric field is built up in the negative x direction, which compensates for the effective force generated by the scattering processes with the drive layer electrons. The drag signal is thus negative.

To substantiate the above qualitative picture, let us analyze the drag between two links using semiclassical transport theory. We first discuss parallel links in some detail, then, briefly discuss the more general case of a finite angle between the links. For a fixed Landau level index, the electronic states in each link are fully labeled by their momentum k. The (nonequilibrium) occupation of states in the links is described by a distribution function $f_\alpha(k)$, where $\alpha = 1, 2$ labels the two layers. We consider the experimental standard setup where a small current flows through the drive layer ($\alpha = 2$) generating a compensating drag voltage in the drag layer ($\alpha = 1$). No net current is allowed to flow in the drag layer. Under these conditions, the linear response of the drag layer to the drive current is determined by the linearized Boltzmann equation

$$\dot{k}_1 \cdot \frac{\partial f_1^0}{\partial k_1} = \left[\frac{\partial f_1}{\partial t}\right]_{\text{coll}} \tag{4}$$

with $\dot{k}_1 = -eE_1$, where E_1 is the electric field leading to the drag voltage. The interlayer collision term is given by

$$\begin{aligned} \left[\frac{\partial f_1}{\partial t}\right]_{\text{coll}}^{12} &= -\int\frac{dk_2}{2\pi}\int\frac{dk_1'}{2\pi}W^{12}_{k_1, k_2;\, k_1', k_2'} \\ &\times[\psi_1(k_1) + \psi_2(k) - \psi_1(k_1') - \psi_2(k_2')] \\ &\times f_1^0(k_1)f_2^0(k_2)[1 - f_1^0(k_1')][1 - f_2^0(k_2')] \\ &\times \delta(\epsilon_{k_1} + \epsilon_{k_2} - \epsilon_{k_1'} - \epsilon_{k_2'}), \end{aligned} \tag{5}$$

where $W^{12}_{k_1, k_2;\, k_1', k_2'}$ is the rate for a single interlayer scattering event $k_1 \longrightarrow k_1'$, $k_2 \longrightarrow k_2'$; and the deviation (from equilibrium) functions ψ_α are defined as usual by $f - f^0 = f^0(1 - f^0)\psi$. There is also an intralayer scattering term with a similar structure. The distribution function in the drive layer is obtained from the single layer Boltzmann equation as $\psi_2(k) = -\frac{\pi}{e v_{F2} T} j_2 v_k$, where j_2 is the drive current and v_{F2} the equilibrium Fermi velocity. For weak interlayer coupling, the deviation function ψ_1 can be neglected in the interlayer collision term, because it would yield a contribution of the order $(W^{12})^2$. To determine the relation between the drive current j_2 and the drag field E_1, we multiply the Boltzmann equation by v_{k_1} and integrate over k_1. The left hand side yields $-e\int\frac{dk_1}{2\pi}f_1^{0'}(\epsilon_{k_1})v_{k_1}^2 E_1$, which tends to $\frac{e}{2\pi} v_{F1}E_1$ at low temperatures, where v_{F1} is the velocity at the Fermi level of the drive layer. Using the antisymmetry of the integrand of Eq. (5) under the exchange of k_1 and k_1', the right hand side can be written as

$$-\frac{\pi j_2}{e v_{F2} T}\int\frac{dk_1}{2\pi}\int\frac{dk_2}{2\pi}\int\frac{dq}{2\pi}W^{12}_{k_1, k_2;\, k_1 + q, k_2 - q}$$

$$\times (v_{k_2} - v_{k_2-q})(v_{k_1} - v_{k_1+q})$$
$$\times \frac{[f_1^0(\epsilon_{k_1}) - f_1^0(\epsilon_{k_1+q})][f_2^0(\epsilon_{k_2}) - f_2^0(\epsilon_{k_2-q})]}{4\sinh^2[(\epsilon_{k_1+q} - \epsilon_{k_1})/2T]} \quad (6)$$
$$\times \delta(\epsilon_{k_1} + \epsilon_{k_2} - \epsilon_{k_1+q} - \epsilon_{k_2-q}).$$

The integrated intralayer scattering contributions cancel due to the condition of the vanishing drag current. Note that the above integral shares several features with the general expression for the drag response function as obtained from the Kubo formula [11, 19]. In particular, it is symmetric in the layer indices, a property that depends crucially on the correct form of $\psi_2(k)$. For a simplified discussion of the most important points, we assume that the interlayer scattering rate W^{12} depends only on the momentum transfers q and energy transfers $\omega = \epsilon_{k_1+q} - \epsilon_{k_1}$ and that the momentum transfers are so small that one can approximate $v_{k+q} - v_k$ by $q(dv_k/dk)$. The drag resistivity $\rho_D = -E_1/j_2$ can then be written as

$$\rho_D = \frac{1}{2\pi e^2} \frac{1}{v_{F1} v_{F2}} \frac{1}{m_1 m_2} \int_0^\infty dq q^2 \int_{-\infty}^{\infty} \frac{d\omega/T}{\sinh^2(\omega/2T)} \quad (7)$$
$$\times W^{12}(q, \omega) \mathrm{Im}\chi_1(q, \omega) \mathrm{Im}\chi_2(q, \omega),$$

where $\chi_\alpha(q, \omega)$ is the dynamical density correlation function in layer α, and the effective masses are given by the curvature of the dispersion relations at the Fermi level

$$\frac{1}{m_\alpha} = \left.\frac{dv_{k\alpha}}{dk}\right|_{k_{F\alpha}}. \quad (8)$$

The integral in Eq. (7) is always positive. The sign of ρ_D is thus given by the sign of the effective masses, that is, by the curvature of the dispersion at the Fermi level. Negative drag is obtained when the dispersion in one layer is electron-like, and holelike in the other. The drag vanishes if the Fermi level in one of the layers is at an inflection point of the dispersion. For the quadratic dispersion $\epsilon_k = k^2/2m$, one has $v_F = k_F/m$ and Eq. (7) reduces to a one-dimensional version of the well-known semiclassical result for drag between free electrons in two dimensions [2]. Returning to Eq. (6), it is not hard to generalize the above results on the sign of ρ_D allowing for larger momentum transfers q and general momentum dependences of W^{12}. Note that, in our case of chiral electrons, no backscattering is possible, unlike the situation in quantum wires [20].

For parallel links, energy and momentum conservation restrict the allowed scattering processes very strongly. At low temperatures, this leads to an exponential suppression of the drag between parallel links. This has nothing to do with the exponential suppression of drag observed in the experiments, since the links are generically not parallel. For nonparallel links, the sum of the momenta of the two links is no longer conserved in the scattering process. Hence, the scattering processes are suppressed much less at low temperatures. Computing the drag between the nonparallel links from the linearized Boltzmann equation (a straightforward generalization of the above steps for parallel links) yields a quadratic temperature dependence at low T. The momentum transfers in the drag and drive links are, however, still correlated for nonparallel links, especially when the angle between the links is not very large, and the relative sign of the curvature of the dispersion in the drive and drag layer, respectively, determines the sign of the drag. The average curvature vanishes for states in the center of the Landau level, while it is positive for energies below and negative for energies above ϵ_0. We thus understand the observation of negative drag when the Landau level in one layer is less than half filled, and more than half filled in the other.

Spin can be easily included in the above picture. Since the interlayer interaction is spin independent, one simply has to sum over the two spin species (up and down) in both the drive and drag layer, while taking the (exchange enhanced) Zeeman spin splitting of the Landau levels into account. If the Fermi level of one layer lies between the centers of the highest occupied Landau levels for up and down spins, respectively, positive and negative contributions to the drag partially cancel each other. The cancellation is complete due to particle-hole symmetry in the case of odd integer filling, as observed in experiment.

Within our semiclassical picture, anomalous drag, especially negative drag, is suppressed at temperatures above the Landau level width, because then the electron and holelike states within the highest occupied level are almost equally populated. This agrees with the results from the Born approximation [14] and also with experiments. For the low temperature asymptotics of the drag, the semiclassical theory yields two different types of behavior depending on the filling. If the Fermi level does not hit any extended states (for either spin species), the drag should vanish exponentially for $T \longrightarrow 0$, since thermal activation or scattering of electrons into extended states is then suppressed by an energy gap. In contrast, for a Fermi level within the extended states band (for at least one spin species), the gap vanishes and the drag obeys generally quadratic low temperature behavior, as obtained for the drag between nonparallel links. Within the Born approximation, no localization occurs and the drag resistance always vanishes quadratically in the low temperature limit [14]. In high mobility samples, localization is negligible at low magnetic fields, while an increasing amount of states gets localized at higher fields [15].

In summary, we have presented a semiclassical theory for electron drag between two parallel two-dimensional electron systems in a strong magnetic field, which provides a transparent picture of the most salient qualitative features of anomalous drag phenomena

observed in recent experiments [6, 7, 10]. Localization plays a role in explaining activated low temperature behavior, but it is not crucial for anomalous (especially negative) drag per se. A quantitative theory of drag, which covers the whole range from low magnetic fields, where the Born approximation is valid [13, 14], to high fields, where localization becomes important, remains an important challenge for the future.

We gratefully acknowledge several important discussions with Leonid Glazman, in particular, on transport and drag in one-dimensional channels. He led us to the Boltzmann equation analysis of drag between links, which substantiated our semiclassical picture considerably. Special thanks go to Rolf Gerhardts for his help in the early stages of this work. We are also grateful for valuable discussions with E. Brener, W. Dietsche, I. Gornyi, K. von Klitzing, K. Muraki, and, especially, Sjoerd Lok.

REFERENCES

1. T. J. Gramila *et al.*, Phys. Rev. Lett. **66**, 1216 (1991).
2. For a review, see A. G. Rojo, J. Phys.: Condens. Matter **11**, R31 (1999).
3. A.-P. Jauho and H. Smith, Phys. Rev. B **47**, 4420 (1993); L. Zheng and A. H. MacDonald, Phys. Rev. B **48**, 8203 (1993).
4. N. P. R. Hill *et al.*, J. Phys.: Condens. Matter **8**, L557 (1996).
5. H. Rubel *et al.*, Phys. Rev. Lett. **78**, 1763 (1997).
6. X. G. Feng *et al.*, Phys. Rev. Lett. **81**, 3219 (1998).
7. J. G. S. Lok *et al.*, Phys. Rev. B **63**, 041305 (2001).
8. M. C. Bønsager *et al.*, Phys. Rev. Lett. **77**, 1366 (1996); Phys. Rev. B **56**, 10 314 (1997).
9. F. von Oppen *et al.*, Phys. Rev. Lett. **87**, 106803 (2001).
10. K. Muraki *et al.*, Phys. Rev. Lett. **92**, 246801 (2004).
11. A. Kamenev and Y. Oreg, Phys. Rev. B **52**, 7516 (1995).
12. J. G. S. Lok *et al.*, Physica E (Amsterdam) **12**, 119 (2002).
13. M. E. Raikh and T. V. Shahbazyan, Phys. Rev. B **47**, 1522 (1993).
14. I. V. Gornyi *et al.*, Phys. Rev. B **70**, 245302 (2004).
15. M. M. Fogler *et al.*, Phys. Rev. B **56**, 6823 (1997).
16. M. Tsukuda, J. Phys. Soc. Jpn. **41**, 1466 (1976); S. V. Iordanskii, Solid State Commun. **43**, 1 (1982); R. F. Kararinov and S. Luryi, Phys. Rev. B **25**, 7626 (1982); S. A. Trugman, Phys. Rev. B **27**, 7539 (1983); B. Shapiro, Phys. Rev. B **33**, 8447 (1986).
17. D. B. Chklovskii *et al.*, Phys. Rev. B **46**, 4026 (1992); N. R. Cooper and J. T. Chalker, Phys. Rev. B **48**, 4530 (1993).
18. J. T. Chalker and P. D. Coddington, J. Phys. C **21**, 2665 (1988).
19. K. Flensberg *et al.*, Phys. Rev. B **52**, 14761 (1995).
20. The role of backscattering versus small momentum transfer scattering in the drag between quantum wires is analyzed in M. Pustilnik *et al.*, Phys. Rev. Lett. **91**, 126805 (2003).

JETP Letters, Vol. 81, No. 10, 2005, pp. 503–508. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 623–629.
Original Russian Text Copyright © 2005 by Borisov, Zverev.

Single-Particle Spectrum of a Dilute 2D Electron Gas

V. V. Borisov and M. V. Zverev

Russian Research Centre Kurchatov Institute, Moscow, 123182 Russia

e-mail: zverev@mbslab.kiae.ru

Received April 13, 2005

Single-particle spectra of a homogeneous 2D electron gas have been calculated in the microscopic functional approach. The effective mass in this system is found to diverge at $r_s \simeq 7$. It is shown that the inclusion of the local exchange and correlation corrections to the effective interaction modifies the instability picture from that obtained in the random phase approximation: in the latter, the instability arises away from the Fermi surface, whereas, with the introduction of the corrections, it manifests itself as the divergence of the effective mass.
© 2005 Pleiades Publishing, Inc.

PACS numbers: 71.10.Hf, 71.27.+a, 75.30.Cr

Experimental studies of the quasi-2D electron systems in silicon transistors (MOSFETs) showed that a decrease in the electron density down to the critical value $n_c \simeq 8 \times 10^{10}$ cm^{-2} is accompanied by the divergence of the effective quasiparticle mass M^* [1, 2], which is the main characteristic of the quasiparticle spectrum $\varepsilon(p)$ in the Landau theory of the Fermi liquid. At the same time, the Stoner factor $(1 + g_0)^{-1}$, where g_0 is the zero harmonic of the Landau spin–spin interaction, is almost invariable [1]. This testifies to the absence of ferromagnetic instability of the quasiparticle system. Since none of the existing experimental data point to the violation of other Pomeranchuk stability conditions, the divergence of the effective mass means that the description of the system in terms of quasiparticles with a zero-temperature momentum distribution $n_{\rm FL}(p) = \theta(p_F - p)$ (in this paper, we consider homogeneous systems) is impossible because of the violation of the stability condition different from the Pomeranchuk conditions [3–5]. In fact, beyond the point where the group velocity of quasiparticles $v_F = d\varepsilon(p)/dp|_F$ changes sign, in the vicinity of the boundary momentum, a region appears where the allowable variations $\delta n(p)$ with the sign opposite to that of $\xi(p) = \varepsilon(p) - \mu$ (μ is the chemical potential) give rise to negative additions $\delta E_0 = \int \xi(p)\,\delta n(p) d\tau$ ($d\tau$ is a phase volume element) to the ground state energy E_0 of the quasiparticle system. However, as it was shown in [3–5], the instability of the quasiparticle distribution $n_{\rm FL}(p)$ means not a failure of the Landau quasiparticle distribution but its rearrangement caused by the necessity to redefine the function $n(p)$ so as to satisfy the minimum condition for the energy functional $E_0[n(p)]$.

A complete microscopic theory of the state beyond the transition point has not yet been developed. However, many properties of this state can be understood with the use of simple phenomenological models [3–7]. From the analysis [6], it was found that these properties are largely determined by the place in the momentum space where the instability of the distribution $n_{\rm FL}(p)$ arises. Since the instability arises at the point $p = p_b$ of the bifurcation formation in the equation

$$\xi(p, [n_{\rm FL}(p)]) = 0, \tag{1}$$

the microscopic calculation of the quasiparticle spectrum $\xi(p, [n_{\rm FL}(p)]$ of the system in terms of the Fermi liquid can be used in developing the microscopic theory of the state beyond the critical point.

In this paper, we carry out a microscopic calculation of the single-particle spectrum of a 2D electron gas, $\xi(p, [n_{\rm FL}(p)])$, with the aim, first, to determine the values of the parameter r_s (the average distance between particles in Bohr radius units) at which the momentum distribution $n_{\rm FL}(p)$ becomes unstable and, second, to determine the point where the bifurcation arises in Eq. (1) and the scenario of the quasiparticle rearrangement. The first microscopic calculation of the single-particle spectrum of a 2D electron gas was recently performed in [8, 9] in the random phase approximation (RPA). In the functional approach considered below, the RPA appears as the first approximation. We take a further step: we take into account the exchange and correlation corrections to the RPA and see how their inclusion affects the behavior of the single-particle spectrum.

We consider a 2D electron gas with the pair interaction potential $V(q) = 2\pi e^2/q$ in terms of the "jellium" model. The extent to which the system is diluted is determined by the value of the parameter $r_s = \sqrt{2} M e^2/p_{\rm F}$. To calculate the single-particle spectrum, we use *ab initio* the functional approach proposed in [10] and the strategy developed in [11]. The cited

0021-3640/05/8110-0503$26.00 © 2005 Pleiades Publishing, Inc.

papers contain all the detailed explanations that may be of interest to the readers.

Let us consider the scheme of calculating the single-particle spectrum of the system under study. We begin with the formula $E_0 = \tau + W$, which represents the ground state energy E_0 of the system as a sum of the energy τ characterizing the system of noninteracting particles and the interaction energy W. The Hartree term W_H of the interaction energy $W = W_H + W_F + W_c$ is cancelled by the uniform background of the positive charge, the Fock energy

$$W_F = \frac{e^2\rho}{2}\int\frac{d^2q}{2\pi q}[S_0(\mathbf{q}) - \rho\delta(\mathbf{q}) - 1] \tag{2}$$

is expressed through the static form factor

$$S_0(\mathbf{q}) = \frac{2}{\rho}\int\frac{d^2p}{(2\pi)^2}n_{FL}(\mathbf{p})[1 - n_{FL}(\mathbf{p}+\mathbf{q})], \tag{3}$$

and the correlation energy

$$W_c = -\frac{1}{2}\int_0^{e^2}de^2\int\frac{d^2q}{2\pi q}\int_0^{\infty}\frac{d\omega}{\pi}\mathrm{Im}[\chi(\mathbf{q},\omega) - \chi_0(\mathbf{q},\omega)] \tag{4}$$

is determined by the difference between the linear response function of the system $\chi(\mathbf{q}, \omega)$ and the response function of noninteracting particles $\chi_0(\mathbf{q}, \omega)$ [12].

These two functions are related by the formula

$$\chi(\mathbf{q},\omega) = \frac{\chi_0(\mathbf{q},\omega)}{1 - R(\mathbf{q},\omega)\chi_0(\mathbf{q},\omega)}, \tag{5}$$

with the effective interaction [10]

$$R(\mathbf{q},\omega) = \frac{\delta^2 W}{\delta\rho(\mathbf{q},\omega)\delta\rho(-\mathbf{q},-\omega)}. \tag{6}$$

In the effective interaction, the correction to the Coulomb potential can be separated into the exchange correction $R_{ex}(\mathbf{q}, \omega) = \delta^2 W_F/\delta\rho(\mathbf{q}, \omega)\delta\rho(-\mathbf{q}, -\omega)$ and the correlation one $R_c(\mathbf{q}, \omega) = \delta^2 W_c/\delta\rho(\mathbf{q}, \omega)\delta\rho(-\mathbf{q}, -\omega)$. In principle, the exchange correction can be calculated in the closed form. The correlation correction itself can be expressed through $R(\mathbf{q}, \omega)$. This makes it possible to formulate an equation for the effective interaction:

$$R(\mathbf{k},\omega) = \frac{2\pi e^2}{k} + R_{ex}(\mathbf{k},\omega) - \frac{1}{2}\frac{\delta^2}{\delta\rho(\mathbf{k},\omega)\delta\rho(-\mathbf{k},-\omega)} \times\int_0^{e^2}de^2\int\frac{d^2q}{2\pi q}\int_0^{\infty}\frac{d\omega'}{\pi}\mathrm{Im}[\chi(\mathbf{q},\omega') - \chi_0(\mathbf{q},\omega')]. \tag{7}$$

To calculate the variational derivatives on the right-hand side of Eq. (7), we follow the method proposed in [10] and apply the local approximation [13]:

$$\frac{\delta^2\chi(\mathbf{q},\omega)}{\delta\rho(\mathbf{k},\varepsilon)\delta\rho(-\mathbf{k},-\varepsilon)} \simeq \frac{1}{2}\left[\frac{d^2\chi(\mathbf{q}+\mathbf{k},\omega+\varepsilon)}{d\rho^2} + \frac{d^2\chi(\mathbf{q}-\mathbf{k},\omega-\varepsilon)}{d\rho^2}\right]. \tag{8}$$

This approximation proved to be effective in application to a 3D electron gas and neutron matter [10]. The equation for the effective interaction, which is frequency independent in the local approximation, takes the form

$$R(k) = \frac{2\pi e^2}{k} + R_{ex}(k) + R_c(k). \tag{9}$$

The local exchange correction $R_{ex}(k)$ is calculated through the static form factor (3) and, with simple algebra, can be expressed as

$$R_{ex}(k) = -\frac{2e^2}{\pi p_F^2}\int_0^{2p_F}\frac{q^2dq}{\sqrt{4p_F^2 - q^2}}\frac{K(4qk/(q+k)^2)}{q+k}, \tag{10}$$

where $K(z)$ is an elliptical integral of the first kind.

The local correlation correction can be calculated by extending the contour of integration with respect to the frequency along the C_I imaginary axis:

$$R_c(k) = -\frac{1}{2}\frac{d^2}{d\rho^2}\int_0^{e^2}de^2\int\frac{d^2q}{2\pi|\mathbf{q}-\mathbf{k}|} \times\int_{C_I}\frac{d\omega}{2\pi i}\chi_0(q,\omega)\left(\frac{1}{1 - R(q)\chi_0(q,\omega)} - 1\right). \tag{11}$$

In our calculations, we make one iteration step: we substitute the effective interaction $R_0(q) = V(q) + R_{ex}(q)$ into the right-hand side of Eq. (11). Integrating with respect to e^2, we obtain

$$R_c(k) = \frac{e^2}{2}\frac{d^2}{d\rho^2}\int\frac{d^2q}{2\pi}\frac{1}{R_0(q)|\mathbf{q}-\mathbf{k}|} \times\int_{C_I}\frac{d\omega}{2\pi i}(R_0(q)\chi_0(q,\omega) + \ln[1 - R_0(q)\chi_0(q,\omega)]). \tag{12}$$

The accuracy that is achieved with the local approximation and only one iteration used for calculating the correlation correction can be estimated by comparing the resulting characteristics of the 2D electron gas with the results of Monte Carlo (MC) calculations. Figure 1 shows the response function in the static limit, $-\chi(q, 0)$, divided by the electron density ρ for $r_s = 1$ and 5. One can see that the MC results [14] are adequately reproduced in our calculations. Note that the inclusion of the exchange correction alone considerably improves the

agreement with the MC data in comparison with the initial RPA calculations.

Figure 2 compares the calculated correlation energy of the 2D electron gas with the MC data [16]. The RPA overestimates the absolute value of the energy by approximately a factor of two. A major part of this discrepancy is eliminated by taking into account the exchange correction, and the inclusion of the correlation correction further reduces the discrepancy to 12% or less.

Here, it is appropriate to discuss why we consider an unpolarized system. In the RPA, the 2D electron gas is completely polarized as early as at $r_s \sim 5.5$ [9], whereas, in the MC calculations, the energy of the polarized state becomes lower than the energy of the polarized state as late as at $r_s \sim 28$ [16]. The first step beyond the RPA, namely, the calculation of the correlation energy with allowance for the correction $R_{ex}(k)$, changes the situation: up to $r_s \sim 30$, the energy of the unpolarized gas remains lower than the energy of the polarized gas. The inclusion of the correlation term does not introduce any qualitative changes in the situation under study.

Let us proceed to calculating the single-particle spectrum:

$$\varepsilon(p) = \frac{\delta E_0}{\delta n(p)} = \frac{p^2}{2M} + \varepsilon_1(p) + \varepsilon_2(p). \qquad (13)$$

In this formula, the term

$$\begin{aligned}\varepsilon_1(p) &= -\frac{1}{2}\int\frac{d^2q}{2\pi}\frac{e^2}{|\mathbf{p}-\mathbf{q}|} \\ &-\frac{1}{2}\int_0^{e^2} de^2\int\frac{d^2q}{2\pi q}\int_0^{\infty}\frac{d\omega}{\pi}\mathrm{Im}\left[\varphi^2(\mathbf{q},\omega)\frac{\delta\chi_0(\mathbf{q},\omega)}{\delta n(p)}\right],\end{aligned} \qquad (14)$$

where $\varphi(q, \omega) = [1 - R(q)\chi_0(q, \omega)]^{-1}$, is due to the variation of the function $\chi_0(q, \omega)$ involved in Eq. (5) for the response function $\chi(q, \omega)$. The term

$$\varepsilon_2(p) = -\frac{1}{2}\int_0^{e^2} de^2\int\frac{d^2q}{2\pi q}\int_0^{\infty}\frac{d\omega}{\pi}\mathrm{Im}\left[\chi^2(\mathbf{q},\omega)\frac{\delta R(q)}{\delta n(p)}\right] \qquad (15)$$

is due to the variation of the effective interaction in Eq. (5). The substitution of the explicit expression for the variation $\delta\chi_0(\mathbf{q}, \omega)/\delta n(p)$ into Eq. (14) yields

$$\begin{aligned}&\varepsilon_1(p) \\ &= -\frac{2e^2}{\pi}p + \int_0^{e^2} de^2\int_{p_F}^{p}\frac{d^2q}{2\pi|\mathbf{p}-\mathbf{q}|}\mathrm{Re}\varphi^2(\mathbf{p}-\mathbf{q}, \varepsilon_p^0 - \varepsilon_q^0) \\ &+ \int_0^{e^2} de^2\int\frac{d^2q}{2\pi q}\int_{-\infty}^{\infty}\frac{dw}{2\pi}(\varphi^2(q, iw) - 1)\frac{\varepsilon_p^0 - \varepsilon_{\mathbf{p}-\mathbf{q}}^0}{(\varepsilon_p^0 - \varepsilon_{\mathbf{p}-\mathbf{q}}^0)^2 + w^2}.\end{aligned} \qquad (16)$$

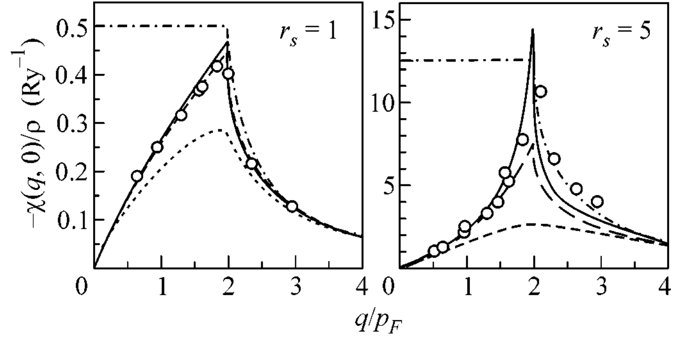

Fig. 1. Response function of the 2D electron gas in the static limit, $-\chi(q, 0)/\rho$, for r_s = (left panel) 1 and (right panel) 5. The MC results [14] are represented by circles, the dash–dotted line shows the response function $-\chi_0(q, 0)/\rho$ [15], and the dotted line represents the RPA calculation. The dashed line shows the calculation including only the correction $R_{ex}(q)$, and the solid line additionally takes into account the correction $R_c(q)$.

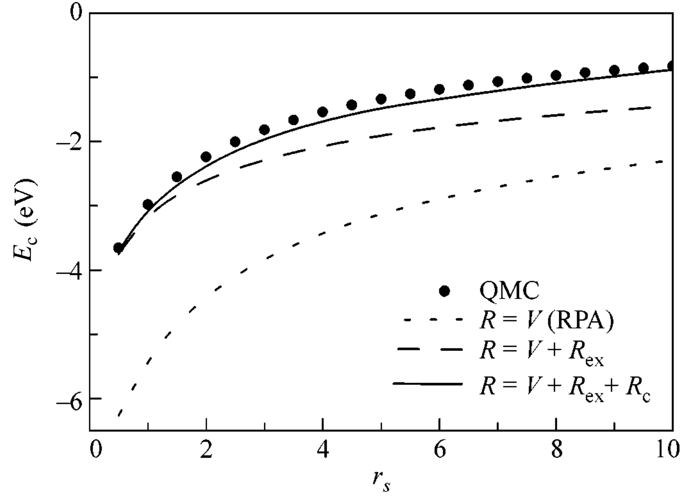

Fig. 2. Correlation energy of the 2D electron gas. The circles represent the MC calculation [16], the dotted line is the RPA calculation, the dashed line shows our calculation taking into account the term $R_0(k)$ of the effective interaction, and the solid line is the calculation with $R(k) = R_0(k) + R_c(k)$.

Thus, if the effective interaction $R(q)$ is determined, the calculation of the term $\varepsilon_1(p)$ of the single-particle spectrum is reduced to calculating the integrals. As for expression (15) for $\varepsilon_2(p)$, it contains not only $R(k)$ but also the variational derivative $\delta R(k)/\delta n(p)$, which requires special calculation. A linear integral equation for this function was derived in [11]. In principle, it can be easily generalized to the 2D case. However, since for the values of $r_s \lesssim 10$, which are considered below, the contribution of $\varepsilon_2(p)$ to the spectrum is small, we can follow [11] and calculate the variation of the local

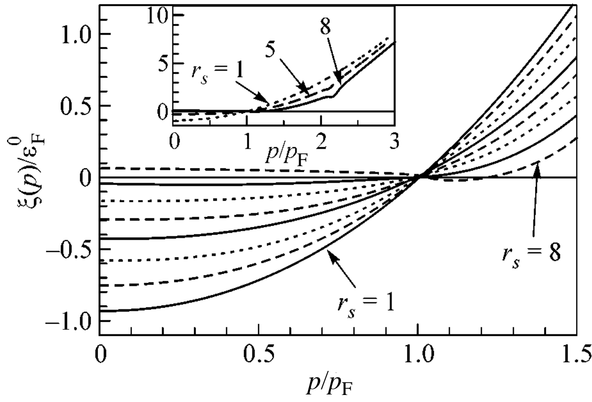

Fig. 3. Single-particle spectra of the 2D electron gas, $\xi(p)$, in ε_F^0 units for the values of the parameter r_s from 1 to 8 with the step $\Delta r_s = 1$. The inset shows the spectra for $r_s = 1$, 5, and 8 in the momentum interval $0 < p/p_F < 3$.

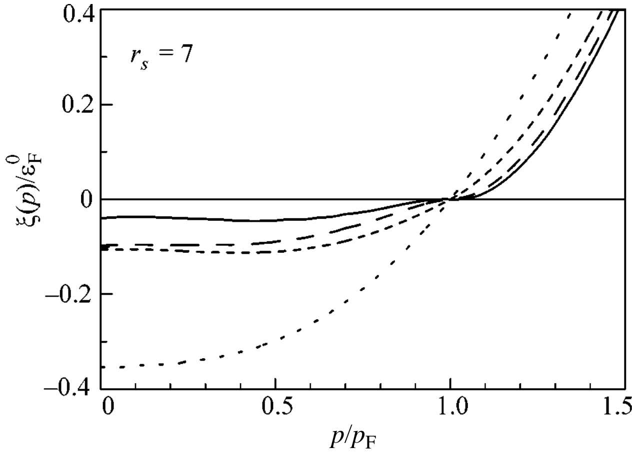

Fig. 4. Single-particle spectra of the 2D electron gas in different approximations for $r_s = 7$. The dotted line shows the RPA calculation, the short-dashed line represents the term $\xi_1(p)$ calculated with the effective interaction $R_0(k)$, the long-dashed line shows the spectrum $\xi_1(p)$ calculated with $R(k) = R_0(k) + R_c(k)$, and the solid line represents the spectrum with the term $\varepsilon_2(p)$.

exchange correction. Varying the static form factor (3) in Eq. (2) and differentiating it twice with respect to the density, we obtain

$$\frac{\delta R_{ex}(k)}{\delta n(p)} = -\frac{\pi e^2}{2p_F^3}\int_0^{2\pi}\left\{-\frac{4}{(l+1)^2}K\left(\frac{2\sqrt{l}}{l+1}\right) + \frac{2}{l+1}\left[\frac{l+1}{l-1}E\left(\frac{2\sqrt{l}}{l+1}\right) + \frac{l-1}{l+1}K\left(\frac{2\sqrt{l}}{l+1}\right)\right]\right\}d\vartheta, \quad (17)$$

where $l = \sqrt{p^2 + k^2 - 2pk\cos\vartheta}/p_F$ and $E(z)$ is an elliptical integral of the second kind. We substitute the variational derivative found above into expression (15) for $\varepsilon_2(p)$, in which we approximate the effective interaction by $R_0(q)$. The integral with respect to the coupling constant can be calculated analytically. As a result, we obtain

$$\varepsilon_2(p) = -\frac{e^2}{2}\int\frac{d^2q}{2\pi}\frac{1}{|\mathbf{p}-\mathbf{q}|}\frac{1}{R_0^2(q)}\frac{\delta R(q)}{\delta n(p)} \times \int_{-\infty}^{\infty}\frac{dw}{2\pi}\left[\frac{R_0(q)\chi_0(q, iw)}{1 - R_0(q)\chi_0(q, iw)} + \ln(1 - R_0(q)\chi_0(q, iw))\right]. \quad (18)$$

The single-particle spectra $\xi(p)$ calculated from Eq. (13), where $\varepsilon_1(p)$ and $\varepsilon_2(p)$ are determined by Eqs. (16) and (18), are shown in Fig. 3 for the values of r_s from 1 to 8 with the step $\Delta r_s = 1$. As r_s increases, the spectrum flattens out in the region of momenta $p < p_F$, and, at $r_s \simeq 7$, the group velocity becomes equal to zero. The inset in Fig. 3 shows the spectra for $r_s = 1$, 5, and 8 in a wider interval of momenta: $0 < p/p_F < 3$. We point to the irregularities in the behavior of spectra near $p \sim 2p_F$, which become more pronounced as r_s increases. These irregularities are caused by the contribution of plasmon poles to the integrals of the quantity $\varphi^2(\mathbf{p} - \mathbf{q}, \varepsilon_p^0 - \varepsilon_q^0)$ with respect to the momentum q in the second term of expression (16) [9]. However, as is seen in Fig. 3, in the region $r_s \lesssim 7$, where the flattening of the spectrum near the Fermi surface clearly manifests itself, these irregularities are still not involved in the formation of the spectrum instability away from p_F. If only the RPA contribution is taken into account, irregularities due to the coupling of single-particle degrees of freedom with plasma excitations are developed much earlier than the flattening of the spectrum near p_F takes place and bifurcation arising in Eq. (1) at $p \sim 2p_F$ appears earlier than the effective mass diverges [9]. Note that, the spectrum calculated in the RPA changes its sign earlier than $M^* \longrightarrow \infty$ in the region of momenta $p \ll p_F$ as well [8, 9].

With the inclusion of the exchange contribution $R_{ex}(k)$, the behavior of the term $\xi_1(p) = \varepsilon_1(p) - \mu$ of the single-particle spectrum remains qualitatively the same as that obtained with the RPA: in both cases, a change of sign occurs in the region of $p \ll p_F$. However, in the RPA, the sign changes at $r_s \simeq 13$, whereas, with the exchange correction, it changes at $r_s \simeq 8$. This is illustrated in Fig. 4, where the spectra calculated in different approximations are displayed. Note that, with the exchange correction taken into account, the effective

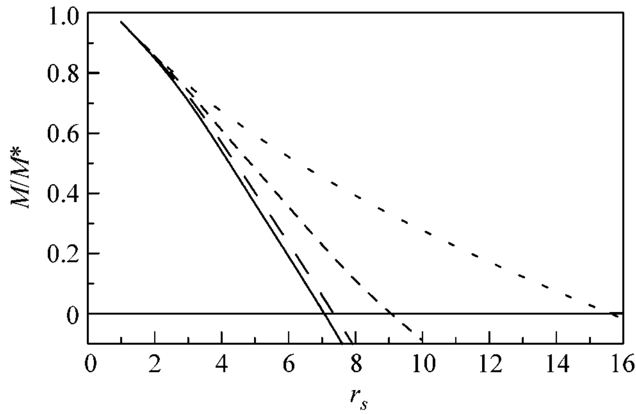

Fig. 5. Ratio M/M^* calculated for the 2D electron gas in different approximations versus the parameter r_s. The notation for the curves is the same as in Fig. 4.

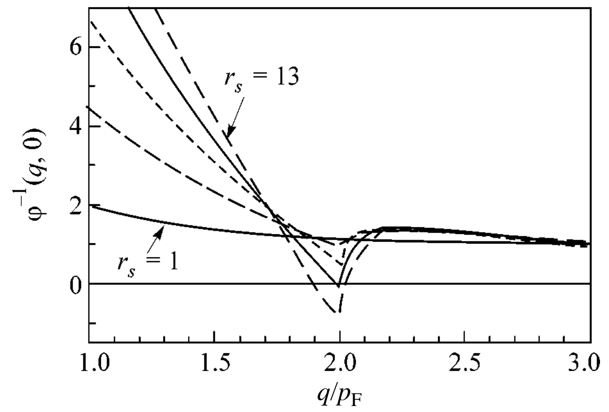

Fig. 6. Static dielectric function $\varphi^{-1}(q, 0)$ calculated for r_s from 1 to 13 with the step $\Delta r_s = 3$ versus q/p_F.

mass $M/M^* = (M/p_F)(d\varepsilon(p)/dp|_{p = p_F}$ diverges at $r_s \simeq 9$ (see Fig. 5). The correlation correction $R_c(k)$, being introduced in the calculations, makes an almost an order of magnitude smaller contribution to the spectrum $R_{ex}(k)$, as compared to the contribution due to the exchange correction, but, owing to the drastic change in the vicinity of $p = p_F$, it makes a considerable contribution to the effective mass, which is illustrated in Fig. 5. As a result, the behavior of the term $\xi_1(p)$ calculated with the effective interaction $R(k) = R_0(k) + R_c(k)$ is entirely different from that calculated with allowance for the exchange correction alone: the instability now occurs immediately on the Fermi surface at $r_s \simeq 7.3$. The term $\xi_2(p)$ is not small in magnitude but has an insignificant slope at the point $p = p_F$ and does not make any considerable contribution to the effective mass: it only shifts the point corresponding to the zero value of the ratio M/M^* from $r_s \simeq 7.3$ to $r_s = 7$ (see Fig. 5). The inclusion of this contribution does not change the character of the instability: the latter arises at the point p_F, i.e., as the divergence of the effective mass. This can be seen from Fig. 4, where the solid line represents the spectrum calculated with allowance for both contributions at $r_s = 7$. The calculation shows that the first and second derivatives of the spectrum simultaneously become equal to zero at the point p_F; therefore, near the Fermi surface, we have $\xi(p) \propto (p - p_F)^3$. We also note that, because of the flatness of the spectrum (within the interval $0 < p < p_F$, its absolute value does not exceed $0.05\varepsilon_F^0$), the corrections ignored by us (the corrections associated, e.g., with the subsequent iteration steps in solving Eqs. (9)–(11) or with the change from the local approximation to the exact calculation of variational derivatives) could shift the point of instability formation p_b to the region of $p < p_F$. Thus, although the situation of the instability formation exactly on the Fermi surface is not universal, precisely this situation is realized in the calculations performed in the framework of approximations used by us in this paper.

Scenarios of the quasiparticle rearrangement initiated by the appearance of bifurcation in Eq. (1) are investigated in [7]. If $p_b < p_F$, i.e., two new roots p_1 and p_2 of Eq. (1) appear within the Fermi circle, an empty ring is formed in the filled Fermi circle. If $p_b > p_F$, the rearranged momentum distribution of quasiparticles $n(p)$ represents a Fermi circle surrounded by a filled ring. In both cases, the momentum distribution of quasiparticles is multiply connected. The appearance of bifurcation in Eq. (1) at the Fermi surface (or very close to it), as in the given calculation ($p_b \simeq p_F$), corresponds to the rearrangement scenario that was studied with the use of models in [3–7, 10] (including the study of the effect of rearrangement on the quasiparticle interaction amplitude [7]). For this scenario, it was found that, in the vicinity of the initial Fermi surface, a fermion condensate is formed: a group of single-particle states with a flat spectrum $\xi(p) \equiv 0$ and a quasiparticle filling $0 < n(p) < 1$. It should be noted that, in all the scenarios, at the critical point and beyond it, the system possesses non-Fermi-liquid properties, which, however, can be explained in quasiparticle terms [17]. For example, in the scenario corresponding to $p_b \simeq p_F$, the appearance of an inflection of the spectrum at the Fermi surface, i.e., the cubic form of the spectrum $\xi(p) \propto (p - p_F)^3$ at the critical point, leads to a quite different temperature dependence of the thermodynamic characteristics of the system at this point, as compared to the conventional Fermi liquid [17].

We consider one more feature of a dilute 2D electron gas. In [18], it was shown that the divergence of M^* is a precursor of the formation of a charge density wave in a 2D or 3D homogeneous Fermi system. As is known, the instability with respect to the formation of this inhomogeneous phase appears when the static dielectric function $\varphi^{-1}(q, 0)$ becomes equal to zero at a certain value of the momentum $q \sim 2p_F$. Figure 6, which

shows the function $\varphi^{-1}(q, 0)$ for r_s from 1 to 13 with a step of 3, demonstrates the appearance of the zero value of this function near $q = 2p_F$ at $r_s = r_s^{CDW} \simeq 10$. Note that, in the RPA, the static dielectric function reaches zero at no value of r_s, because the product $V(q)\chi_0(q, 0)$ is negative for any value of q. The product $R_0(q)\chi_0(q, 0)$ is everywhere negative as well. However, when the correlation correction is taken into account, the effective interaction $R(q) = R_0(q) + R_c(q)$ changes its sign in the vicinity of $q \sim 2p_F$ at $r_s \simeq 3$. As a result, in this vicinity, an increase in r_s is accompanied by an increase in the positive product $R(q)\chi(q, 0)$ and, at $r_s^{CDW} \simeq 10$ at the point $q_c \simeq 2p_F$, this product compensates for the unity in the expression for $\varphi^{-1}(q, 0)$. The pole that appears in response function (5) at $q = q_c$ and $\omega = 0$ leads to an instability with respect to the condensation of charge density waves with the vector q_c [12]. The comparison of the values $r_s^\infty \simeq 7$ and $r_s^{CDW} \simeq 10$ shows which of the two instabilities arises first: according to [18], in a 2D electron gas, an increase in the parameter r_s first leads to the divergence of the effective mass and then to the condensation of the charge density waves.

The value of $r_s^\infty \simeq 7$ determined above reasonably agrees with the value $r_s^c \simeq 9.3$ [2] corresponding to the density $n_c \simeq 8 \times 10^{10}$ cm^{-2} at which the divergence of the effective mass is observed for the electron system of an Si MOSFET [1, 2]. However, some comments are necessary on this point. Although the experimental value of r_s^c takes into account the corresponding values of the band mass, degeneracy multiplicity, and permittivity, a quantitative comparison requires the calculation of the spectrum of the quasi-2D electron gas layer at the semiconductor–insulator boundary with the interaction taking into account the finite thickness of the layer. In addition, an actual system contains disordered charged impurities, the scattering from which may contribute to the quasiparticle dispersion. In this paper, we did not intend to give a quantitative description of experimental data. The main result obtained by us is as follows: the calculation of the single-particle spectrum of a 2D electron gas on the basis of the functional approach shows that the inclusion of the exchange and correlation corrections to the effective interaction, first, noticeably shifts the value of r_s^∞, at which the effective quasiparticle mass diverges, from the RPA result [8, 9] and, second, changes the instability picture. In the RPA, the instability arises away from the Fermi surface, which corresponds to the quasiparticle rearrangement scenario with the formation of a multiply connected Fermi surface. In our calculations, the instability arises as the divergence of the effective mass; i.e., the first derivative of the single-particle spectrum $\xi(p)$ becomes equal to zero at the Fermi surface simultaneously with its second derivative. This form of instability initiates the Fermi-condensate scenario of rearrangement in the quasiparticle system.

We are grateful to V.A. Khodel for many useful discussions and to V.T. Dolgopolov, Yu.M. Kagan, V.M. Pudalov, A.A. Shashkin, V.M. Yakovenko, S. Das Sarma, and Y. Zhang for discussing the problems considered in this paper. This work was supported in part by the Russian Ministry of Science (project no. NSh-1885.2003.2), the NSF (grant no. 9900713), the McDonnel Center for Space Sciences, the Dynasty Foundation (a student's grant), and the Russian Research Centre Kurchatov Institute (young scientist's grant no. 3, 2004). One of us (M.V.Z.) is grateful to the Department of Physics of Washington University, St. Louis, for hospitality extended to him during his visit in summer 2004, when this work was carried out.

REFERENCES

1. A. A. Shashkin, S. V. Kravchenko, V. T. Dolgopolov, *et al.*, Phys. Rev. B **66**, 073303 (2002).
2. V. M. Pudalov, M. E. Gershenson, H. Kojima, *et al.*, Phys. Rev. Lett. **88**, 196404 (2002).
3. V. A. Khodel and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **51**, 488 (1990) [JETP Lett. **51**, 553 (1990)].
4. G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 208 (1991) [JETP Lett. **53**, 222 (1991)].
5. P. Nozières, J. Phys. I **2**, 443 (1992).
6. V. A. Khodel, M. V. Zverev, and J. W. Clark, JETP Lett. (in press).
7. M. Baldo, V. V. Borisov, J. W. Clark, *et al.*, J. Phys.: Condens. Matter **16**, 6431 (2004).
8. Y. Zhang and S. Das Sarma, cond-mat/0312565.
9. Y. Zhang, V. M. Yakovenko, and S. Das Sarma, cond-mat/0410039.
10. V. A. Khodel, V. V. Khodel, and V. R. Shaginyan, Phys. Rep. **249**, 1 (1994).
11. M. V. Zverev, V. A. Khodel, and V. R. Shaginyan, Zh. Éksp. Teor. Fiz. **109**, 1054 (1996) [JETP **82**, 567 (1996)].
12. D. Pines and P. Nozières, *Theory of Quantum Liquids* (Benjamin, New York, 1966; Mir, Moscow, 1967).
13. Y. Kavasoe, H. Yasuhara, and M. Watabe, J. Phys. C: Solid State Phys. **10**, 3923 (1977).
14. S. Moroni, D. M. Ceperley, and G. Senatore, Phys. Rev. Lett. **69**, 1837 (1992).
15. F. Stern, Phys. Rev. Lett. **18**, 546 (1967).
16. Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **48**, 12 037 (1993).
17. J. W. Clark, V. A. Khodel, and M. V. Zverev, Phys. Rev. B **71**, 012401 (2005).
18. V. A. Khodel', V. R. Shaginyan, and M. V. Zverev, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 242 (1997) [JETP Lett. **65**, 253 (1997)].

Translated by E. Golyamina

JETP Letters, Vol. 81, No. 10, 2005, pp. 509–513. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 630–635.
Original Russian Text Copyright © 2005 by Golovenchits, Sanina.

Dielectric and Magnetic Properties of the Multiferroic $Tb_{(1-x)}Bi_xMnO_3$: Electric Dipole Glass and Self-Organization of Charge Carriers

E. I. Golovenchits and V. A. Sanina

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: e.golovenchits@mail.ioffe.ru
Received April 13, 2005

Single crystals of the new multiferroic $Tb_{(1-x)}Bi_xMnO_3$ have been grown and studied. A semiconductor compound with $x = 0.05$ is investigated in most detail. At temperatures $T \geq 165$ K, the electric dipole glass state is realized in the crystal. Localized charge carriers form conducting drops of electrons and holes, which are located predominantly in thin layers at the boundaries of polar domains. When drops escape as the temperature increases, jumps in conductance and capacitance are observed. The state of drops is controlled by low bias voltage. The long-range magnetic order arises at temperatures $T \leq 90$ K. Negative magnetoresistance is observed at temperatures of the existence of localized charge carriers. *© 2005 Pleiades Publishing, Inc.*

PACS numbers: 75.47.Lx, 76.50.+g, 77.80.–e

We present the results of investigation of the dielectric constant, magnetic susceptibility, and magnetoresistance of the new multiferroic $Tb_{(1-x)}Bi_xMnO_3$, which is a solid solution of the initial compounds $TbMnO_3$ and $BiMnO_3$. The latter compound has long been known as a multiferroic with ferromagnetic and ferroelectric orderings at the Curie temperatures T_C = 105 and 750–800 K, respectively, and has monoclinic symmetry with the space group *C*2 [1, 2]. As was recently found, the $TbMnO_3$ compound, which has a structure of a rhombically distorted perovskite (space group *Pbnm*), is a multiferroic with magnetic and ferroelectric orderings at temperatures of 40 and 30 K, respectively [3]. At close ordering temperatures, coupling between the magnetic and ferroelectric states are maximal. Growing crystals of the solid solutions $Tb_{(1-x)}Bi_xMnO_3$, we hoped to obtain multiferroics for small x values with the perovskite structure that have the temperatures of magnetic and ferroelectric orderings, which remain close to each other, higher than those for $TbMnO_3$. The $Tb_{(1-x)}Bi_xMnO_3$ crystals obtained for $x \leq 0.2$ really have the perovskite structure (space group *Pbnm*), have the magnetic and polar (electric dipole glass) orders, and are insulators or semiconductors depending on x. This paper presents the results of investigation of the semiconductor $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$. The symmetry and composition of the crystal are determined by the x-ray structural phase analysis and fluorescence x-ray method, respectively. Nonstoichiometry in oxygen leads to the appearance of electron charge carriers, which ensure the addition of the corresponding number of Mn^{4+} ions to the basic Mn^{3+} ions. As a result, the $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$ compound is a strongly compensated semiconductor.

The capacitance and conductance were measured in the frequency range 0.012–100 kHz and the temperature range 5–350 K by a universal LCR-819 meter. Three capacitors were made of samples with approximately identical areas (~3 mm^2) and various thicknesses (0.4–1.5 mm). The contacts of the capacitors are created by gold deposition. The approximate coincidence of the specific conductance for samples with various thicknesses indicates that the measured quantities characterize the bulk properties of the capacitors.

In the $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$ crystal at temperatures $T > 150$ K, a state with very high capacitance (~10^3 pF) arises, which corresponds to the effective dielectric constant $\varepsilon_{eff} \sim 5 \times 10^4$. Such high capacitances are usually characteristic of layered structures: ferroelectric–metal systems and *p–n* junctions in semiconductors. Jumps in capacitance are observed at temperatures that increase with the frequency (see Fig. 1). For low temperatures up to 150 K, the capacitance is low (3–4 pF) and is independent of the temperature. As is seen in Fig. 1, the capacitance increases irregularly with frequency: maximum jumps in capacitance are observed for low frequencies (up to 1 kHz). In the temperature range 180–290 K, an increase in capacitance slows down and then a new increase begins with $T > 290$–300 K. As the frequency increases, the capacitance jumps first decrease, and, beginning with a frequency of 20 kHz, the capacitance remains nearly the same as that for low temperatures up to a temperature of $T \simeq 240$ K, near which the anomalous behavior of the capacitance,

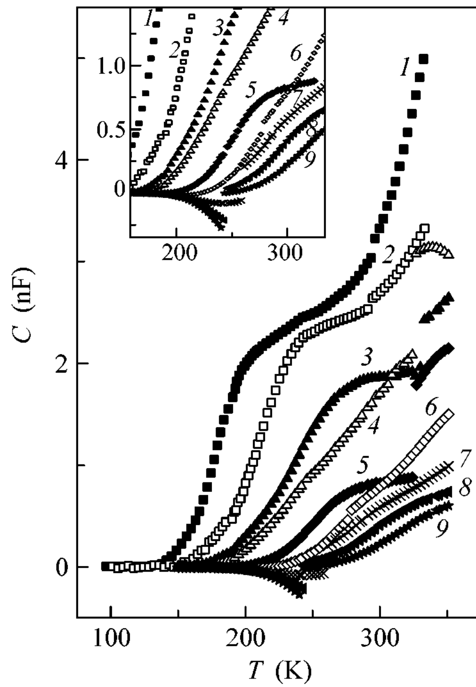

Fig. 1. Temperature dependence of capacitance for frequencies (*1*) 0.012, (*2*) 0.1, (*3*) 0.5, (*4*) 1.0, (*5*) 5.0, (*6*) 10.0, (*7*) 20.0, (*8*) 40.0, and (*9*) 66.7 kHz. Contacts are deposited on the developed (101) plane of the crystal. The alternating electric field **e** || [101]. The inset: the region near the capacitance jumps in a larger scale.

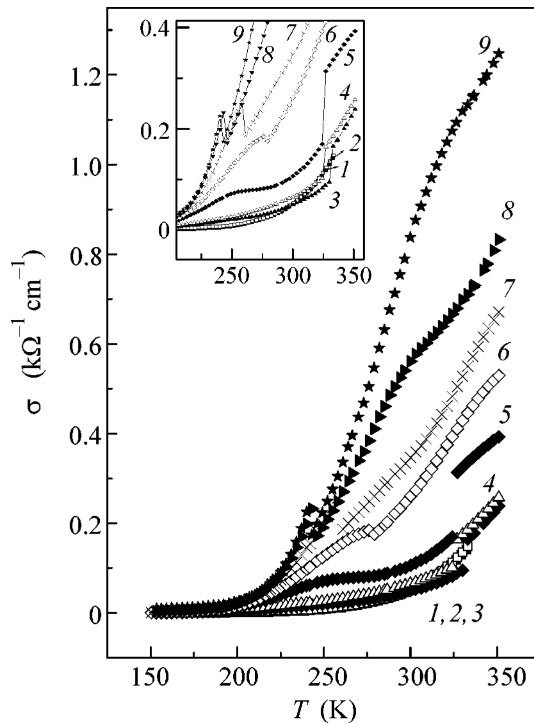

Fig. 2. Same as in Fig. 1, but for the conductance.

jump, and subsequent increase with temperature are observed (see Fig. 1).

The frequency dispersion of capacitance indicates that the state arising in the crystal at $T > 150$ K is inhomogeneous. The dielectric response of the crystal at a given frequency is observed predominantly from restricted domains of the crystal that have the lifetime satisfying the condition $\tau = 1/\omega$. According to Fig. 1, the dielectric state of larger domains of the crystal that give response at low frequency differs substantially from the state of small domains that give response at high frequency.

The dc σ_{dc} and ac σ_{ac} conductances were measured by the four-terminal method and capacitance method, respectively. The dc conductance σ_{dc} is approximately equal to the ac conductance σ_{ac} for low frequencies (from 12 to 500 Hz). For temperatures below 330 K, the low-frequency conductance is of hopping character with the thermal activation barrier $E_{A\sigma} \simeq 215$ meV. We attribute a jump increase in the conductance near 330 K to the transition from the hopping to band conductance.

As is seen in Fig. 2, σ_{ac} increases with the frequency. The observed frequency dependence of the conductance is characteristic of amorphous semiconductors [4]. However, in contrast to them, the conductance for frequencies above 20 kHz falls stepwise at the temperature $T \simeq 240$ K, which coincides with the temperature of the step increase in capacitance at the same frequencies (see Fig. 1). We think that charge carriers are localized in small restricted domains of the crystal, whereas large domains are dielectric. At $T \simeq 240$ K, the state of localized charge carriers in small restricted domains of the crystal changes stepwise. They transit instantaneously to the bottom of the conduction band and the state of these domains becomes more dielectric. In this case, the through conductance at low frequencies increases slightly (see Fig. 2).

Using the temperature dependences of the low-frequency capacitance, one can evaluate the activation barrier for the polarizability E_{AP} of large domains of the crystal. If T_f is the temperature of the maximum of the derivative of $C(T)$ at a given frequency ω, it is easy to show that the Arrhenius law is valid and $\tau = 1/\omega = \tau_0\exp(E_{AP}/kT_f)$, where $E_{AP} = 275$ meV. We note that the activation barriers E_{AP} and $E_{A\sigma}$ differ from each other. At temperatures of a sharp increase in the low-frequency capacitance (Fig. 1), maxima in dielectric losses do not appear (see Fig. 2), because conductance-induced losses are much larger.

Dispersion curves for capacitance (polarizability), particularly for low frequencies (Fig. 1), are similar to those usually observed for structural electric dipole glasses (such as spin glass) [5, 6]. We think that, as is usual for glasses, the frequency dispersion in our case is due to the presence of relaxators with a broad spectrum of the relaxation times ($\tau_{min} \ll \tau \ll \tau_{max}$). However, thorough analysis reveals certain differences of the

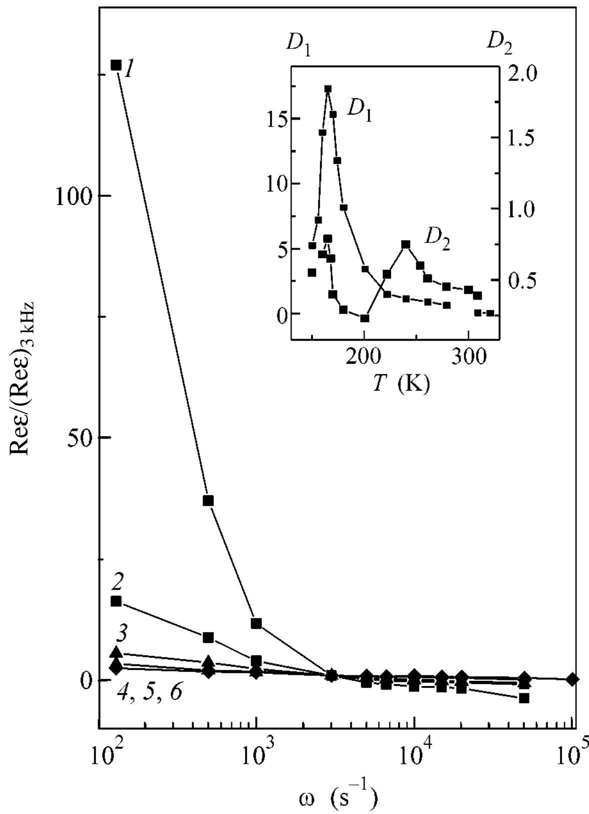

Fig. 3. Frequency dependence of the relative real part of Reε for temperatures (*1*) 180, (*2*) 200, (*3*) 220, (*4*) 240, (*5*) 280, and (*6*) 320 K. The inset: the temperature dependence of the densities of states D_1 and D_2.

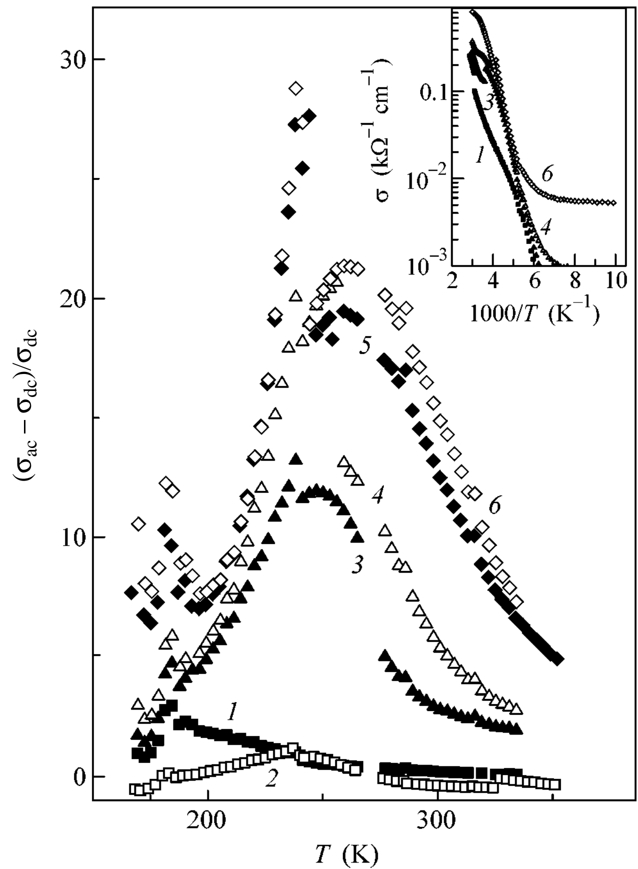

Fig. 4. Temperature dependence of the relative conductance σ_{ac} for frequencies (*1*) 1.0, (*2*) 5.0, (*3*) 10.0, (*4*) 20.0, (*5*) 40.0, and (*6*) 50.0 kHz. The inset: σ_{ac} vs. $1000/T$.

behavior observed here from the usual behavior for glasses. Figure 3 shows the frequency dependences of the real part of the dielectric constant (Reε), which enable one to obtain the density of states of relaxators $D(1/\omega, T)$ [5, 6]. A linear dependence of Reε on lnω is usually observed in spin glass systems, which indicates that the density of states of relaxators $D(1/\omega, T)$ is independent of $1/\omega$ within the limit of this dependence. In the case under consideration (see Fig. 3), there are two sections with a constant density of states:

$$D_1: (1/\omega)_{cr} < 1/\omega < (1/\omega)_{max},$$

$$D_2: (1/\omega)_{min} < 1/\omega < (1/\omega)_{cr}),$$

where $(1/\omega)_{cr} = 2 \times 10^{-5}$ s. As is seen in the inset in Fig. 3, for both sets of relaxators D_1 and D_2, there is a maximum at $T \simeq 165$ K corresponding to the glass freezing temperature. It is seen that an additional maximum in D_2 exists at $T \simeq 240$ K for smaller domains of structural correlations. We recall that jumps in capacitance and conductance at high frequencies are observed near 240 K. Thus, electric dipole glass arises in the $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$ crystal at temperatures above 165 K, and its state changes near 240 K.

We discuss in more detail the features of the behavior of high-frequency conductance. As is seen in Fig. 4, the conductance σ_{ac} is much higher than σ_{dc} in the temperature range 200–300 K. For higher temperatures, when the band conductance arises, there is no difference between two types of conductance; i.e., the frequency dispersion disappears. The jump in σ_{ac} at $T \simeq$ 240 K is manifested against the broad-maximum background in σ_{ac} in Fig. 4. It is natural to attribute this maximum to localized charge carriers. Assuming that the maximum in σ_{ac} arises under the condition $\omega\tau = 1$ and taking $\tau = \tau_0\exp(E_A/kT)$, we obtain $E_A = 280$ meV; i.e., $E_A \simeq E_{AP}$. Thus, the state of localized charge carriers is determined by the polar state of the crystal. In this case, we think that largest restricted domains in the crystal with the density of state D_1 are dielectric and have large polarizability, whereas the small domains with the density of state D_2 contain localized charge carriers that change their state at $T \simeq 240$ K.

The jumps in conductance and capacitance for high frequencies depend substantially on the bias voltage $U_=$

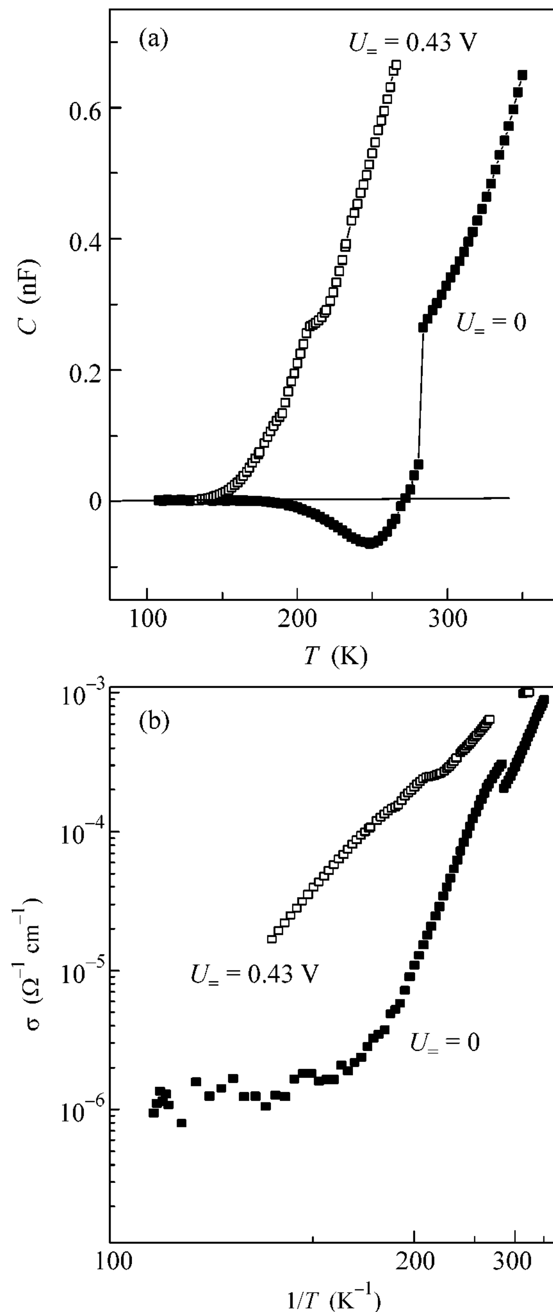

Fig. 5. Effect of the bias voltage ($U_= \parallel e$) at a frequency of 20 kHz on the (a) capacitance and (b) conductance.

(see Fig. 5). As $U_=$ increases, the state changes first slightly and then stepwise at $U_= \simeq 0.43$ V. Hence, a bias of about 5 V/cm eliminates the inhibition of growth and the jump in capacitance. In this case, an increase in capacitance and conductance begins with a temperature of 165 K and the activation barrier for charge carriers is the same as that for σ_{dc}.

Let us analyze the experimental data. As was mentioned above, the $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$ crystal contains carriers of two signs—electrons and holes with nearly identical concentrations. In addition, there is a wide set of limited polar domains (electric dipole glass); i.e., charge carriers are in a medium with a random potential. For this situation, the Efros–Shklovskii model [7] is applicable, which describes the formation of conducting drops of electrons and holes in an amorphous semiconductor with strong compensation.

The random potential of the lattice in the $Tb_{0.95}Bi_{0.05}MnO_{3+0.005}$ crystal is formed due to the coexistence of a wide set of polar domains of different sizes and polarization orientation. In this situation, localized charge carriers are self-organized. They tend to be located at the boundaries of polar domains with oppositely oriented polarizations for which the jump in the electric field is maximal. Such carriers are most localized and they reduce the high-frequency polarizability (capacitance) in a temperature range where this localization holds. In this case, there are charge carriers that are localized at the boundaries of polar domains with smaller jumps in the electric field and they form a broad maximum (see Fig. 4).

Barriers for localized charge carriers at high frequencies in the temperature range to the conductance jump are in the range 0.21–0.28 eV and increase with the frequency (see the inset in Fig. 4). Taking into account that the atomic scale of one lattice constant corresponds to a barrier of 1 eV, one can evaluate the thickness of the layer, where carriers are located, as 4–5 lattice constants. For such small distance between carriers, strong Coulomb repulsion arises, which reduces the activation barrier for the "evaporation" of pairs of these carriers (see [4]). Charge carriers can evidently be localized at the boundaries between polar domains whose width is such that the Coulomb repulsion does not exceed the polarization barrier. In other words, the critical width of the boundary and, correspondingly, the highest frequency F_{cr} must exist to which jumps in conductance and capacitance can be observed. Measurements for $F = 1$ MHz show the absence of jumps. Therefore, 100 kHz $< F_{cr} <$ 1 MHz. We think that jumps in capacitance and conductance for high frequencies near 240 K are associated with the evaporation of drops of charge carriers localized at the most thin boundaries of polar domains. It is also clear why small bias voltage, which is efficiently enhanced by many orders of magnitude when it is applied to such thin layers, reduces the polarization barrier and evaporation of drops.

Analysis of the temperature dependence of the magnetic susceptibility (Fig. 6) shows the presence of anomalies at $T \simeq 6$ and 90 K, which testify to the magnetic phase transitions. The first and second transitions most probably occur in the Tb and Mn ion subsystems, respectively. The inset in Fig. 6 indicates the presence of negative magnetoresistance in the temperature range where localized charge carriers exist. There is an obvious correlation between the form of the $\sigma_{ac}(T)$ dependence (Fig. 4) and the temperature dependence of the magnetoresistance (the inset in Fig. 6). Thus, polar

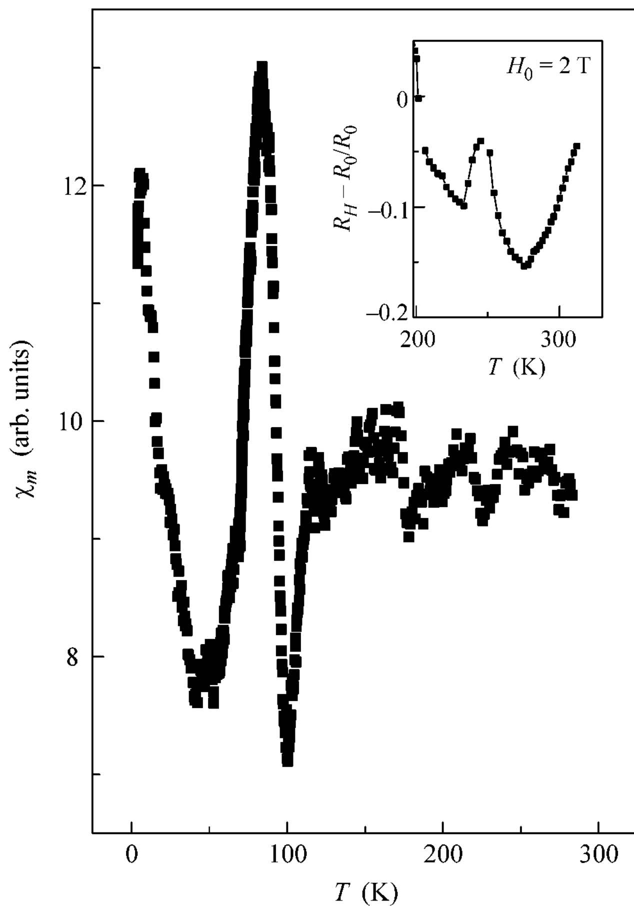

Fig. 6. Temperature dependence of the magnetic susceptibility for the varying magnetic field $\mathbf{h} \perp b$ at a frequency of 10 kHz. The inset: the negative magnetoresistance measured with direct current in the magnetic field H_0 = 2 T, $H_0 \perp b$.

domains of different sizes, as well as Mn^{3+} and Mn^{4+} ions, coexist in the manganite under consideration. In this manganite, the phase separation, which is standard for manganites with colossal magnetoresistance, occurs specifically. Indeed, polar domains in electric dipole glass are dielectric and contain predominantly Mn^{3+} ions, whereas the conducting drops at the boundaries of the polar domains that exist at temperatures of 190–290 K are ferromagnetically correlated (superparamagnetic) and are responsible for the appearance of negative magnetoresistance.

Thus, the $Tb_{0.95}Bi_{0.05}MnO_{3 + 0.005}$ solid solution with the low concentration of Bi ions has the symmetry of rhombically distorted perovskite and is multiferroic for which the temperatures of the magnetic ordering and appearance of the polar state (electric dipole glass) are much higher than the respective temperatures for the pure compound $TbMnO_3$. The compound studied in this work is a strongly compensated semiconductor, where the self-organization of localized charge carriers is observed.

This work was supported by the Russian Foundation for Basic Research (project no. 05-02-16328), by the Presidium of the Russian Academy of Sciences (fundamental research project “Quantum Macrophysics”), and by the Division of Physical Sciences of the Russian Academy of Sciences (project “Phase Transitions in Condensed Media”).

REFERENCES

1. V. A. Bokov, I. E. Mel’nikova, S. A. Kizhaev, *et al.*, Fiz. Tverd. Tela (Leningrad) **7**, 3695 (1965) [Sov. Phys. Solid State **7**, 2993 (1965)].
2. T. Kimura, S. Kawamoto, I. Yamada, *et al.*, Phys. Rev. B **67**, 180401(R) (2003).
3. T. Kimura, T. Goto, H. Shintani, *et al.*, Nature **426**, 55 (2003).
4. A. R. Long, Adv. Phys. **31**, 587 (1982).
5. S. L. Ginzburg, *Irreversible Processes in Spin Glasses* (Nauka, Moscow, 1989), p. 152 [in Russian].
6. E. Courtens, Phys. Rev. B **33**, 2975 (1986).
7. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer, New York, 1984).

Translated by R. Tyapaev

JETP Letters, Vol. 81, No. 10, 2005, pp. 514–518. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 636–641.
Original Russian Text Copyright © 2005 by Lutsev, Stogniĭ, Novitskiĭ.

Giant Injection Magnetoresistance in Gallium Arsenide/Granulated Film Heterostructures with Nanosize Cobalt Inclusions

L. V. Lutsev[1], A. I. Stogniĭ[2], and N. N. Novitskiĭ[3]

[1] *OAO Ferrite-Domain Research Institute, St. Petersburg, 196084 Russia*
e-mail: lutsev@domen.ru
[2] *Minsk Research Institute of Radiomaterials, Minsk, 220024 Belarus*
[3] *Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences, ul. Brovki 17, Minsk, 220072 Belarus*

Received March 25, 2005; in final form, April 15, 2005

Electron transport has been studied in the gallium arsenide/granulated SiO_2 film heterostructure with Co nanoparticles and in the gallium arsenide/TiO_2 film heterostructure with Co island sublayers. When electrons are injected from a film into a semiconductor, a new phenomenon is observed, which is called injection magnetoresistance. For the SiO_2(Co)/GaAs structure with 60 at. % Co in a magnetic field of 23 kOe at a voltage of 50 V, the injection magnetoresistance reaches 5200% at room temperature. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 73.20.–r, 73.40.–c, 75.70.Pa

1. Introduction. The main problem of spintronics are the effective generation of spin-polarized current, the injection of spin-polarized electrons into a nonmagnetic material, and the possibility of the spatial transport and change in the spin polarization [1]. For applications, it is extremely desirable to realize these processes at room temperature and to ensure the complete polarization of electrons. The spin transport and the corresponding magnetoresistance have been observed on the following types of heterostructures.

(i) Magnetic material/nonmagnetic semiconductor, where the magnetic material is a magnetic semiconductor [2–4], a half-metallic ferromagnet, or a ferromagnetic metal [5–7]. The spin injection from the magnetic material into the magnetic semiconductor is observed at low temperatures. At room temperature, the effects of spin-polarized transport are small or vanish.

(ii) Layered structures consisting of alternating ferromagnetic and nonmagnetic metals [8–10]. The resistance of these structures is minimal when the magnetic moments of the ferromagnetic layers are parallel to each other and is maximal for the antiparallel orientation. For three-layer structures, the typical values of giant magnetoresistance at room temperature lie in the range 5–8%.

(iii) Magnetic tunnel structures [11–16]. These structures consist of half-metallic or metallic ferromagnetic layers separated by an insulating layer. Tunneling magnetoresistance at room temperature lies from 20 to 50%.

(iv) Granulated films consisting of nanosized magnetic metal particles in a dielectric matrix [17–22]. The granulated films exhibit both positive and negative magnetoresistance due to spin-dependent tunneling between nanoparticles. The values of the magnetoresistance lie in the range of 3–10%.

We analyzed the effect of the positive injection magnetoresistance in heterostructures of gallium arsenide/granulated SiO_2 film with Co nanoparticles and gallium arsenide/TiO_2 film with Co island sublayers. For the SiO_2(Co)/GaAs structure with 60 at. % Co in a surface-parallel magnetic field of 23 kOe at a voltage of 50 V, the injection magnetoresistance reaches 5200% at room temperature. In this case, the 520-nm SiO_2 film contains 3.5-nm Co particles and is in the region of the percolation threshold. These injection magnetoresistance values are two orders of magnitude larger than the maximum values of the giant magnetoresistance and tunneling magnetoresistance that are observed in layered metallic and magnetic tunneling structures, as well as than the magnetoresistance of granulated structures.

2. Experiment. Experiments were carried out with heterostructures $(SiO_2)_{100-x}Co_x$/GaAs [or shorter SiO_2(Co)/GaAs] consisting of the amorphous SiO_2 film with Co nanoparticles on the GaAs substrate and with heterostructures $(TiO_2)_{100-x}Co_x$/GaAs [or shorter TiO_2(Co)/GaAs] consisting of the amorphous TiO_2 film with Co island sublayers on the GaAs substrate. The 0.4-mm *n*-GaAs substrates had an orientation of (100), a resistivity of 0.9–1.0 Ω cm, and a carrier concentration of 10^{15} cm^{-3}. The electron mobility was in the range $7.0–7.2 \times 10^3$ $cm^2/(s\ V)$. Before the deposition of the films, the substrates were polished by a low energy

0021-3640/05/8110-0514$26.00 © 2005 Pleiades Publishing, Inc.

oxygen ion beam using the technique described in [23, 24]. The roughness of the polished surfaces did not exceed 0.5 nm.

The SiO_2(Co) films were deposited by the ion-beam sputtering method from a cobalt–quartz composite target onto the GaAs substrates heated to 200°C. The concentration of Co nanoparticles in SiO_2 was specified by the ratio of the areas of cobalt and quartz. The composition of the deposited structures was evaluated by x-ray structure analysis. For the SiO_2(Co) structures under investigation, the Co content and thickness were 39 at. % (450 nm), 60 at. % (520 nm), and 85 at. % (600 nm). The mean size of the Co particles was estimated by means of small-angle x-ray scattering and increased with x from 2.7 nm for $x = 39$ at. % to 3.5 nm for $x = 60$ at. %. With these sizes, the Co particles are in the single-domain ferromagnetic state [19, 25, 26]. As the Co content increased, the resistivity of the SiO_2(Co) films decreased from 2.0 Ω cm (39 at. %) through 3.0×10^{-3} Ω cm (60 at. %) to 7.0×10^{-4} Ω cm (85 at. %).

The TiO_2(Co) films were layer-by-layer grown using the ion-beam sputtering method from the separated TiO_2 and Co targets. In contrast with SiO_2(Co) films, the granulated TiO_2(Co) film has the layered island structure. The first layer deposited on the GaAs substrate was the layer of Co islands. Ten layers of Co islands and ten layers of TiO_2 were deposited so that the $(Co/TiO_2)_{10}$ structure was formed. The thicknesses of the island layers were identical for all the samples under investigation and were equal to 2.7 nm. The thicknesses of the TiO_2 layers were different for the samples under investigation and decreased as the Co content increases. The composition of the films was determined by the neutron activation and x-ray analyses. For the samples under investigation, the Co content x and the total thickness of the TiO_2(Co) structure were 34 at. % (45 nm), 55 at. % (40 nm), and 76 at. % (35 nm). For comparison of the characteristics of the heterostructures with granulated films, Co films were also deposited on GaAs substrates.

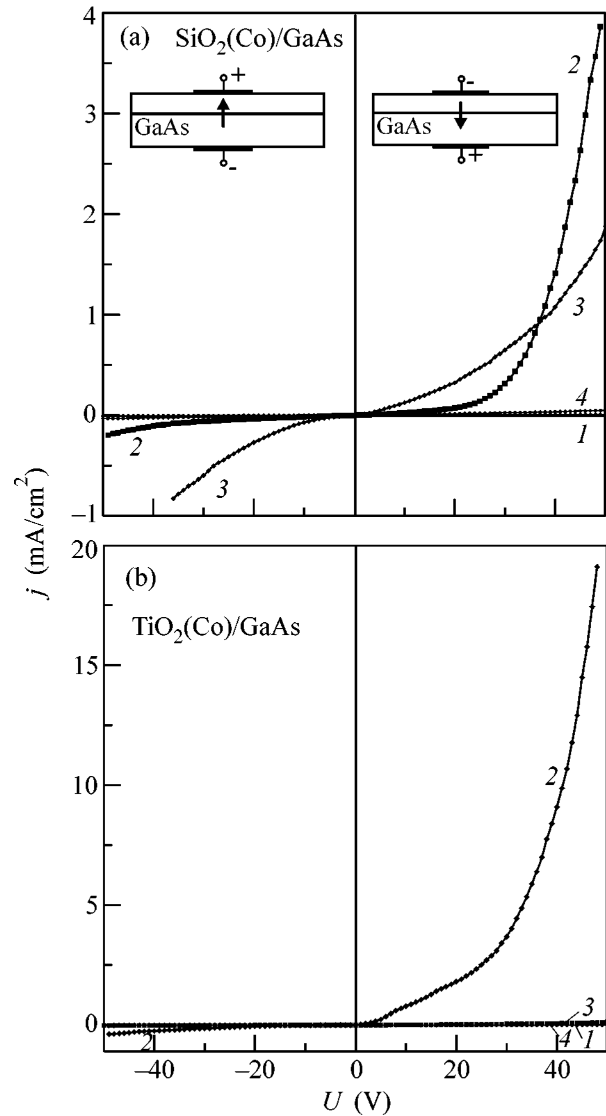

Fig. 1. Current–voltage characteristics for (a) the SiO_2(Co)/GaAs structure with a Co content of (*1*) 39, (*2*) 60, (*3*) 85, and (*4*) 100 at. % and (b) the TiO_2(Co)/GaAs structure with a Co content of (*1*) 34, (*2*) 55, (*3*) 76, and (*4*) 100 at. %.

Figure 1 shows the current–voltage characteristics for the SiO_2(Co)/GaAs and TiO_2(Co)/GaAs structures at room temperature. One contact was on the GaAs substrate, and the other, on the granulated film. The sizes of the samples were equal to $3 \times 3 \times 0.4$ mm. The size of a sample in the current flow direction was equal to 0.4 mm. For structures with low and high Co contents, as well as for the Co/GaAs structure, the current–voltage characteristics are close to the ohm characteristic. On the contrary, the TiO_2(Co)/GaAs samples with the Co content $x = 55$ at. % and the SiO_2(Co)/GaAs samples with $x = 60$ and 85 at. % exhibit the diode-type current–voltage characteristic. For these structures, electrons at the positive applied voltage are injected from the granulated film into the GaAs substrate and the current density j is high. If the applied voltage U is negative, electrons drift from GaAs to the film and the current density is low. We emphasize that, for the samples with a high Co content [$x \geq 60$ at. % for SiO_2(Co)/GaAs and $x \geq 55$ at. % for TiO_2(Co)/GaAs], the resistivity of GaAs is higher than the resistivity of the film and the applied voltage primarily falls on the semiconductor substrate.

Figure 2 illustrates the effect of the magnetic field $H = 23$ kOe on the current–voltage characteristic for the injection of electrons into the semiconductor for the SiO_2(Co)/GaAs structure with 60 at. % Co. The magnetic field is parallel to the film surface. For $U > 20$ V, the magnetic field suppresses the injection and the cur-

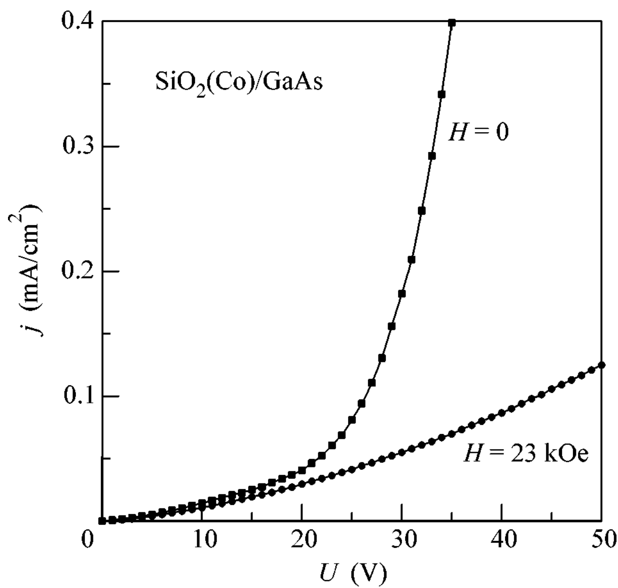

Fig. 2. Current–voltage characteristic for the injection of electrons into the semiconductor for the $SiO_2(Co)/GaAs$ structure with 60 at. % Co in the absence of a magnetic field and in the magnetic field $H = 23$ kOe.

rent–voltage characteristic approaches the ohm characteristic.

Figure 3a shows the relative current density j/j_0, where j_0 is the current density in the absence of a magnetic field, as a function of the magnetic field H for various applied voltages at room temperature for the $SiO_2(Co)/GaAs$ structure with 60 at. % Co. At negative voltages, electrons drift from GaAs to the $SiO_2(Co)$ film and the effect of the magnetic field on the current is insignificant. At positive voltages, electrons are injected into the GaAs substrate and the current decreases significantly as the magnetic field H increases. If the magnetic field is parallel to the film surface, the injection current j approaches a certain limit as the magnetic field H increases; i.e., saturation is observed. If the magnetic field is perpendicular to the film surface, the dependence of the current on the magnetic field H is weaker and saturation is not observed in the field range 0–23 kOe. Figure 3b shows the relative current density j/j_0 as a function of the magnetic field H parallel to the film for various applied voltages in the $TiO_2(Co)/GaAs$ structure with 55 at. % Co. The dependence of the injection current on the magnetic field for the $TiO_2(Co)/GaAs$ structure is weaker than that for the $SiO_2(Co)/GaAs$ structure.

By analogy with the coefficients of giant magnetoresistance and tunneling magnetoresistance in [8–15], we define the injection magnetoresistance coefficient *IMR* as the ratio

$$IMR = \frac{R(H) - R_0}{R_0} = \frac{j_0 - j(H)}{j(H)},$$

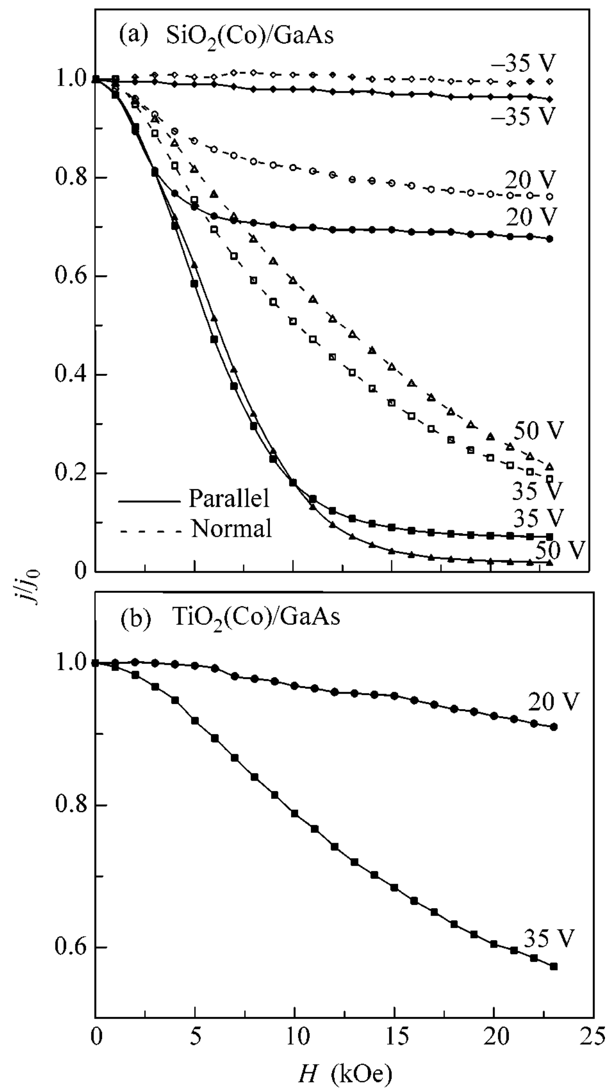

Fig. 3. Relative current density j/j_0 vs. the magnetic field H for various applied voltages for (a) the $SiO_2(Co)/GaAs$ structure with 60 at. % Co in the magnetic fields parallel and perpendicular to the film and (b) the $TiO_2(Co)/GaAs$ structure with 55 at. % Co in the parallel magnetic field.

where R_0 and $R(H)$ are the resistances of the GaAs/granulated film structure in the absence and presence of the magnetic field H, respectively, and $j(H)$ is the injection current density in the magnetic field H. The coefficient *IMR* for the $SiO_2(Co)/GaAs$ and $TiO_2(Co)/GaAs$ structures at the applied voltage $U = 35$ V is shown in Fig. 4 as a function of the magnetic field H parallel to the film. For the $SiO_2(Co)/GaAs$ structure with 60 at. % Co, the saturation of *IMR* is observed for high magnetic fields H. At the voltage $U = 50$ V in the magnetic field $H = 23$ kOe, the coefficient *IMR* for this structure reaches 52 (5200%). We emphasize that the magnetoresistance effect for single GaAs samples is not observed in such magnetic fields.

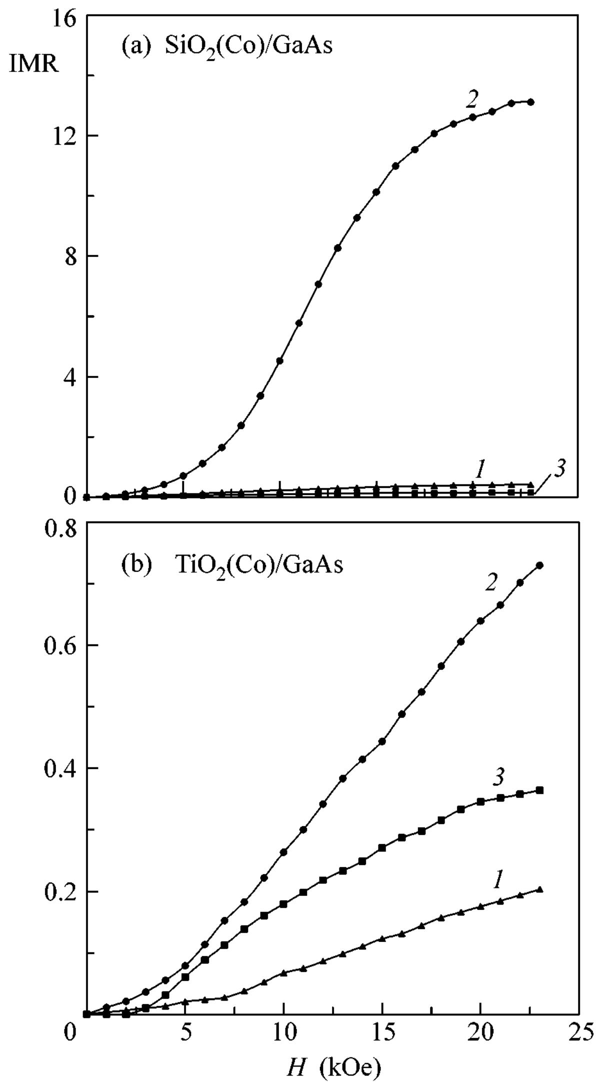

Fig. 4. Injection magnetoresistance coefficient *IMR* vs. the magnetic field H at the voltage U = 35 V for (a) the SiO_2(Co)/GaAs structure with a Co content of (*1*) 39, (*2*) 60, and (*3*) 85 at. % and (b) the TiO_2(Co)/GaAs structure with a Co content of (*1*) 34, (*2*) 55, and (*3*) 76 at. %.

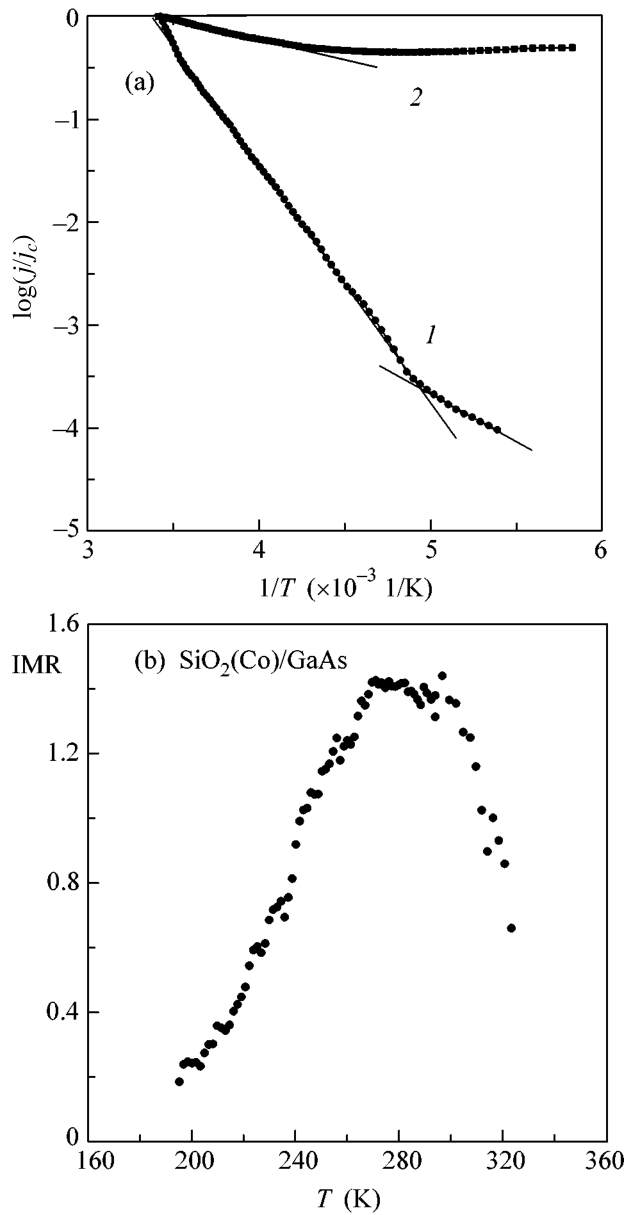

Fig. 5. Temperature dependences of (a) the injection current j for (*1*) the SiO_2(Co)/GaAs structure with 60 at. % Co and (*2*) the TiO_2(Co)/GaAs structure with 55 at. % Co and (b) the injection magnetoresistance coefficient *IMR* for the SiO_2(Co)/GaAs structure with 60 at. % Co at the voltage U = 35 V in the magnetic field H = 2.6 kOe.

The logarithm of the injection current at U = 35 V for the SiO_2(Co)/GaAs structure with 60 at. % Co and for the TiO_2(Co)/GaAs structure with 55 at. % Co is shown in Fig. 5a as a function of the inverse temperature $1/T$. The current density j is normalized to the current density j_c measured at room temperature T = 292 K. It is seen that the injection current for the SiO_2(Co)/GaAs structure is approximated by the law $A_1\exp(-\varepsilon_1/kT) + A_2\exp(-\varepsilon_2/kT)$ with the activation energies ε_1 = 0.47 eV and ε_2 = 0.19 eV. The injection current for the TiO_2(Co)/GaAs structure is described by the law with the activation energy ε = 0.071 eV for temperatures near room temperature. Figure 5b shows the temperature dependence of the coefficient *IMR* for the SiO_2(Co)/GaAs structure with 60 at. % Co at the voltage U = 35 V in the magnetic field H = 2.6 kOe parallel to the film. The coefficient *IMR* has a maximum at T = 280 K.

3. Discussion. The following properties distinguish the injection magnetoresistance effect from the magnetodiode effect [27, 28].

(i) The magnetodiode effect becomes more pronounced as the temperature decreases. On the contrary,

the injection magnetoresistance effect reaches the maximum at a certain temperature and becomes less pronounced as the temperature decreases.

(ii) In contrast to the injection magnetoresistance effect, the magnetodiode effect vanishes in the magnetic field parallel to the current.

(iii) The magnetodiode effect becomes more pronounced as the magnetic field increases, whereas the coefficient *IMR* is saturated in high magnetic fields.

(iv) The magnetodiode effect depends on the geometry and thickness of a semiconductor. On the contrary, the injection magnetoresistance effect depends strongly on the properties of the granulated film deposited on the semiconductor. In particular, the coefficients *IMR* for the SiO_2(Co)/GaAs structures with 60 and 85 at. % Co differ from each other by two orders of magnitude at the same injection current.

Taking into account the above discussion and the significantly lower magnetoresistance of granulated films, one can assume that the injection magnetoresistance effect is governed by the spin-dependent scattering of injected spin-polarized electrons on a potential barrier formed near the GaAs/granulated film interface.

4. Conclusions. A new phenomenon, injection magnetoresistance, has been investigated in the gallium arsenide/granulated SiO_2 film heterostructure with Co nanoparticles and in the gallium arsenide/TiO_2 film heterostructure with Co island sublayers. For the SiO_2(Co)/GaAs structure with 60 at. % Co in a surface-parallel magnetic field of 23 kOe at a voltage of 50 V, the coefficient of injection magnetoresistance reaches 5200% at room temperature. This value is two orders of magnitude larger than the maximum values of the giant magnetoresistance and tunneling magnetoresistance that are observed in layered metallic and magnetic tunneling structures.

We are grateful to R.V. Pisarev, V.N. Gridnev, V.V. Pavlov, P.A. Usachev, and A.M. Kalashnikova for stimulating discussions, L.A. Kalyuzhnaya for performing x-ray spectroscopic analysis, and V.I. Siklitskiĭ and M.V. Baĭdakova for the measurements by the small-angle x-ray scattering method.

REFERENCES

1. S. A. Wolf, D. D. Awschalom, R. A. Buhrman, *et al.*, Science **294**, 1488 (2001).
2. G. Schmidt, G. Richter, P. Grabs, *et al.*, Phys. Rev. Lett. **87**, 227 203 (2001).
3. Y. Ohno, D. K. Yong, B. Beschoten, *et al.*, Nature **402**, 790 (1999).
4. B. T. Jonker, Y. D. Park, B. R. Bennett, *et al.*, Phys. Rev. B **62**, 8180 (2000).
5. Z. H. Xiong, Di Wu, Z. V. Vardeny, *et al.*, Nature **427**, 821 (2004).
6. A. T. Hanbicki, B. T. Jonker, G. Istkos, *et al.*, Appl. Phys. Lett. **80**, 1240 (2002).
7. A. Hirohata, Y. B. Xu, C. M. Guertler, *et al.*, J. Appl. Phys. **87**, 4670 (2000).
8. M. N. Baibich, J. M. Broto, A. Fert, *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).
9. J. Bass and W. P. Pratt, Jr., J. Magn. Magn. Mater. **200**, 274 (1999).
10. M. A. M. Gijs and G. E. W. Bauer, Adv. Phys. **46**, 285 (1997).
11. J. S. Moodera, L. R. Kinder, T. M. Wong, *et al.*, Phys. Rev. Lett. **74**, 3273 (1995).
12. J. S. Moodera and G. Mathon, J. Magn. Magn. Mater. **200**, 248 (1999).
13. Xiu-Feng Han, M. Oogane, H. Kubota, *et al.*, Appl. Phys. Lett. **77**, 283 (2000).
14. E. Y. Tsymbal, O. N. Mryasov, and P. R. LeClair, J. Phys.: Condens. Matter **15**, R109 (2003).
15. J. M. De Teresa, A. Barthelemy, A. Fert, *et al.*, Science **286**, 507 (1999).
16. X. Jiang, R. Wang, R. Shelby, *et al.*, Phys. Rev. Lett. **94**, 056 601 (2005).
17. B. Dieny, S. Sankar, M. R. McCartney, *et al.*, J. Magn. Magn. Mater. **185**, 283 (1998).
18. S. Mitani, K. Takanashi, K. Yakushiji, *et al.*, J. Appl. Phys. **83**, 6524 (1998).
19. K. Yakushiji, S. Mitani, K. Takanashi, *et al.*, J. Magn. Magn. Mater. **212**, 75 (2000).
20. L. V. Lutsev, Yu. E. Kalinin, A. V. Sitnikov, and O. V. Stogneĭ, Fiz. Tverd. Tela (St. Petersburg) **44**, 1802 (2002) [Phys. Solid State **44**, 1889 (2002)].
21. B. A. Aronzon, A. E. Varfolomeev, D. Yu. Kovalev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 944 (1999) [Phys. Solid State **41**, 857 (1999)].
22. A. E. Varfolomeev and M. V. Sedova, Fiz. Tverd. Tela (St. Petersburg) **45**, 500 (2003) [Phys. Solid State **45**, 529 (2003)].
23. A. I. Stogniĭ, N. N. Novitskiĭ, and O. M. Stukalov, Pis'ma Zh. Tekh. Fiz. **28** (1), 39 (2002) [Tech. Phys. Lett. **28**, 17 (2002)].
24. A. I. Stogniĭ, N. N. Novitskiĭ, and O. M. Stukalov, Pis'ma Zh. Tekh. Fiz. **29** (2), 6 (2003) [Tech. Phys. Lett. **29**, 43 (2003)].
25. S. Barzilai, Y. Goldstein, I. Balberg, *et al.*, Phys. Rev. B **23**, 1809 (1981).
26. S. Sankar, D. Dender, J. A. Borchers, *et al.*, J. Magn. Magn. Mater. **221**, 1 (2000).
27. V. I. Stafeev and É. I. Karakushan, *Magnetodiodes: New Semiconductor Devices with High Sensibility to Magnetic Field* (Nauka, Moscow, 1975) [in Russian].
28. A. A. Abramov and I. N. Gorbatyĭ, Fiz. Tekh. Poluprovodn. (St. Petersburg) **36**, 847 (2002) [Semiconductors **36**, 793 (2002)].

Translated by R. Tyapaev

JETP Letters, Vol. 81, No. 10, 2005, pp. 519–522. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 642–645.
Original Russian Text Copyright © 2005 by Benemanskaya, Frank-Kamenetskaya.

Charge Accumulation Layer on the *n*-GaN (0001) Surface with Ultrathin Ba Coatings

G. V. Benemanskaya and G. É. Frank-Kamenetskaya

Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia

Received March 16, 2005; in final form, April 21, 2005

The adsorption of Ba on the *n*-type GaN(0001) surface is studied. It is found that submonolayer Ba coatings induce cardinal changes in the electronic properties of the surface with the formation of a charge accumulation layer in the region of the near-surface band bending. The excitation of the Ba/*n*-GaN system by light from the region of GaN transparency results in photoemission. The lowest value of the work function corresponds to ~1.90 eV at a Ba coverage of ~0.4 ML. Two surface bands induced by Ba adsorption are found in the surface photoemission spectra. © *2005 Pleiades Publishing, Inc.*

PACS numbers: 73.20.–r, 79.60.Dp

A powerful advance in the technology of group III nitrides, as well as in the development of devices on their basis, has been observed in recent years. At the same time, under the pressure of the necessity of obtaining a rapid practical result, basic studies of group III nitrides and, in particular, studies of phenomena on the surface of epitaxial layers of group III nitrides have been insufficient. The recent works have demonstrated that the surface makes a significant contribution to many processes. For example, such a detrimental effect as the current collapse of field-effect transistors depends on the quality of the surface of group III nitride compounds. So far, it is unclear whether the occurrence of a negative affinity at the surface of *p*-type GaN and *p*-type AlN after cesium deposition is a typical property. The properties of *n*-type surfaces are very poorly understood at all, though these surfaces are an integral part of devices based on these materials.

Studying the electronic properties of ultrathin interfaces forming upon the adsorption of metals on the GaN surface is of both fundamental and applied importance [1–3]. The adsorption of Ba on the GaN surface has not yet been studied. The adsorption of Cs was studied by UV photoemission spectroscopy for *p*-type GaN(0001) [4, 5] and for *n*-type GaN(0001) at a temperature of 150 K [6]. We have recently studied the Cs/*n*-GaN(0001) by threshold photoemission spectroscopy at a temperature of 300 K [7, 8]. It was found that the work function $\varphi \sim 4.3$ eV for the pure *n*-GaN(0001) surface decreases to $\varphi \sim 1.35$ eV after the adsorption of Cs. It should be noted that such a sharp decrease in the work function upon Cs adsorption is observed only at the *n*-type GaN(0001) surface, whereas the lowest value of the work function at the *p*-type GaN(0001) surface equals $\varphi \sim 3.5$ eV [5]. Photoemission with a high quantum yield was detected for the Cs/*n*-GaN(0001) system upon light excitation from the transparency region of GaN. It was found that Cs adsorption induced the formation of a charge accumulation layer (AL) in the near-surface region of GaN. The nature of the photoemission was related to the excitation of electrons from the AL, which formed because of the near-surface bending of the conduction band. This is a fundamental difference from all the known photocathodes in which photoemission is due to the excitation of electrons from the valence band.

In this connection, it is of particular interest to study the electronic properties of the *n*-GaN surface upon Ba adsorption and to reveal the specific features of the formation of the Ba/*n*-GaN interface as compared to Cs/*n*-GaN. In this work, the Ba/*n*-GaN(0001) system with various submonolayer Ba coatings is studied using threshold photoemission spectroscopy. It is found that the adsorption of Ba induces a number of qualitative changes in the electronic properties of the near-surface region with the formation of an AL similar to the adsorption of Cs. Moreover, the formation of induced surface bands has been detected upon Ba adsorption.

Threshold photoemission spectroscopy involves *s*- and *p*-polarized excitations and has been successfully used previously for studying the adsorption of metals on the surface of Si(111) and Si(100) semiconductors [9] and on the GaAs(100) surface [10]. This method is based both on the separation of bulk and surface photoemissions and on the near-threshold enhancement of photoemission from surface states in the case of *p* polarization. For *s* polarization, bulk states are excited. In this case, photoemission $I_s(h\nu)$ occurs from the near-surface region, which is determined by the escape depth of 20–30 nm for low-energy electrons. Conventional quasi-bulk photoemission is associated with the excitation of valence-band electrons; hence, the photoemission threshold $h\nu_s$ for *s*-polarized light corresponds to the ionization energy, which equals the

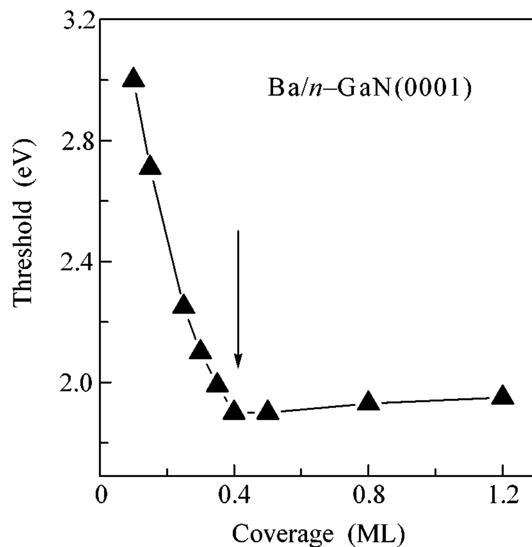

Fig. 1. Photoemission thresholds $h\nu_s = h\nu_p = \varphi$ vs. the Ba coverage for the n-GaN(0001) surface.

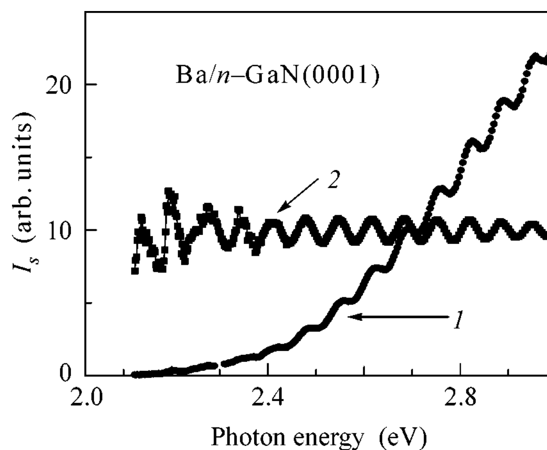

Fig. 2. (*1*) Photoemission spectrum $I_s(h\nu)$ for the Ba/n-GaN(0001) system at a Ba coverage of 0.4 ML and (*2*) extracted modulation intensity spectrum.

position of the edge E_V at the surface with respect to the vacuum level E_{VAC}. In the unconventional case when an AL occurs in the conduction band, which was found for the Cs/n-GaN system, the photoemission threshold $h\nu_s$ corresponds to the work function. The normal component of p-polarized light can excite surface states localized at a depth of ~0.5 nm. Such states are investigated by surface photoemission spectra $I_p(h\nu)/I_s(h\nu)$ for p- and s-polarized excitations, respectively. A detailed description of the threshold photoemission spectroscopy technique is given in [9–11].

The experiment was performed *in situ* in a vacuum of $P < 1 \times 10^{-10}$ Torr at room temperature. The n-type sample doped with silicon (2×10^{17} cm^{-3}) ~ 4 μm thick represented an epitaxial GaN(0001) layer grown on a sapphire substrate by MOCVD epitaxy. The sample was annealed directly in a vacuum at a temperature of ~750°C. Atomic pure Ba was deposited on the sample surface from a standard source. The Ba coverages were varied in the range 0.1–2 ML and were determined *in situ* by the procedure in [8]. Note that 1 ML corresponds to a concentration equal to 9.89×10^{14} atom/cm^2, at which the GaN(0001) 1 × 1 surface contains one metal adatom per gallium atom. Photoemission was excited by s- or p-polarized monochromatic light incident on the sample at an angle of 45°.

It was found that Ba adsorption on the n-GaN(0001) surface results in a sharp decrease in the photoemission thresholds $h\nu_s$ and $h\nu_p$ for s- or p-polarized excitation, respectively. The equality of thresholds $h\nu_s = h\nu_p$ was found in the entire range of coverages. The variation of thresholds as functions of the Ba coverage is shown in Fig. 1. Points in the curve in Fig. 1 were obtained from an approximation of the spectral dependences of photoemission $I_s(h\nu)$ and $I_p(h\nu)$, which were measured for each Ba coverage. It was found that the minimum value of the threshold energies corresponds to (1.90 ± 0.05) eV at a surface Ba concentration of ~4×10^{14} atom/cm^2. The determination of Ba coverages was carried out with regard to the sticking coefficient of Ba adatoms, which equals one and does not vary at least within a monolayer. At the minimum value of the thresholds, the Ba coverage lies in a submonolayer range and corresponds to ~0.4 ML. We also note that Ba coatings are stable at room temperature and are retained for many hours.

Studying the character of the spectral dependences of bulk photoemission $I_s(h\nu)$ at various Ba coverages showed that the spectra near the threshold obey the Fowler law $I_s(h\nu) \sim (h\nu - h\nu_s)^2$. It is well known that the law is valid for the photoemission from metal bulk. For semiconductors, this law is obeyed, for example, at the photoemission $I_p(h\nu)$ from quasi-metallic surface bands lying in the band gap at the Fermi level [9]. In our case of the Ba/n-GaN interface, the Fowler law was valid for bulk rather than surface photoemission $I_s(h\nu)$. Note that GaN belongs to wide band-gap semiconductors with a band gap of 3.4 eV. Photoemission under excitation in the GaN transparency region cannot be caused by the excitation of valence-band states. Hence, the detected photoemission arises from quasi-metallic states in the conduction band. In this case, the band edge E_c at the surface lies below the Fermi level, which corresponds to the formation of a charge accumulation layer (see Fig. 4 below). Hence, the bulk photoemission threshold $h\nu_s$ corresponds to the work function $h\nu_s = \varphi$.

Thus, it was found that the nature of the threshold photoemission for Ba/n-GaN, as well as for Cs/n-GaN, is associated with the formation of a charge AL in the near-surface band-bending region. A comparison of the $I_s(h\nu)$ spectra for the Ba/n-GaN and Cs/n-GaN systems with the use of the same GaN sample as a substrate

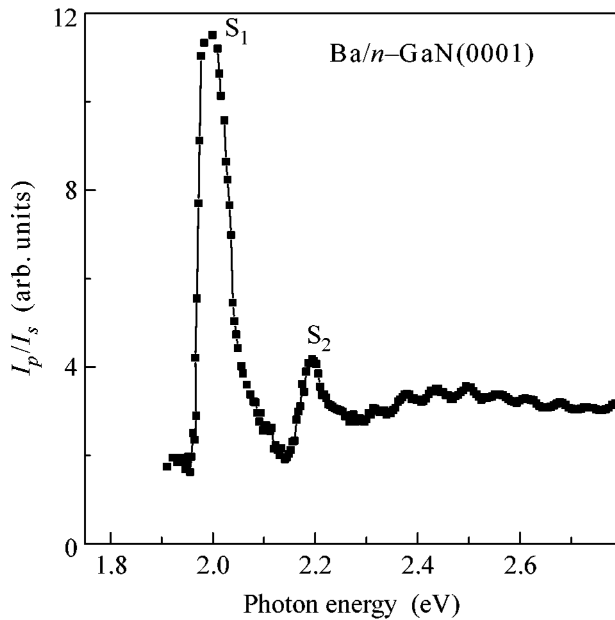

Fig. 3. Surface photoemission spectrum I_p/I_s for a Ba coverage of 0.5 ML.

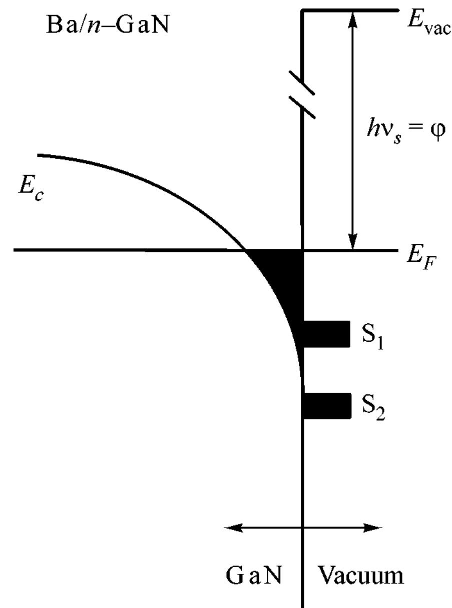

Fig. 4. Schematic diagram of the charge accumulation layer and surface states S_1 and S_2 induced by Ba adsorption at a coverage of 0.5 ML on the *n*-GaN(0001) surface.

shows that the energy parameters of the AL in the case of Ba adsorption are at least half as much as for Cs adsorption.

The threshold photoemission spectrum $I_s(h\nu)$ at a Ba coverage of 0.4 ML is shown in Fig. 2, curve *1*. The character of the spectrum appears to be extremely unusual because of the occurrence of pronounced oscillations in the photocurrent. It is found that the oscillation period in this spectrum is constant with respect to the energy $\Delta \sim 0.07$ eV and is retained in the spectra recorded for other Ba coverages. This effect was first observed for the Cs/*n*-GaN(0001) system in our works [7, 8]. The interference of light from the transparency region that occurs in the plane-parallel plate of the GaN sample should be primarily considered as the mechanism of the occurrence of oscillations in the photoemission spectra. In this case, photoemission from the subsurface accumulation layer can be excited by light quanta with an energy higher than the work function $h\nu \geq \varphi$. The excitation of AL electrons in the plane-parallel plate of the GaN sample with a surface AL can be carried out by light that falls both on the vacuum side and multiply on the side of the plane-parallel plate of the GaN sample. Thus, we observe an unconventional situation of the mutual effect of interference and photoemission. The modulation spectrum was extracted from the experimental spectrum $I_s(h\nu)$ using a computer program for Fabry–Perot interference (Fig. 2, curve *2*). The calculations were performed with regard to the refractive index 2.32 in the optical excitation region of GaN. Optimizing the calculation parameters shows that the exact oscillation period $\Delta = 0.69$ eV corresponds to a thickness of 3.83 μm for a plane-parallel plate. The calculated value coincides well with the data for the size of the epitaxial layer for the GaN sample (~4 μm) obtained in the course of growth.

During the formation of the metal–semiconductor interface, the redistribution of the electron density in the bond region is mainly due to the valence electrons of the adatom and the surface states of the substrate. Therefore, the main changes in the 2D electronic structure during the formation of the interface should be expected at the Fermi level. We studied the surface photoemission spectra $I_p(h\nu)/I_s(h\nu)$ for the Ba/*n*-GaN(0001) interface. It was found that two filled surface bands S_1 and S_2 form successively in the spectra starting with a Ba coverage of ~0.3 ML. The surface photoemission spectrum I_p/I_s is presented in Fig. 3 for a coverage of 0.5 ML. The formation of the surface bands induced by Ba adsorption ends at 0.5–0.6 ML, and, further, up to the Ba monolayer coverage, the electronic spectrum of surface states remains virtually unchanged. It is determined that the S_1 and S_2 bands are located below the Fermi level by ~0.15 and ~0.30 eV, respectively. There are no surface states in this energy range at the pure GaN(0001) surface [4, 5]. It should also be noted that induced surface states have not been observed previously for various interfaces of the metal/*n*-GaN and metal/*p*-GaN type [4–8].

The nature of the detected 2D electronic states is evidently associated with the local interaction of Ba adatoms and dangled bonds of surface Ga atoms. Ba atoms have two valence electrons with a complicated *s*–*p*–*d* structure of low-lying electronic states. Experimental and theoretical studies performed for the

Ba/W(110) system showed that at least two surface bands exist in the vicinity of E_F [12]. Therefore, the occurrence of several 2D bands in the process of Ba/*n*-GaN interface formation can be explained by the participation of various *sp*, *ds*, and *dp* hybridized states. The existence of 2D states in the vicinity of E_F points to the formation of a strong covalent bond between adsorbed Ba atoms and the GaN surface.

The electronic structure of the near-surface region of the *n*-GaN(0001) substrate and the Ba/*n*-GaN(0001) interface is schematically shown in Fig. 4, which summarizes the results obtained in this work. The formation of the charge accumulation layer in the conduction band is due to the surface band bending induced by Ba adsorption. The formation of the interface also initiates the formation of 2D electronic bands directly at the surface, namely, the induced S_1 and S_2 surface states.

We are grateful to V.S. Vikhnin and V.Yu. Davydov for fruitful discussions and to S.N. Timoshnev for help in the experimental work. This work was supported by the Russian Foundation for Basic Research, project no. 04-02-17621, and by the Presidium of the Russian Academy of Sciences.

REFERENCES

1. M. Razeghi and A. Rogalski, J. Appl. Phys. **79**, 7433 (1996).
2. S. J. Pearton, J. C. Zolper, R. J. Shul, *et al.*, J. Appl. Phys. **86**, 1 (1999).
3. F. Machuca, Y. Sun, Z. Liu, *et al.*, J. Vac. Sci. Technol. B **18**, 3042 (2000).
4. M. Eyckeler, W. Mönch, T. U. Kampen, *et al.*, J. Vac. Sci. Technol. B **16**, 2224 (1998).
5. C. I. Wu and A. Kahn, J. Appl. Phys. **86**, 3209 (1999); Appl. Surf. Sci. **162–163**, 250 (2000).
6. T. U. Kampen, M. Eyckeler, and W. Mönch, Appl. Surf. Sci. **123–124**, 28 (1998).
7. I. V. Afanas'ev, G. V. Benemanskaya, V. S. Vikhnin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 270 (2003) [JETP Lett. **77**, 226 (2003)].
8. G. V. Benemanskaya, V. S. Vikhnin, G. E. Frank-Kamenetskaya, *et al.*, Appl. Phys. Lett. **85**, 1365 (2004).
9. G. V. Benemanskaya, D. V. Daineka, and G. E. Frank-Kamenetskaya, J. Phys.: Condens. Matter **11**, 6679 (1999).
10. G. V. Benemanskaya, D. V. Daĭneka, and G. É. Frank-Kamenetskaya, Zh. Éksp. Teor. Fiz. **119**, 342 (2001) [JETP **92**, 297 (2001)].
11. A. Liebsch, G. V. Benemanskaya, and M. N. Lapushkin, Surf. Sci. **302**, 303 (1994).
12. G. V. Benemanskaya, O. P. Burmistrowa, and M. N. Lapushkin, Phys. Lett. A **137**, 139 (1989).

Translated by A. Bagatur'yants

JETP Letters, Vol. 81, No. 10, 2005, pp. 523–526. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 646–649.
Original Russian Text Copyright © 2005 by Bykov, Kalagin, Bakarov.

Magnetophonon Resonance in a GaAs Quantum Well with AlAs/GaAs Superlattice Barriers at High Filling Factors

A. A. Bykov, A. K. Kalagin, and A. K. Bakarov

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia

e-mail: bykov@thermo.isp.nsc.ru

Received April 21, 2005

The magnetotransport of a high-mobility 2D electron gas in single GaAs quantum wells with AlAs/GaAs superlattice barriers is studied at high filling factors. For the selectively doped structures under study in the temperature range from 10 to 25 K, magnetoresistance oscillations periodic in the inverse magnetic field are observed with their frequency being proportional to the Fermi wave vector of the 2D electron gas. The experimental results are explained by the interaction of the 2D electron gas with leaky interface acoustic phonons. *© 2005 Pleiades Publishing, Inc.*

PACS numbers: 73.23.–b, 73.40.Gk

Magnetophonon resonance caused by the resonant absorption and emission of optical phonons was predicted in [1]. By now, this phenomenon has been studied in sufficient detail in bulk semiconductors [2–4] and in 2D electron systems formed on the basis of selectively doped semiconductor structures [5–7]. The resonant electron–phonon interaction (scattering) occurs in a strong magnetic field when the following conditions are satisfied: $\omega_{LO} = j\omega_c = jeB/m^*$, where ω_{LO} is the frequency of a longitudinal optical phonon, ω_c is the cyclotron frequency, and j is a positive integer number. In magnetoresistance, the magnetophonon resonance manifests itself as a series of oscillations exponentially decreasing in amplitude [4]:

$$\Delta R_{xx}/R_{xx} \sim \cos(2\pi\omega_{LO}/\omega_c)\exp(-\gamma\omega_{LO}/\omega_c), \quad (1)$$

where γ is the damping factor. These oscillations, as the Shubnikov–de Haas oscillations, are periodic in $1/B$, but their period, in contrast to the latter oscillations, does not depend on the charge carrier concentration.

In selectively doped GaAs/AlGaAs heterojunctions and in GaAs quantum wells, the magnetophonon resonance caused by the interaction of the 2D electron gas with longitudinal optical phonons is only observed at relatively high temperatures $T = 100$–180 K [5]. This is explained by the fact that the number of phonons depends on temperature according to the exponential law $n_{ph} \sim \exp(-\hbar\omega_{LO}/k_BT)$. Therefore, the magnetophonon resonance does not manifest itself at low temperatures, when $\hbar\omega_{LO} \gg k_BT$. However, because of the inevitable decrease in the mobility of the 2D electron gas with increasing temperature, the magnetophonon resonance also does not occur at excessively high temperatures, because it is necessary that the strong magnetic field condition $\omega_c\tau_{tr} > 1$ be satisfied for the charge carriers. These boundary conditions determine the temperature range within which the magnetophonon resonance manifests itself.

Until recently, it was believed that the interaction of a 2D electron gas with acoustic phonons cannot be resonant. However, this generally accepted point of view was disproved in [8], which resulted in the experimental observation of a new class of magnetic-field oscillations arising in the magnetoresistance of a high-mobility 2D electron gas in GaAs/AlGaAs heterojunctions under the effect of microwave radiation [9]. It was found that, at large filling factors, i.e., when the inequality $E_F/\hbar\omega_c \gg 1$ (E_F is the electron energy at the Fermi level) is satisfied, the selection rules realized in the system are such that they allow the manifestation of the magnetophonon resonance due to the interaction of the 2D electron gas with acoustic phonons.

The essence of these selection rules consists in that, at high filling factors, the transition of an electron from one Landau level to another should be accompanied by a change in the electron momentum according to the condition $\Delta k_x \sim 2k_F$, where the x axis coincides with the direction of the bias current and k_F is the Fermi wave vector. Such selection rules in the momentum space correspond to the condition $\Delta Y \sim 2R_c$ in the real space, where ΔY is the displacement of the center of the orbit and R_c is the cyclotron radius. Precisely this modulation of scattering in the momentum space leads to the new class of oscillations in the magnetoresistance of the 2D electron gas in strong magnetic fields at high filling factors [8–10].

In this paper, we report on the study of the magnetoresistance of a high-mobility 2D electron gas in GaAs quantum wells with AlAs/GaAs superlattice barriers in the temperature range from 4.2 to 25 K. We have found that, in strong magnetic fields, as the temperature

0021-3640/05/8110-0523$26.00 © 2005 Pleiades Publishing, Inc.

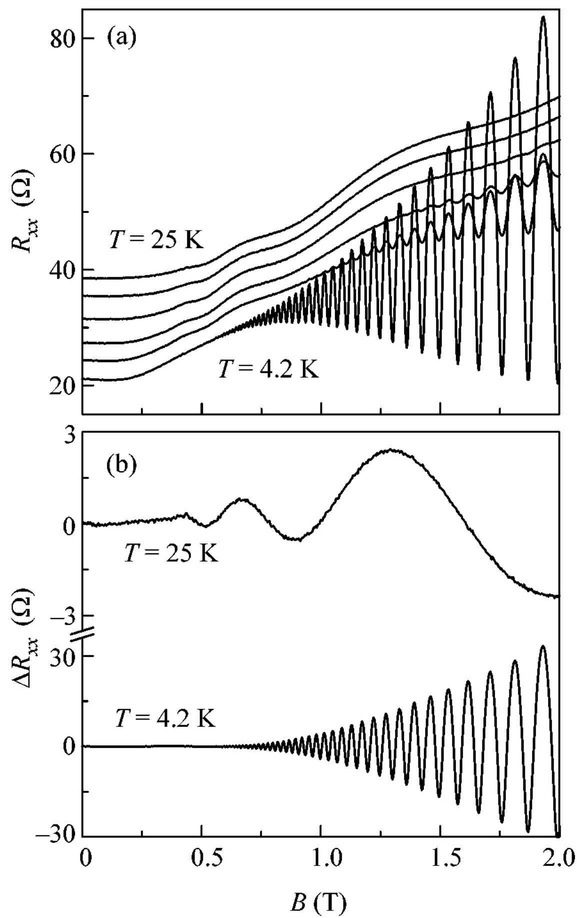

Fig. 1. (a) Experimental dependences $R_{xx}(B)$ in the temperature range from 4.2 to 25 K for a sample with $n = 1.4 \times 10^{12}$ cm^{-2} and (b) dependences $\Delta R_{xx}(B)$ for temperatures of 4.2 and 25 K.

increases, the magnetoresistance of the selectively doped structures under study exhibits oscillations periodic in $1/B$. The frequency of these oscillations proved to be proportional to the Fermi wave vector, and their amplitude reached its maximum value in the temperature range from 10 to 25 K.

The structures under study were selectively doped single GaAs quantum wells with AlAs/GaAs superlattice barriers [11, 12]. They were grown by molecular beam epitaxy (MBE) on (100)GaAs substrates. The width of a GaAs quantum well was 13 nm. The source of 2D electrons were δ-doped silicon layers formed in the superlattice on both sides of the quantum well. The MBE structures under study only differed in the silicon concentration in the δ-doped layers. The mobility and concentration of the 2D electron gas in the initial MBE structures at liquid helium temperature were $\mu = (0.5–1) \times 10^6$ cm^2/V s and $n = (0.9–1.4) \times 10^{12}$ cm^{-2}, respectively. We studied the magnetoresistance for two groups of samples that differed in mobility and in the concentration of the 2D electron and were made on the basis of two different MBE structures. Within each of the groups, the behavior of the magnetoresistance was virtually the same for different samples. The measurements were performed in magnetic fields up to 2 T on Hall bars with a width of 50 μm and a distance of 200 μm between the potential contact pads. The resistance was measured with an ac current of 10^{-6} A in the temperature range from 4.2 to 25 K.

Figure 1a shows the typical field dependences of the magnetoresistance of 2D electron gas in the temperature range from 4.2 to 25 K for the samples that were made on the basis of the structure with the higher mobility. The samples of this group also had a higher concentration of 2D electron gas, as compared to the other group. One can see that, as the temperature increases, the amplitude of the Shubnikov–de Haas oscillations decreases, the resistance at zero magnetic field increases, and a fluctuating component appears on the background of positive magnetoresistance. Figure 1b shows the fluctuating components of magnetoresistance of 2D electron gas at $T = 4.2$ and 25 K. The curves $\Delta R_{xx}(B)$ represent the difference between the experimental dependences and the averaged components.

An analysis of the dependence $\Delta R_{xx}(B)$ (Fig. 1b) obtained at $T = 4.2$ K showed that, in this case, the fluctuating component only contained the Shubnikov–de Haas oscillations. Within the experimental error, the period of these oscillations in an inverse magnetic field corresponded to the concentration of the 2D electron gas that was determined from the dependences $R_{xy}(B)$. The analysis of the dependences $R_{xy}(B)$ showed that, within the experimental error, this concentration did not depend on T in the whole temperature range under study. At $T = 25$ K, the dependence $\Delta R_{xx}(B)$ no longer exhibits oscillations with a period corresponding to the Hall concentration of the 2D electron gas. At this temperature, oscillations that, according to the analysis, are also periodic in $1/B$ but with a greater period are observed. As it was noted in [13], one of the possible sources of these oscillations is the filling of the second dimensional quantization level. However, a numerical Fourier analysis of the dependences $\Delta R_{xx}(B)$ showed that, at $T = 4.2$ K, only one periodic component manifests itself and this component corresponds to the Hall concentration of the 2D electron gas. This fact makes it impossible to explain the presence of the component periodic in $1/B$ at $T = 25$ K by the filling of the second dimensional quantization level.

Figure 2a shows the typical dependences of the magnetoresistance of the 2D electron gas in the temperature range from 4.2 to 25 K for the samples made on the basis of the MBE structure with the lower concentration. The samples of this group were characterized by a lower mobility, as compared to the samples from the other group. For this series of samples, the behavior of $R_{xx}(B)$ with increasing temperature is similar to that shown in Fig. 1a for the higher-mobility samples. Figure 2b represents the fluctuating components of the

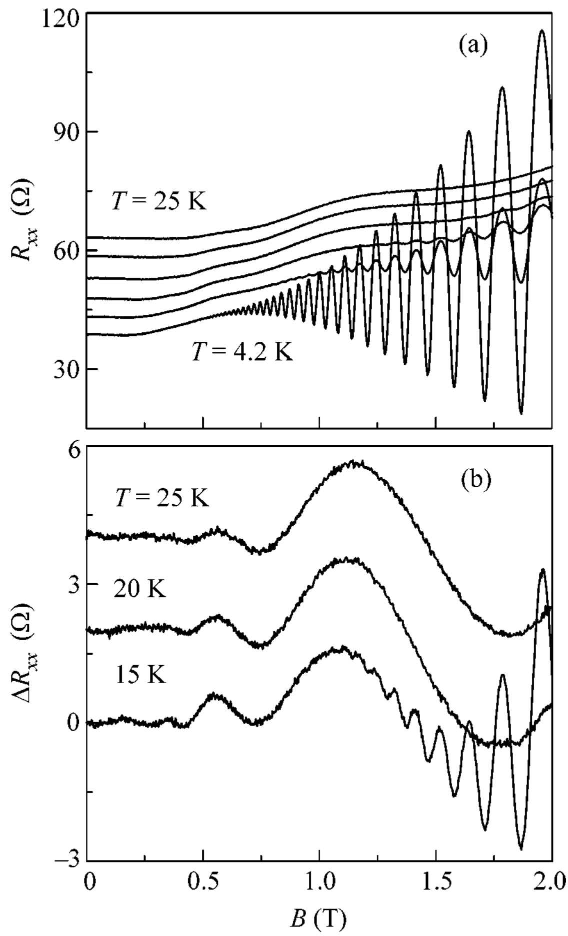

Fig. 2. (a) Experimental dependences $R_{xx}(B)$ in the temperature range from 4.2 to 25 K for a sample with $n = 1.0 \times 10^{12}$ cm^{-2} and (b) dependences $\Delta R_{xx}(B)$ for various temperatures. The curves for 20 and 25 K are shifted upwards along the y axis for clarity.

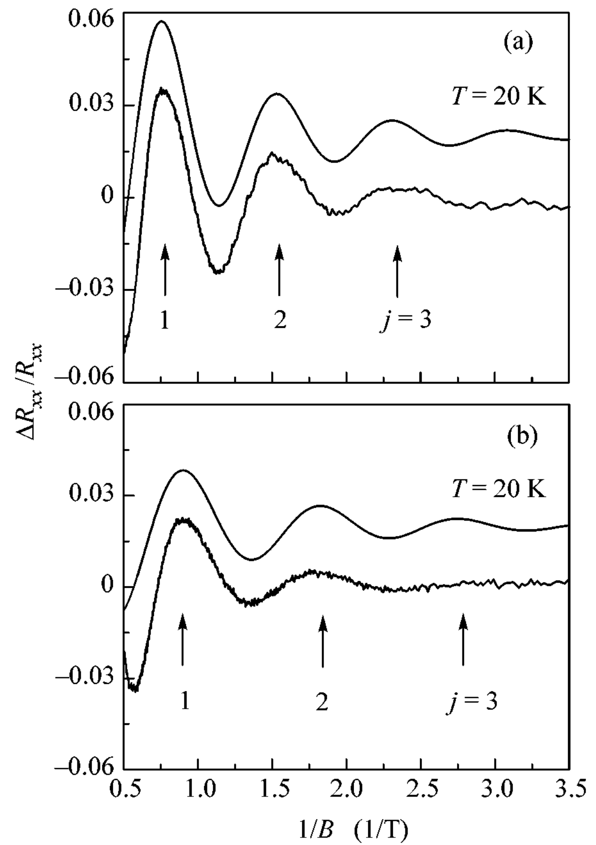

Fig. 3. Plots of $\Delta R_{xx}/R_{xx}$ vs. $1/B$ at T = 20 K for n = (a) 1.4×10^{12} and (b) 1.0×10^{12} cm^{-2}. The thick lines represent the experiment, and the thin lines, the calculation by Eq. (2). The curves shown by the thin lines are shifted upwards along the y axis for clarity. Numbers 1–3 indicate the numbers of the oscillation maxima.

magnetoresistance of 2D electron gas for different temperatures. For this series of samples, the analysis of the periodic component of $\Delta R_{xx}(B)$ showed that its amplitude is maximum at $T \sim 20$ K.

The facts that the oscillations observed at $T > 4.2$ K are periodic in the inverse magnetic field and that their amplitude depends on temperature suggest that these oscillations are caused by the magnetophonon resonance. However, an unusual feature of the magnetophonon resonance observed in the selectively doped structures under study is the manifestation of this resonance at temperatures much lower than 100 K [14]. This experimental fact rules out the interaction of the 2D electron gas with longitudinal optical phonons in the GaAs and AlAs layers. Moreover, this behavior cannot be explained by the convolution of the phonon spectrum in the superlattices lying on both sides of the GaAs quantum well, because, in this case, only the spectrum of the transverse phonons is transformed while these phonons cannot manifest themselves in the interaction with the 2D electron gas.

Figure 3 displays the dependences of $\Delta R_{xx}/R_{xx}$ on $1/B$. One can see that the period of the observed oscillations depends on the concentration of the 2D electron gas. This fact also points to a fundamental difference between the magnetophonon resonance observed by us and the one that was observed at higher temperatures in a similar 2D electron system [14]. An analysis of the dependence of the magnetoresistance oscillation frequency on the concentration of the 2D electron gas at temperatures of 10–20 K showed that the frequency is proportional to k_F. This result allows us to assume that, as in [8–10], the oscillations observed by us in the magnetoresistance of 2D electron gas in GaAs quantum wells with AlAs/GaAs superlattice barriers are caused by the electron transitions from one Landau level to another with a change of momentum by $\Delta k_x \sim 2k_F$. In this case, the magnetophonon resonance manifesting itself in the GaAs wells with AlAs/GaAs superlattice barriers in the temperature range from 10 to 25 K can qualitatively be explained by the resonant absorption

and emission of leaky interface acoustic phonons [15] with a wave vector $q \sim 2k_F$ [8–10].

The conditions of absorption and emission of interface acoustic waves are determined by the relation $2k_F u = j\omega_c$, where $u \approx \omega/q$ is the propagation velocity of leaky interface acoustic waves. Knowing the period of oscillations caused by the magnetophonon resonance in GaAs wells with AlAs/GaAs superlattice barriers, we estimated the quantity u, which proved to be ~5.9 km/s. This value slightly differs from the velocity of leaky interface acoustic waves along the GaAs/AlGaAs heterojunction [15]. In addition, unlike [8, 15], only one interface mode manifests itself in our samples rather than two modes; i.e., the magnetoresistance at T = 10–25 K exhibits oscillations with only one frequency. We explain this distinction by the difference between the layered structure of GaAs quantum wells with AlAs/GaAs superlattice barriers studied by us [11, 12] and the layered structure of high-mobility GaAs/AlGaAs heterostructures [8].

Using the relation $k_B T_c = u\hbar(2k_F)$, we estimated the characteristic temperature T_c at which the magnetophonon resonance due to the interface acoustic phonons should manifest itself. For samples with the higher mobility, this temperature was found to be about 27 K, and for samples with the lower mobility, about 22 K, which fully agrees with the temperature dependence of the magnetoresistance oscillation amplitude observed in the experiment. We also compared the experimental dependences of $\Delta R_{xx}/R_{xx}$ on $1/B$ with the curves calculated by the formula

$$\Delta R_{xx}/R_{xx} = A\cos(4\pi k_F u/\omega_c)\exp(-\gamma 2k_F u/\omega_c), \quad (2)$$

where A is a dimensionless constant. Formula (2) was obtained from formula (1) by replacing ω_{LO} with $2k_F u$. From Fig. 3, one can see that, for the magnetophonon resonance caused by the interaction of the 2D electron gas with leaky interface acoustic phonons, the amplitude of magnetoresistance oscillations decreases according to the law given by formula (1). Here, the only fitting parameter was the constant A, which proved to be equal to 0.05 for the samples with the lower concentration and to 0.1 for the samples with the higher concentration. We explain this difference by the fact that the samples with the lower concentration of the 2D electron gas were characterized by the lower mobility, which was lower by approximately a factor of two. The parameter γ for the samples under study at T = 20 K proved to be equal to unity.

Thus, we have studied the magnetotransport properties of a high-mobility 2D electron gas in single GaAs quantum wells with AlAs/GaAs superlattice barriers at high filling factors. For the selectively doped structures under study, in the temperature range from 10 to 25 K, we observed magnetoresistance oscillations periodic in an inverse magnetic field with the oscillation frequency proportional to the Fermi wave vector of the 2D electron gas. The experimental results obtained by us are explained by the resonant interaction of the 2D electron gas with leaky interface acoustic waves characterized by the velocity $u \sim 5.9$ km/s and the wave vector $q \sim 2k_F$.

We are grateful to A. Pogosov and A. Milekhin for useful discussions. This work was supported by the Russian Foundation for Basic Research, project no. 04-02-16789, and INTAS, grant no. 03-51-6453.

REFERENCES

1. V. L. Gurevich and Yu. Firsov, Zh. Éksp. Teor. Fiz. **40**, 199 (1961) [Sov. Phys. JETP **13**, 137 (1961)].
2. S. M. Puri and T. H. Geballe, Bull. Am. Phys. Soc. **8**, 309 (1963).
3. Yu. A. Firsov, V. L. Gurevich, R. V. Parfeniev, and S. S. Shalyt, Phys. Rev. Lett. **12**, 660 (1964).
4. R. A. Stradling and R. Eood, J. Phys. C **1**, 1711 (1968).
5. D. C. Tsui, T. Englert, A. Y. Cho, and A. C. Gossard, Phys. Rev. Lett. **44**, 341 (1980).
6. N. Mori and T. Ando, Phys. Rev. B **40**, 6175 (1989).
7. V. V. Afonin, V. L. Gurevich, and R. Laiho, Phys. Rev. B **62**, 15913 (2000).
8. M. A. Zudov, I. V. Ponomarev, A. L. Efros, *et al.*, Phys. Rev. Lett. **86**, 3614 (2001).
9. M. A. Zudov, R. R. Du, J. A. Simmons, and J. L. Reno, Phys. Rev. B **64**, 201311 (2001).
10. C. L. Yang, J. Zhang, R. R. Du, *et al.*, Phys. Rev. Lett. **89**, 076801 (2002).
11. K.-J. Friedland, R. Hey, H. Kostial, *et al.*, Phys. Rev. Lett. **77**, 4616 (1996).
12. A. A. Bykov, A. K. Bakarov, L. V. Litvin, and A. I. Toropov, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 300 (2000) [JETP Lett. **72**, 209 (2000)].
13. A. A. Bykov, A. K. Kalagin, and A. K. Bakarov, Pis'ma Zh. Éksp. Teor. Fiz. **81**, 498 (2005) [JETP Lett. **81** (8), 406 (2005)].
14. C. Faugeras, D. K. Maude, G. Martinez, *et al.*, Phys. Rev. B **69**, 073405 (2004).
15. I. V. Ponomarev and A. L. Efros, Phys. Rev. B **63**, 165305 (2001).

Translated by E. Golyamina

JETP Letters, Vol. 81, No. 10, 2005, pp. 527–532. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 650–655.
Original Russian Text Copyright © 2005 by Belyavsky, Kapaev, Kopaev.

Topology of the Fermi Surface and the Coexistence of Orbital Antiferromagnetism and Superconductivity in Cuprates

V. I. Belyavsky, V. V. Kapaev, and Yu. V. Kopaev

Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 119991 Russia

e-mail: kopaev@sci.lebedev.ru

Received April 28, 2005

It has been shown that the strong coupling model taking into account a rise in the spin antiferromagnetic insulating state explains the doping dependence of the topology and shape of the Fermi contour of superconducting cuprates. Hole pockets with shadow bands in the second Brillouin zone form the Fermi contour with perfect ordinary and mirror nesting, which ensures the coexistence of orbital antiferromagnetism and superconductivity with a large pair momentum for $T < T_C$. The weak pseudogap region ($T_* < T < T^*$) corresponds to the orbital antiferromagnetic ordering, which coexists with the incoherent state of superconducting pairs with large momenta in the strong pseudogap region ($T_C < T < T_*$). *© 2005 Pleiades Publishing, Inc.*

PACS numbers: 78.47.+p, 78.66.–w

1. Angle-resolved photoemission spectroscopy (ARPES) indicates that there is the only electron-spectrum band in cuprates [1] to which the Fermi contour belongs. The shape of the Fermi contour depends strongly on the doping level [2, 3]. For relatively low hole concentrations, the Fermi contour has the shape of a closed square with smoothed corners that is centered at the point (π, π) and whose sides are parallel to the boundaries of the 2D *crystallographic* Brillouin zone. As the doping level increases, the square is turned by an angle of $\pi/4$ and its center is shifted to the point (0, 0), which is accompanied by the transition from hole single-particle excitations to electronic excitations.

Thus, the *large* Fermi contour appears, which corresponds to Luttinger's theorem and bounds the momentum space region containing $(1 + x)$ holes [4]. This property contradicts the expectation that the doping of the parent antiferromagnetic insulator must lead to the Fermi contour in the form of hole pockets that are situated in the lower antiferromagnetic band and have centers at the boundary of the 2D *magnetic* Brillouin zone.

The large Fermi contour corresponds to the maximum spectral intensity of the ARPES signal. Lower maxima of the intensity form the fine structure of the Fermi contour (*umklapp bands*), which is associated with the hopping processes due to the existence of the superstructure in the cuprate plane of the Bi2212 compound studied in [1].

More lower maxima form *shadow bands* observed in [5]. The appearance of shadow bands due to short-range antiferromagnetic correlations was predicted in [6], and these bands exactly correspond to bands that remain in the second magnetic zone after the reduction to the first magnetic Brillouin zone. The low spectral intensity of shadow bands can be explained [1] by the smallness of the matrix elements that determine this intensity in the second magnetic zone. Shadow bands are mirror reflections of sections of the main Fermi contour (*main bands*) with respect to the boundary of the magnetic Brillouin zone and, together with these sections, reproduce the structure of hole pockets.

In underdoped cuprates, a closed Fermi contour is observed for $T < T^*$, where T^* determines the upper boundary of the pseudogap regime. As the temperature decreases, a break arises in the Fermi contour at $T \approx T^*$ near the *antinodal* points corresponding to the directions $[0, \pi]$ so that the Fermi contour has the form of arcs [7] symmetric with respect to the *nodal* directions $[\pi, \pi]$. Beyond these arcs, the minimum energy of single-particle excitations is nonzero, which is interpreted as the appearance of the pseudogap. As the temperature decreases, the pseudogap region increases, and point loci corresponding to the pseudogap and superconducting gap coincide with each other at the temperature T_C of the superconducting transition.

Chakravarty *et al.* [8] assumed that the pseudogap state corresponds to the orbital antiferromagnetic ordering in the form of a charge current density wave with *d*-wave symmetry (*d*-density wave). Since the transition from the pseudogap state to the normal state is apparently similar to a crossover rather than to a phase transition, the orbital antiferromagnetic order considered in [8] was called hidden, because it is difficult to observe it experimentally. The *d*-density wave concept of ordering allows the interpretation of such a transition as an infinite order phase transition [9] that is

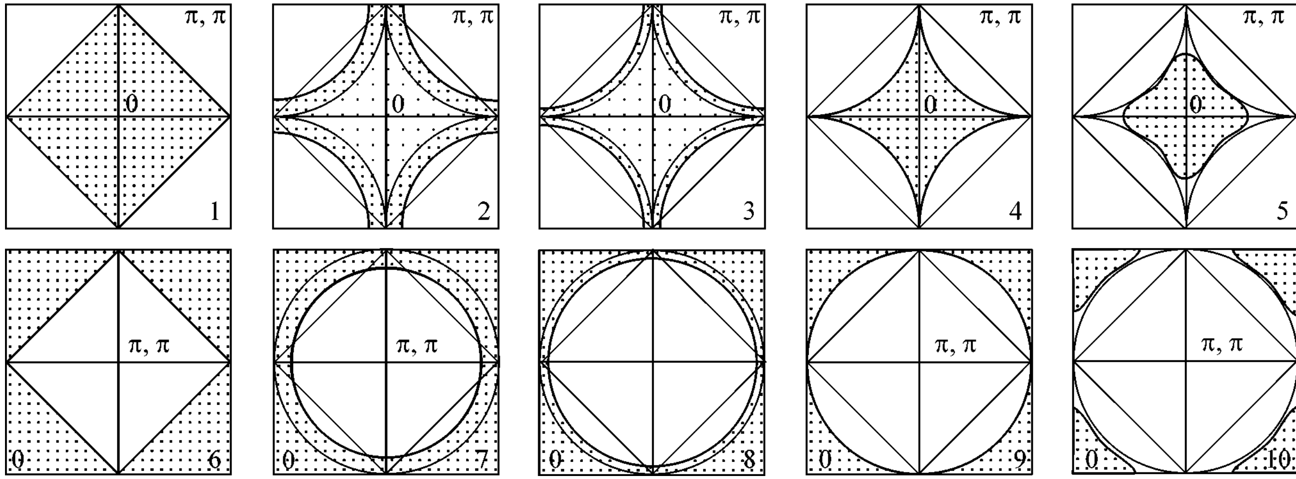

Fig. 1. Doping-induced evolution of (upper panels) electron and (lower panels) hole Fermi contours. Electron filling regions are distinguished. The necessary comments are given in the text.

free of singularities of the thermodynamic quantities that are inherent in a second order phase transition. A natural consequence of this concept is the considerable difference between the intensities of the spectral sections of the Fermi contour in the first and second magnetic Brillouin zones [10].

The aim of this work is to explain the behavior of the Fermi contour in the hole-doped cuprates that is associated with the insulating instability of the system close to the parent antiferromagnetic insulator.

2. The strong-coupling model taking into account interaction in the second coordination sphere ($t - t'$ model) leads to the dispersion law

$$\varepsilon(k_x, k_y) = -2t(\cos k_x + \cos k_y) + 4t'\cos k_x \cos k_y, \quad (1)$$

which well describes the observed shape of the Fermi contour [11] for x_{opt} with the appropriate choice of the parameter t'/t.

The evolution of the Fermi contour with variation in the hole doping level can be *qualitatively* described in the $t - t'$ model of a *rigid band* [12], whose minimum and maximum under the condition $t > 2t'$ are located at the points (0, 0) and (π, π), respectively. The upper panels of Fig. 1 (schemes 1–5) show the position of the Fermi contour separating the regions of (distinguished regions) occupied and unoccupied *electronic* states in the 2D crystallographic Brillouin zone centered at the point (0, 0) for various doping levels. Schemes 6–10 in the lower panels of Fig. 1 repeat schemes 1–5 in the Brillouin zone centered at the point (π, π). Schemes 1 and 6 correspond to the nearest-neighbor approximation ($t' = 0$) for half filling. In this case, the Fermi contour has the shape of a square with the nesting vector [π, π] and passes through the *saddle points* (±π, 0) and (0, ±π) of the electron dispersion law. Thus, the Fermi contour in schemes 1 and 6 is a separatrix separating two families of closed isolines, one of which is centered at the minimum of the energy band and the second, at its maximum.

For $0 < t' < 2t$, the separatrix changes its shape, but it passes through the same saddle points (as is shown in all schemes except for 1 and 6) and thereby bounds the momentum-space region smaller than the half-filling region, which corresponds to the isoline for an energy higher than the energy of a saddle point (schemes 2 and 7). Correspondingly, the Fermi contour is a closed line bounding the region of *unoccupied* states with the center at the point (π, π), which is the *topological center* of the Fermi contour (scheme 7).

When doping is weak, the introduction of additional holes shifts the Fermi contour to the separatrix (schemes 3 and 8). In this case, the topological center of the Fermi contour remains at the point (π, π). As the doping level increases, the separatrix coincides with the Fermi contour at a certain hole concentration $x = x_t$ (schemes 4 and 9) and this concentration corresponds to the displacement of the topological center of the Fermi contour (for $x > x_t$, the closed line in scheme 5 bounds the region of occupied *electronic* states) to the point (0, 0). Such a displacement can be considered as the *electronic topological transition* associated with the van Hove singularity at a saddle point [13].

The doping-induced evolution of the Fermi contour that is schematically shown in Fig. 1 corresponds to the evolution observed in ARPES experiments in the LSCO and BSCCO cuprate compounds. However, a number of substantial features of the behavior of the Fermi contour in the doping process, in particular, the structure of hole pockets and the corresponding shadow bands, cannot be explained in the rigid-band scheme, because a rise in the antiferromagnetic insulating state is disregarded in this scheme.

3. The long-range spin (triplet) antiferromagnetic order in cuprates in the form of the spin density wave exists in a comparatively narrow range of doping levels

near half filling ($0 < x \lesssim 0.03$), but fluctuation of the spin density wave are observed up to the optimum doping level $x_{\rm opt}$. Translational symmetry is broken in this case. Therefore, electronic states must be classified within the 2D *magnetic* Brillouin zone at the boundary of which the *s*-wave insulating gap Δ_S arises (it is a monotonic function of the doping level).

The band with the dispersion law $\varepsilon(\mathbf{k})$ is split into two subbands

$$\varepsilon_{1,2}(\mathbf{k}) = \varepsilon_+(\mathbf{k}, \mathbf{Q}) \pm \sqrt{\varepsilon_-^2(\mathbf{k}, \mathbf{Q}) + \Delta_S^2}, \quad (2)$$

where $2\varepsilon_\pm(\mathbf{k}, \mathbf{Q}) = \varepsilon(\mathbf{k}) \pm \varepsilon(\mathbf{k} + \mathbf{Q})$ and $\mathbf{Q} = (\pi, \pi)$. The signs minus and plus correspond to lower $\varepsilon_1(\mathbf{k})$ and upper $\varepsilon_2(\mathbf{k})$ subbands, respectively. Since $\Delta_S = \Delta_S(x)$, dispersion law (2) also depends on the doping level.

The area of the magnetic Brillouin zone corresponds to half filling ($x = 0$). For weak doping, the hole filling of the lower energy subband (unoccupied states in the upper electron subband) exists in almost the entire first magnetic zone [centered at the point (π, π), see Fig. 2a], except for comparatively small vicinities of the points of the type $(\pi, 0)$ with the electron filling of the upper electron subband; i.e., the total concentration of holes in the first magnetic zone is approximately equal to 1.

In addition, the first magnetic Brillouin zone includes hole pockets (distinguished in Fig. 2b) corresponding to the filling of the upper hole energy subband (unoccupied states in the lower electron subband). Almost all the doping-introduced holes are located in these pockets; i.e., the area of the pockets in the first magnetic zone is approximately equal to x. These extra (over the insulating gap) holes are main charge carriers in underdoped cuprates, and only they are involved in superconducting pairing and determine, in particular, the superfluid density $n_s \sim x$.

The area of hole pockets increases with the doping level until the time when the topological center of the Fermi contour at $x = x_t$ (schemes 4 and 9 in Fig. 1) is displaced from (π, π) to $(0, 0)$, which is accompanied by the change in the charge sign of the charge carriers. For $x > x_t$, the electron concentration in the lower energy subband is equal to $1 - x$.

Superconductivity disappears at a certain doping level near the point of the electronic topological transition x_t. There are reasons [14] to think that the large Fermi contour of strongly overdoped cuprates in the superconducting state is simply connected and centered at the point (π, π), i.e., corresponds to scheme 8 in Fig. 1.

In the crystallographic Brillouin zone, the Fermi contour (the line separating the regions of occupied and unoccupied states in Fig. 2a) bounds, according to Luttinger's theorem, the momentum-space region whose area corresponds to the total hole concentration $1 + x$. The reduction to the magnetic Brillouin zone is accompanied by the transformation of the Fermi contour arcs (marked by digits in Fig. 2a) inside this region (thick

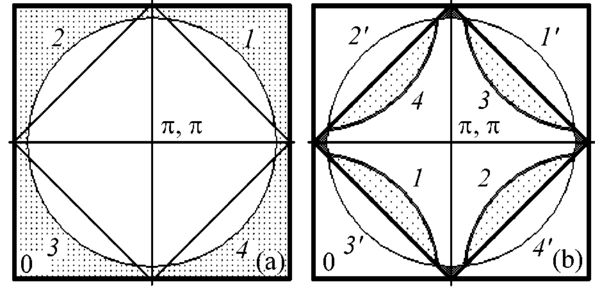

Fig. 2. Reduction to the magnetic Brillouin zone centered at the point (π, π) (hole representation). The doping level corresponds to scheme 8 in Fig. 1. Digits in panel (a) mark the Fermi contour arcs shown in the scheme of extended bands. Digits with prime in panel (b) mark the Fermi contour arcs in the second magnetic Brillouin zone that correspond to shadow bands. The hole-filling regions of the upper hole energy band are distinguished within the first magnetic Brillouin zone. The dark regions near the points of the type $(\pi, 0)$ correspond to the electron filling of the lower hole subband in the first magnetic zone.

lines marked by the corresponding digits in Fig. 2b). In the second magnetic Brillouin zone, the Fermi contour arcs remain at the same places as in the crystallographic Brillouin zone (marked by digits with primes in Fig. 2b).

4. The insulating gap depending on the doping level can be considered as a fitting parameter that is additional to t'/t and allows the reproduction of finer details of the Fermi contour compared to the simple pattern following from the $t - t'$ model. The introduction of an additional fitting parameter immediately in the strong coupling model, i.e., the inclusion of interaction in the third coordination sphere (the $t - t' - t''$ model [11]) cannot lead to the appearance of shadow bands, because this interaction is not associated with the breaking of the translational symmetry and provides the description of the observed change in the large Fermi contour only with very nonrealistic values of the parameters t'/t and t''/t. In contrast, dispersion law (2) is naturally defined in the magnetic Brillouin zone, and the large Fermi contour is obtained by its simple expansion in the crystallographic Brillouin zone as is shown in Fig. 2.

Taking into account the value and the doping dependence of the fitting parameter $\Delta_S(x)$, one can verify that dispersion law (2) is in complete agreement with the observed doping-induced evolution of the Fermi contour [2]. Figure 3 shows the distribution of isolines in the lower subband $\varepsilon_1(\mathbf{k})$ for $t'/t = 0.3$ and weak doping corresponding to $\Delta_S(x) = 0.3$. The isolines shown in Fig. 3 are hole pockets centered at the points of the type $(\pi/2, \pi/2)$.

For weak doping $x < x_t$, the sections of the large Fermi contour that belong to the lower energy subband are parts of hole pockets situated in the second magnetic Brillouin zone (see Fig. 3). These sections are connected with each other by isolines belonging to the

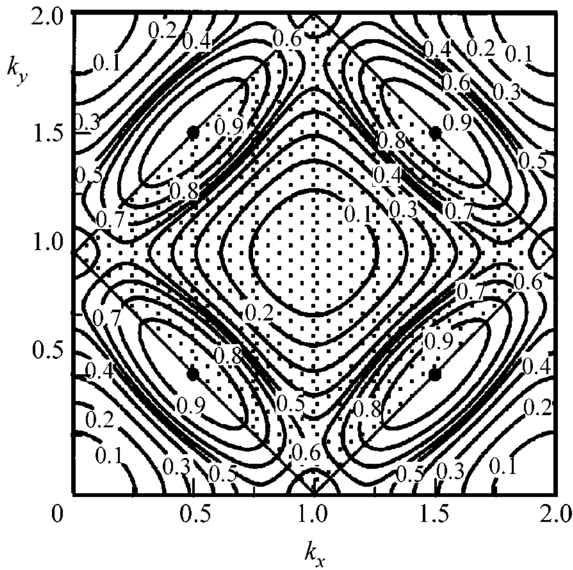

Fig. 3. Isoline distribution in the lower subband. The filling factors are shown near the isolines. The first magnetic Brillouin zone is distinguished. Weak doping corresponds to the Fermi contour in the form of hole pockets whose centers are shown.

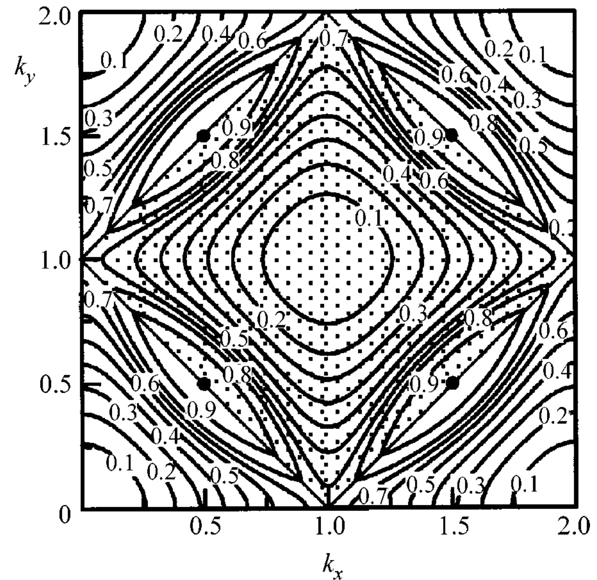

Fig. 4. Isoline distribution in the lower subband for strong doping. The first magnetic Brillouin zone is distinguished. The Fermi contour is centered at the point (0, 0) and corresponds to the filling factor $\lesssim 0.7$.

upper subband; i.e., the upper and lower subbands overlap and the Fermi contour for weak doping belongs to both subbands. The closed large Fermi contour centered at the point (π, π) has the shape of a distorted square with smoothed angles and sides parallel to the boundaries of the crystallographic Brillouin zone.

The appearance of the insulating gap leads to the splitting of the saddle point: the separatrix separating the families of closed isolines is seen in Fig. 3. An increase in the doping level reduces the insulating gap and correspondingly changes the structure of isolines.

As the doping level increases, the areas of the hole pockets proportional to x increase up to the time when all electrons leave the upper electron subband and the electronic topological transition occurs in the lower subband with the displacement of the topological center of the large Fermi contour from the point (π, π) to the point (0, 0) as is shown in Fig. 4 for $\Delta_S(x) = 0.001$. The large Fermi contour bounding the electron filling region with the area $1 - x$ has the shape of a distorted square with smoothed corners and sides parallel to the diagonal of the crystallographic Brillouin zone. The saddle point remains unsplit and is located at the points of the type (π, 0), as in the $t - t'$ model of the hard zone.

We emphasize that the Fermi contour features under consideration that are associated with the appearance of the spin density wave are manifested for underdoped cuprates at temperatures $T > T^*$. In the pseudogap state, i.e., in the range $T_C < T < T^*$, the appearance of the *d*-density wave associated with the orbital antiferromagnetic order gives rise to the appearance of the insulating gap (which appears as a pseudogap) on the Fermi contour in the form of hole pockets. This appearance is responsible for the Fermi contour arcs observed in ARPES near the nodal points. Further doping (beyond the superconducting phase region) leads to the displacement of the topological center of the Fermi contour to the point (0, 0).

5. The filling of the energy subbands is determined by the coherence factors $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ in the Bogoliubov transformation, which diagonalizes the Hamiltonian describing the electron–hole pairing, as a result of which the spin density wave arises. The filling probability is characterized by the coefficient $v_{\mathbf{k}}^2$, which is on the order of unity within the first magnetic zone. For this reason, the spectral intensity of the ARPES signal is high for hole pocket arcs belonging to the first magnetic zone (outer arcs in the crystallographic Brillouin zone, see Fig. 2a). The transition to the second magnetic Brillouin zone corresponds to the change $\mathbf{k} \longrightarrow \mathbf{k} + \mathbf{Q}$. Since $v_{\mathbf{k}+\mathbf{Q}} = u_{\mathbf{k}}$, the ARPES response from the second magnetic zone is significantly weakened (as in the appearance of the *d*-density wave [10]); i.e., the Fermi contour arcs in this zone correspond to very low spectral intensity. For this reason, these arcs can be identified with shadow energy bands.

The insulating *s*-wave gap associated with the reduction of the translational symmetry of the cuprate plane when the spin density wave arises is formed at the boundary of the magnetic Brillouin zone. This gap cannot be identified with the pseudogap that arises on the Fermi contour [7] and has the *d*-wave symmetry. Hence, the appearance of the pseudogap state can be attributed to the orbital antiferromagnetic ordering

leading to the *d*-density wave. The perfect nesting of the Fermi contour in the form of hole pockets with the nesting vector $\mathbf{Q} = (\pi, \pi)$, which connects different pockets as is seen in Fig. 3, must lead to the appearance of the *d*-density wave order, which can coexist with the spin density wave order in the region of weak doping ($x \lesssim x_{opt}$).

The spectral intensities of the ARPES signal that correspond to the spin density wave and *d*-density wave are determined by their coherence factors depending on the doping level. The redistribution of the intensities between two channels of dielectric pairing with an increase in the doping level [15] is likely manifested in a decrease in the ARPES signal associated with the spin density wave state with the simultaneous enhancement of the signal, which can be associated with the *d*-density wave state.

The long-range spin antiferromagnetic order in the doping region corresponding to the superconducting state is evidently absent, and the degree of the short-range order decreases as x increases. However, the efficient susceptibility to the orbital antiferromagnetic order, as well as the susceptibility to the appearance of the superconducting order with the same total pair momentum $\mathbf{Q}$, has a maximum at $\mathbf{Q} = (\pi, \pi)$. This circumstance enables one to introduce the mean value of the insulating gap associated with the appearance of the *d*-density wave.

Owing to the formation of the *d*-density wave energy gap, the chemical potential lies within this gap, and the system in the pseudogap state behaves as an insulator.

6. The Fermi contour in the form of hole pockets exhibits the perfect *mirror nesting* feature [16], which is a favorable condition for superconducting pairing with a large total momentum coinciding in this case with $\mathbf{Q} = (\pi, \pi)$. The feature of this pairing in this case is that the momenta of particles forming a pair belong to different Brillouin zones; i.e., one of them refers to the main Fermi contour branch and the second, to the shadow Fermi contour branch. These different Brillouin zones correspond to different coherence factors arising in superconducting pairing. As the doping level increases, the area of pockets increases; i.e., the region of the kinematic constraint for superconducting pairing expands. At the same time, the intensity of such pairing that is proportional to the spectral density in the shadow branch decreases. For this reason, T_C as a function of the doping level must have a maximum.

Pairing with a large momentum of a pair for repulsive pairing interaction leads to the complex two-component order parameter [17] whose absolute value is determined by the binding energy of the pair and the relative phase is naturally related [18] to the orbital antiferromagnetic current *d*-density wave state in the cuprate plane [19]. The order parameter appearing after superconducting pairing for repulsive interaction as a function of the momentum of the relative motion of the pair in the kinematic-constraint region has a line of zeros that intersects the Fermi contour [20] and can corresponds to either extended *s*-wave ($s + g$) symmetry or *d*-wave symmetry [21, 22]. The orbital antiferromagnetic order parameter associated with the appearance of the *d*-density wave has *d*-wave symmetry and vanishes at four intersection points of the main branches of the Fermi contour with the diagonals of the crystallographic Brillouin zone. The phase of the superconducting order parameter that corresponds to the coexistence of superconductivity and orbital antiferromagnetism is equal to π at these points [18]. Thus, the existence of zeros of the two-component superconducting order parameter at *different* points of the Fermi contour for charge (absolute value) and current (phase) degrees of freedom ensures the stability of the superconducting state with respect to electron scattering on nonmagnetic impurities.

As the doping level increases, the amplitude of the orbital antiferromagnetic order parameter decreases. Correspondingly, the deviation of the relative phase of the superconducting order parameter from π also decreases. The superconducting region on the temperature–doping phase diagram is divided into two phases [18]. In one of these phases, the superconducting and orbital antiferromagnetic ordered states coexist, and, in the other phase corresponding to the higher doping level, the orbital antiferromagnetic order is absent (the relative phase of the superconducting order parameter is equal to π). The superconducting order parameter has zeros on the Fermi contour and is sensitive to electron scattering on impurities. This sensitivity may become one of the causes of the suppression of superconductivity with an increase in the doping level.

In underdoped cuprates, an increase in temperature destroys phase coherence at $T = T_C$; i.e., the mean value of the common phase of the complex two-component order parameter vanishes. The absolute values of the components of the order parameter and its relative phase (whose deviation from π corresponds to the manifestation of the current degree of freedom as the orbital antiferromagnetic ordering [18]) are also nonzero in a certain range $T_C < T < T^*$ above T_C. Thus, the long-range orbital antiferromagnetic order, which coexists with superconductivity for $T < T_C$, continues to exist for $T_C < T < T^*$ and can be attributed to the pseudogap state [8], which corresponds to the charge current density wave (*d*-density wave) responsible for the insulating gap arising on the Fermi contour.

In the temperature range $T_C < T < T_* < T^*$, developed fluctuations exist in the superconducting order parameter, which are associated with the creation of incoherent superconducting pairs with large momenta. These fluctuations give rise to the effective increase in the energy gap in the spectrum of single-particle excitations, which can be interpreted as a *strong pseudogap* in the range $T_C < T < T_*$. Fluctuations in the superconducting order parameter are suppressed in the range

$T_* < T < T^*$, and this range characterized by the insulating orbital antiferromagnetic gap can be referred to the *weak pseudogap*.

This work was supported in part by the Russian Foundation for Basic Research.

REFERENCES

1. H. Ding, A. F. Bellman, J. C. Campuzano, *et al.*, Phys. Rev. Lett. **76**, 1533 (1996).
2. A. Fujimori, A. Ino, T. Yoshida, *et al.*, Physica C (Amsterdam) **341–348**, 2067 (2000).
3. A. Ino, C. Kim, M. Nakamura, *et al.*, Phys. Rev. B **65**, 094504 (2002).
4. H. Ding, M. R. Norman, T. Yokoya, *et al.*, Phys. Rev. Lett. **78**, 2628 (1997).
5. P. Aebi, J. Osterwalder, P. Schwaller, *et al.*, Phys. Rev. Lett. **72**, 2757 (1994).
6. A. P. Kampf and J. R. Schrieffer, Phys. Rev. B **42**, 7967 (1990).
7. M. R. Norman, H. Ding, M. Randeria, *et al.*, Nature **392**, 157 (1998).
8. S. Chakravarty, R. B. Laughlin, D. K. Morr, and C. Nayak, Phys. Rev. B **63**, 094503 (2001).
9. S. Chakravarty, Phys. Rev. B **66**, 224505 (2002).
10. S. Chakravarty, C. Nayak, and S. Tewari, cond-mat/0306084.
11. A. I. Liechtenstein, O. Gunnarsson, O. K. Andersen, and R. M. Martin, Phys. Rev. B **54**, 12 505 (1996).
12. Z.-X. Shen, W. E. Spicer, D. M. King, *et al.*, Science **267**, 343 (1995).
13. I. M. Lifshits, Zh. Éksp. Teor. Fiz. **38**, 1569 (1960) [Sov. Phys. JETP **11**, 1130 (1960)].
14. M. Plate, J. D. F. Mottershead, I. S. Elfimov, *et al.*, cond-mat/0503117.
15. A. Ino, C. Kim, M. Nakamura, *et al.,* Phys. Rev. B **62**, 4137 (2000).
16. V. I. Belyavskiĭ, V. V. Kapaev, and Yu. V. Kopaev, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 51 (2002) [JETP Lett. **76**, 44 (2002)].
17. V. I. Belyavskiĭ and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **127**, 45 (2005) [JETP **100**, 39 (2005)].
18. V. I. Belyavskiĭ, Yu. V. Kopaev, and M. Yu. Smirnov, Zh. Éksp. Teor. Fiz. (2005) (in press).
19. D. A. Ivanov, P. A. Lee, and X.-G. Wen, Phys. Rev. Lett. **84**, 3958 (2000).
20. V. I. Belyavskiĭ, Yu. V. Kopaev, V. M. Sofronov, and S. V. Shevtsov, Zh. Éksp. Teor. Fiz. **124**, 1149 (2003) [JETP **97**, 1032 (2003)].
21. G. Zhao, Phys. Rev. B **64**, 024503 (2001).
22. B. H. Brandow, Phys. Rev. B **65**, 054 503 (2002).

Translated by R. Tyapaev

JETP Letters, Vol. 81, No. 10, 2005, pp. 533–537. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 656–660.
Original Russian Text Copyright © 2005 by Podlivaev, Openov.

Stone–Wales Transformation Paths in Fullerene C_{60}

A. I. Podlivaev and L. A. Openov*

Moscow Engineering Physics Institute (State University), Kashirskoe sh. 31, Moscow, 115409 Russia

** e-mail: opn@supercon.mephi.ru*

Received April 28, 2005

The mechanisms of formation of a metastable defect isomer of fullerene C_{60} due to the Stone–Wales transformation are theoretically studied. It is demonstrated that the paths of the "dynamic" Stone–Wales transformation at a high (sufficient for overcoming potential barriers) temperature can differ from the two "adiabatic" transformation paths discussed in the literature. This behavior is due to the presence of a great near-flat segment of the potential-energy surface in the neighborhood of metastable states. Moreover, the sequence of rupture and formation of interatomic bonds is other than that in the case of the adiabatic transformation. *© 2005 Pleiades Publishing, Inc.*

PACS numbers: 36.40.Cg, 61.48.+c, 71.15.Pd

Various models of fullerene growth have been proposed [1]. Because a sphere-shaped cluster of C_{60} is formed at a high temperature under actual operating conditions, its structure can be significantly different from the structure of ideal buckminsterfullerene with I_h icosahedral symmetry. On the "surface" of buckminsterfullerene, carbon atoms are arranged at the vertexes of 20 hexagons and 12 pentagons, which are isolated from each other. Therefore, regardless of the mechanism of formation of such a spheroidal cluster from graphite fragments and/or carbon dimers, the question arises of the paths of its evolution into an equilibrium configuration of buckminsterfullerene, that is, the question of the mechanisms of defect annealing that reduces the potential energy E_{pot} of the cluster.

Defect annealing occurs by local rearrangements of C–C bonds and is the reverse process of defect formation. The so-called Stone–Wales transformation [2], which consists in the rearrangement of two C–C bonds in buckminsterfullerene (see Fig. 1), results in a defect isomer closest to buckminsterfullerene in terms of energy. This metastable isomer exhibits C_{2v} symmetry and contains two pairs of pentagons with common sides. Among a great number of other isomers, this isomer stands out as the last segment (before buckminsterfullerene) in the chain of sequentially decreasing energies of a C_{60} cluster on defect annealing [3].

In this context, attention was focused on studying changes in the mutual arrangement of atoms in the Stone–Wales transformation and on determining the height U of the minimum energy barrier to this transformation [4–10]. In this case, two different transformation paths were considered: (I) the rotation of a C–C bond shared by two hexagons through 90° so that all of the atoms remained on the cluster "surface" and (II) the rupture of a C–C bond shared by pentagons and hexagons; thereafter, one of the atoms initially rose above the cluster "surface" and then relaxed to form new bonds. In accordance with the calculations of E_{pot} as a function of atomic coordinates $\{\mathbf{R}_i\}$ performed using the density functional theory [5–10], the barrier height for path I was U_I = 6–8 eV, whereas U_{II} = 6.64–7.6 eV for path II. A considerable scatter in the values of U_I and U_{II} does not allow one to determine unambiguously which of the two paths exhibits a lower barrier and, consequently, by which mechanism the Stone–Wales transformation occurs. Even with the use of the same exchange-correlation functional and equal basis-function sets (or equal cutoff energies in a plane-wave basis) for calculating U_I and U_{II}, different results were obtained: $U_I = U_{II} = 7.27$ eV [10], $U_I = 6.30$ eV $< U_{II} =$ 6.64 eV [9], and $U_I = 8.1$ eV $> U_{II} = 7.6$ eV [6, 7].

Note that the determination of the height of a barrier and the type of a transition state based on the shape analysis of a potential-energy surface implies an adiabatic transition of the cluster from one state to another. That is, in such a transition, one or more properly cho-

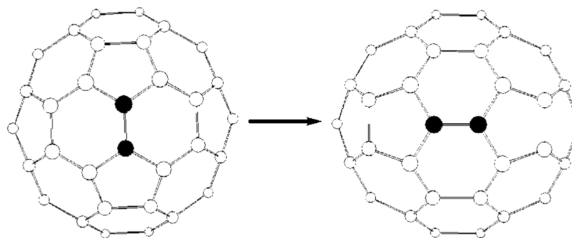

Fig. 1. Formation of two pairs of pentagons with common sides on the rearrangement of two C–C bonds in fullerene C_{60} (Stone–Wales transformation). For clarity, atoms in the background are not shown.

0021-3640/05/8110-0533$26.00 © 2005 Pleiades Publishing, Inc.

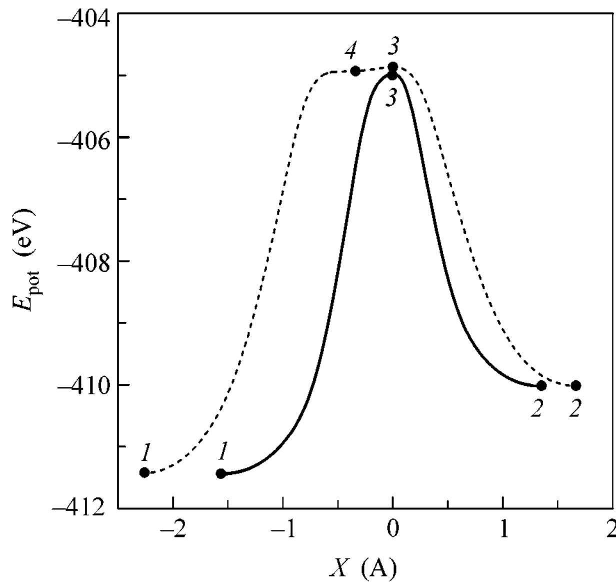

Fig. 2. Potential energy E_{pot} of the C_{60} cluster vs. the reaction coordinate X in the neighborhoods of (*1*) equilibrium and (*2*) metastable atomic configurations in the Stone–Wales transformation (Fig. 1). The energy of 60 isolated carbon atoms was taken as zero energy. The solid and dashed lines indicate paths I and II, respectively (see the text); points *3* are $E_{pot}(X)$ maximum points [saddle points for $E_{pot}(\{\mathbf{R}_i\})$]; and *4* is a d^2E_{pot}/dX^2 minimum point for path II. The path length along a trajectory that passes through the corresponding saddle point in $(3n - 6)$-dimensional space and joins buckminsterfullerene with the metastable isomer was chosen as the reaction coordinate. $X = 0$ at saddle points.

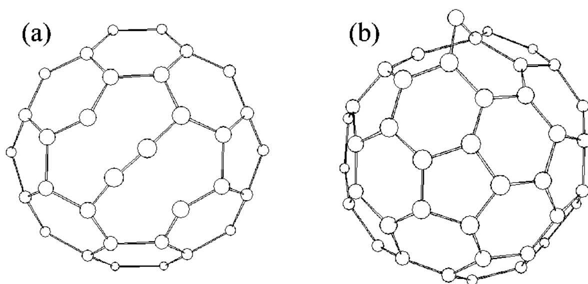

Fig. 3. Atomic configurations of transition states *3* in Fig. 2 in the Stone–Wales transformation via (a) path I (symmetric transition state) and (b) path II (asymmetric transition state). Atoms in the background and C–C bonds more than 2 Å in length are not shown.

sen interatomic distances change in accordance with a certain law with a small step along the reaction coordinate, whereas relaxation to a minimum (with consideration for imposed restrictions) total energy is performed at each step in all of the other degrees of freedom. Physically, an adiabatic transition corresponds to a forced deformation of the cluster at $T = 0$, and it must not occur identically to transitions at a finite (especially, high) temperature. In this work, we studied Stone–Wales transformation paths both in an adiabatic transition and at a high temperature (sufficient for overcoming a potential barrier) using a molecular dynamic simulation of the "life" of buckminsterfullerene. A comparison between the results allowed us, on the one hand, to find common features in the low-temperature and high-temperature buckminsterfullerene ⟶ metastable isomer transitions and, on the other hand, to reveal considerable differences between them.

We used a tight binding potential for the calculations. This potential incorporates the contribution of all the valence electrons (four electrons from each atom) to the total energy in an explicit form, and it is actually an n-body potential, where n is the number of atoms in the cluster. Because of this, this potential surpasses simple but less reliable empirical potentials. Although this method is not as strict as *ab initio* approaches, it is highly competitive with them in the accuracy of the description of carbon systems [11]. Moreover, this method considerably facilitates the simulation of dynamic processes because it is not resource-intensive. Previously, we used this method for an analysis of the thermal stability of metastable C_8 and C_{20} clusters [12–14]. For buckminsterfullerene, it gave the binding energy $E_b = 60E(C_1) - E(C_{60}) = 6.86$ eV/atom and the bond lengths $l = 1.396$ and 1.458 Å, which are consistent with the experimental values of E_b = 6.97–7.01 eV/atom [15] and $l = 1.402$ and 1.462 Å [16]. The HOMO–LUMO gap $\Delta = 1.62$ eV is also consistent with the experimental value of $\Delta = 1.6$–1.8 eV [17].

Initially, we studied the mechanisms of the buckminsterfullerene ⟶ metastable isomer isomerization by adiabatic transitions. For this purpose, we calculated the "potential relief" $E_{pot}(\{\mathbf{R}_i\})$ of the C_{60} cluster in the vicinity of an equilibrium atomic configuration (buckminsterfullerene) and the atomic configuration closest in energy (metastable isomer) for the above two Stone–Wales transformation paths. Figure 2 shows the results of this calculation. The energy of the metastable isomer is higher than the energy of buckminsterfullerene by $\Delta E = 1.42$ eV, which is consistent with the value of $\Delta E = 1.4$–1.7 eV obtained by the density functional method [5–10]. In both path I and path II, the dependence of E_{pot} on reaction coordinate X exhibits the only stationary point (a maximum), which is a saddle point for the $E_{pot}(\{\mathbf{R}_i\})$ surface and determines the barrier height. According to our calculations, the value of U_I = 6.46 eV is somewhat lower than $U_{II} = 6.58$ eV.

Figure 3a shows the atomic configuration of the transition state for path I. It is "symmetric" in accordance with available published data [6–10]. On the contrary, the transition state for path II is asymmetric (Fig. 3b). One of the atoms is arranged over the cluster surface, and it forms two bonds (*sp* hybridization). In this case, another ("surface") atom has four neighbors (sp^3 hybridization). This transition state has been

reported previously [4, 6, 7, 9, 10]. Note that path II exhibits an almost horizontal segment near a maximum in the $E_{pot}(X)$ curve. The atomic configuration that corresponds to the value of X at which the second derivative $d^2E_{pot}(X)/dX^2$ exhibits a minimum (Fig. 2, point *4*) is visually similar to the transition state (Fig. 3b). Murry *et al.* [6, 7] stated that this configuration corresponds to a local minimum of E_{pot} (that is, actually, to another metastable state). In recent publications [9, 10], no additional minimum of E_{pot} was reported.

Nevertheless, we found that both of the above Stone–Wales transformation paths are characterized by barriers with approximately equal heights: $U_I = 6.46$ eV and $U_{II} = 6.58$ eV. These data are consistent with first-principle calculations [5–10], which gave similar values of U_I and U_{II} in a range of 7 ± 1 eV. One might expect that the Stone–Wales transformation under conditions of a real experiment (that is, due to thermally activated processes) occurs via both of these paths with approximately equal probabilities. To test this hypothesis, we numerically simulated the dynamics of buckminsterfullerene at $T = 4000$–5000 K and a fixed total energy of the cluster, $E = E_{kin} + E_{pot} = \text{const}$. The cluster temperature was found from the equation [18] $\frac{1}{2}k_BT(3n - 6) = E_{kin}$, where $n = 60$, and E_{kin} is the kinetic energy in the system of the center of gravity averaged over 10^3 steps of molecular dynamics (the single-step time $t_0 = 2.72 \times 10^{-16}$ s was about 0.01 of the period of the highest frequency mode of cluster vibrations). On this formulation of the problem, the temperature T is a measure of energy of relative atomic motion [18]. In this case, the value of $E_{kin} = 30$–40 eV is sufficient for overcoming the potential barriers $U_{I,II} \approx 6.5$ eV (although $E_{kin} \gg U_{I,II}$, buckminsterfullerene isomerization processes at $T \approx 4500$ K occur rarely, approximately once per 10^6t_0).

We analyzed in detail 45 buckminsterfullerene $\longrightarrow$ metastable isomer transitions and vice versa. We found that the Stone–Wales transformation occurred via path I (or by a visually very similar mechanism) in the majority of cases (37) (see Fig. 4); it occurred via path II only in three instances (see Fig. 5). Thus, although both of the paths are almost equally favorable in terms of energy (see Fig. 2), path I is much more preferable from a dynamic standpoint. We believe that this is due to the dramatic difference between the frequency factors A_I and A_{II} in the Arrhenius temperature dependence of the rates of transition $k_{I,II} = A_{I,II}\exp(-U_{I,II}/k_BT)$. Consequently, in the consideration of Stone–Wales transformation mechanisms, it is likely that attention should be focused on an analysis of the frequency factors of transitions via these two paths rather than on a further increase in the accuracy of the calculation of the barrier heights $U_{I,II}$ (and on determining which of these barriers is lower even by fractions of eV).

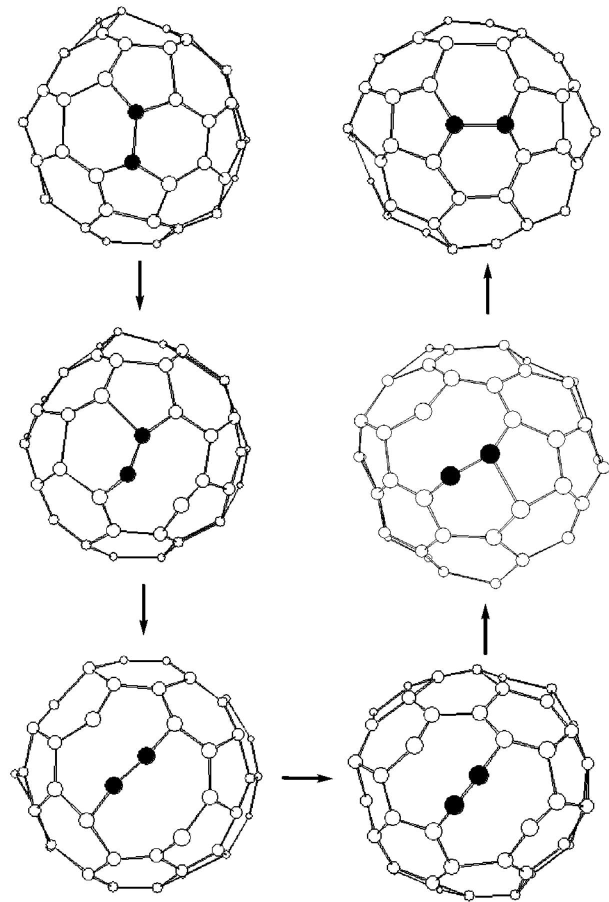

Fig. 4. Snapshots of the C_{60} cluster in the course of the dynamic Stone–Wales transformation via path I. Initial temperature $T_{ini} = 4465 \pm 5$ K. The first configuration corresponds to the time $t' = 0.245$ ns, and the subsequent configurations correspond to the times $t' + \Delta t$, where $\Delta t = 5.4$, 17.7, 69.9, 73.4, and 85.7 fs, respectively. Atoms in the background and C–C bonds more than 2 Å in length are not shown.

In addition to the transformation of buckminsterfullerene via paths I and II, we also observed transitions accompanied by the formation of either a large “window” or three adjacent windows and the appearance of a “branch” of three atoms, which was perpendicular to the cluster surface. All of these intermediate configurations correspond to potential-energy surface segments above the saddle points of transitions via paths I and II, that is, to other transformation paths with higher barriers. At $T = 5000$–5500 K, the relative number of these transitions increases dramatically. They can occasionally result in the formation of isomers with higher energies (for example, fullerenes with three or more pairs of adjacent pentagons or nonclassical fullerenes with tetragons and heptagons).

Note that, in the dynamic transformation of buckminsterfullerene via path I, as a rule, one or two atoms in the C–C bond that rotated through 90° (Figs. 1 and 3a) rose above the cluster surface through different

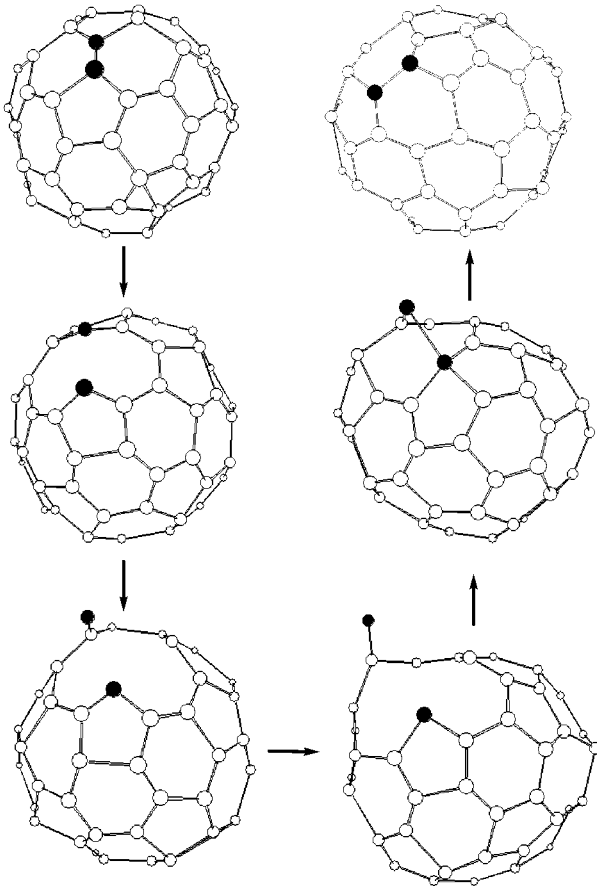

Fig. 5. Same as in Fig. 4, but for the transformation via path II. $T_{ini} = 4680 \pm 5$ K; $t' = 0.471$ ns; $\Delta t = 27, 54, 136, 299$, and 354 fs, respectively.

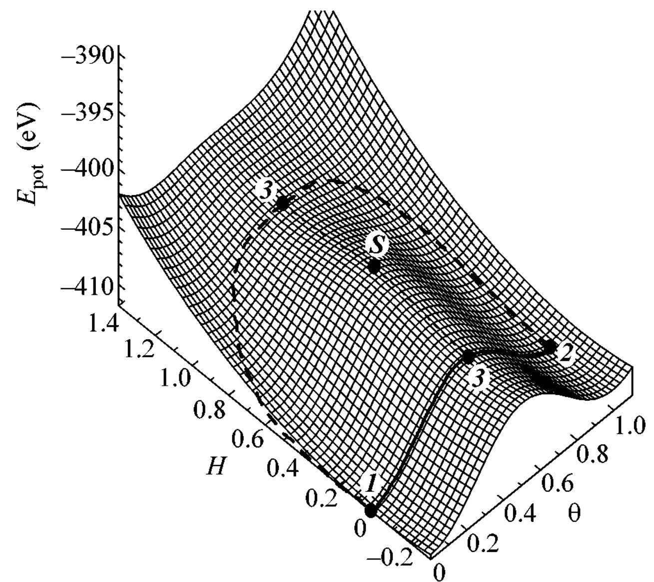

Fig. 6. Potential energy E_{pot} of the C_{60} cluster vs. two reaction coordinates H and θ in the neighborhoods of (*1*) equilibrium and (*2*) metastable atomic configurations. The energy of 60 isolated carbon atoms was taken as zero energy. The solid and dashed lines indicate paths I and II, respectively; (*3*) indicates corresponding saddle points; and S is a second-order stationary point (see the text). The reaction coordinates were chosen as follows: H is the height through which an atom of the rotated C–C bond raised over the cluster surface normalized to a maximum height through which this atom rose in its motion along path II ($H = 0$ for path I; a maximum value of $H = 1$ was at the saddle point of path II) and θ is the angle of rotation of the C–C bond normalized to 90° (θ = 0 for buckminsterfullerene, θ = 1 for the metastable isomer, and θ ≈ 0.5 for the saddle points and the second-order stationary point S).

heights every time as in the transformation via path II (but, unlike path II, without the formation of an sp^3 configuration). This suggests that there are a great number of configurations with energies close to the saddle-point energies of paths I and II. That is, in other words, a segment in the potential-energy surface between two saddle points is almost planar. To test this hypothesis, we calculated E_{pot} as a function of two reaction coordinates H and θ and found that, indeed, the $E_{pot}(H, \theta)$ function exhibits a very great radius of curvature along the coordinate H (Fig. 6). Note that a second-order stationary point (characterized by two imaginary frequencies) occurred between the saddle points of paths I and II. The energy of this stationary point is 0.56 eV higher than the energy of the saddle point of path II. This stationary point corresponds to a configuration in which one of the atoms of the rotated C–C bond rose over the cluster surface (however, to a lesser degree than in the transformation via path II).

In conclusion, let us consider the nature of the rupture and formation of interatomic bonds in the dynamic Stone–Wales transformation. Figure 4 shows a sequence of snapshots of the cluster in the course of the transformation via path I. It can be seen that the transformation began with the rupture of a single bond rather than the concerted rupture of two bonds shared by pentagons and hexagons, as in an adiabatic case [10]. Physically, this is easy to understand: a distortion of the ideal buckminsterfullerene structure due to the thermal vibrations of constituent atoms caused a symmetry breakdown in the arrangement of these two bonds. Because of this, one of these bonds became weaker than the other, and it was ruptured initially. Subsequently, the second bond was ruptured. Thereafter, a configuration similar to the transition state in an adiabatic case was formed (Fig. 3a). At the next step, new bonds were formed (consecutively rather than simultaneously) to complete the transformation. Note that the overall process occurred very rapidly and took only $\Delta t_I \approx 86$ fs. The transformation via path II began with the rupture of either a bond shared by pentagons and hexagons or a bond shared by two hexagons. For the latter case, Fig. 5 shows a sequence of snapshots of the C_{60} cluster. The configuration corresponding to the transition state (Fig. 3b) can be clearly seen in Fig. 5. The transformation time $\Delta t_{II} \approx 354$ fs is much longer than

the value of Δt_I; this is consistent with the greater "length" of path II (see Fig. 6).

This work was supported by the U.S. Civilian Research and Development Foundation (award "Research and Education Center for Basic Investigation of Matter Under Extreme Conditions").

REFERENCES

1. Yu. E. Lozovik and A. M. Popov, Usp. Fiz. Nauk **167**, 751 (1997) [Phys. Usp. **40**, 717 (1997)].
2. A. J. Stone and D. J. Wales, Chem. Phys. Lett. **128**, 501 (1986).
3. S. J. Austin, P. W. Fowler, D. E. Manolopoulos, and F. Zerbetto, Chem. Phys. Lett. **235**, 146 (1995).
4. T. R. Walsh and D. J. Wales, J. Chem. Phys. **109**, 6691 (1998).
5. J.-Y. Yi and J. Bernholc, J. Chem. Phys. **96**, 8634 (1992).
6. R. L. Murry, D. L. Strout, G. K. Odom, and G. E. Scuseria, Nature **366**, 665 (1993).
7. R. L. Murry, D. L. Strout, and G. E. Scuseria, Int. J. Mass Spectrom. Ion Processes **138**, 113 (1994).
8. B. R. Eggen, M. I. Heggie, G. Jungnickel, *et al.*, Science **272**, 87 (1996).
9. Y. Kumeda and D. J. Wales, Chem. Phys. Lett. **374**, 125 (2003).
10. H. F. Bettinger, B. I. Yakobsen, and G. E. Scuseria, J. Am. Chem. Soc. **125**, 5572 (2003).
11. C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho, J. Phys.: Condens. Matter **4**, 6047 (1992).
12. L. A. Openov and V. F. Elesin, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 695 (1998) [JETP Lett. **68**, 726 (1998)].
13. V. F. Elesin, A. I. Podlivaev, and L. A. Openov, Phys. Low-Dim. Struct. **11–12**, 91 (2000).
14. I. V. Davydov, A. I. Podlivaev, and L. A. Openov, Fiz. Tverd. Tela (St. Petersburg) **47**, 751 (2005) [Phys. Solid State **47**, 778 (2005)].
15. H. S. Chen, A. R. Kortan, R. C. Haddon, *et al.*, Appl. Phys. Lett. **59**, 2956 (1991).
16. A. A. Quong, M. R. Pederson, and J. L. Feldman, Solid State Commun. **87**, 535 (1993).
17. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes* (Academic, San Diego, 1996).
18. C. Xu and G. E. Scuseria, Phys. Rev. Lett. **72**, 669 (1994).

Translated by V. Makhlyarchuk

JETP Letters, Vol. 81, No. 10, 2005, pp. 538–541. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 81, No. 10, 2005, pp. 661–664.
Original Russian Text Copyright © 2005 by Tarasov, Kuz'min.

Concept of a Mixer Based on a Cold-Electron Bolometer

M. Tarasov[1, 3] and L. Kuz'min[2, 3]

[1] *Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, 125009 Russia*
e-mail: tarasov@hitech.cplire.ru
[2] *Institute of Nuclear Physics, Moscow State University, Moscow, 119992 Russia*
[3] *Chalmers University of Technology, SE41296 Gothenburg, Sweden*

Received December 27, 2004; in final form, April 25, 2005

A phase-sensitive terahertz heterodyne mixer of a new type based on a cold-electron bolometer is proposed. In this mixer, a normal-metal thin-film absorber is connected to a planar antenna via superconductor–insulator–normal metal (SIN) tunnel junctions, thus forming a SINIS structure. The SINIS mixer combines the advantages of a hot-electron bolometer (HEB), such as a high signal frequency at a small local oscillator power, with the advantages of an SIS mixer, including low noise level, a high intermediate frequency, and wide working temperature range (up to a critical temperature of the superconductor). In contrast to the HEB and SIS mixers, the proposed device is less sensitive to external magnetic noise and exhibits no additional noise related to the superconducting transition and the Josephson effect. *© 2005 Pleiades Publishing, Inc.*

PACS numbers: 74.50.+r, 85.25.Pb

Introduction. At present, high-sensitivity superconductor mixers for the terahertz frequency range are of two main types, which are based on a superconductor–insulator–superconductor (SIS) structure and a hot-electron bolometer (HEB). The SIS mixer allows a noise temperature lower than 100 K (which is equivalent to several quantum noise limits $T_n^q = hf/k$) to be reached at frequencies below 1 THz [1]. However, above 700 GHz (which corresponds to the energy gap for niobium) or twice as high a frequency (for niobium nitride), the noise temperature of SIS mixers increases sharply. In HEB mixers [2], the noise temperatures at frequencies above 1 THz are on the order of 1000 K and can be approximately evaluated as $T_n = 10hf/k$ [3]. The principle of the HEB mixer operation is substantially different from that of traditional modulation type mixers (such as the SIS mixer), in which the conductivity of a nonlinear element is modulated at a local oscillator (LO) frequency. The conductivity of an HEB mixer cannot be modulated at such a high LO frequency because this structure is too inertial and cannot follow a terahertz heterodyne frequency. Instead, this slow mixer detects an average interference signal, this resembling the situation in a Fourier spectrometer, where an acoustooptical detector or some other slow detector measures an interference signal at the output of a Michelson interferometer. Both SIS and HEB type mixers have their own advantages and drawbacks with respect to the frequency range, sensitivity to thermal and magnetic fluctuations, excess noise, and losses.

Recently, a cold-electron bolometer (CEB) was proposed [4] and successfully implemented as the incoherent detector with a noise-equivalent power (NEP) of 10^{-18} W/Hz$^{1/2}$. This device offers a number of advantages, the most important being the possibility of electron cooling, which provides an increase in the response signal and a decrease in the noise level. This possibility opens the way to the creation of low-noise detectors of both incoherent and phase-sensitive types. Such a CEB mixer design combines the elements of SIS and HEB structures, comprising a thin-film absorber connected to a planar antenna via superconductor–insulator–normal metal (SIN) tunneling junctions. In contrast to an incoherent CEB operating at millikelvin temperatures, CEB mixers operate in the region of conventional helium temperatures. This circumstance increases the response frequency from 10–100 MHz up to 1–10 GHz, which corresponds to the standard heterodyne frequencies.

Bolometric mixer operation principle. A power mixer can be considered as the combination of a power detector and an interferometer, in which the interference of the detected and heterodyne signals leads to the appearance of an interference component of the detected power. In the typical case, whereby the LO power is much higher than that of the signal to be detected, we can describe the output signal as

$$E(t) = \cos\omega t + \alpha\cos(\omega+\delta)t$$
$$= (1-\alpha)\cos\omega t + \alpha[\cos\omega t + \cos(\omega+\delta)t] = E1 + E2.$$

Let us consider the second term in the case when the two amplitudes are equal:

$$\cos\omega t + \cos(\omega+\delta)t = 2\cos\frac{2\omega+\delta}{2}t\cos\frac{\delta}{2}t. \quad (1)$$

This interference signal, representing a combination of two waves, heats the absorber. The total power absorbed in the bolometer at an arbitrary ratio of the component powers is

$$P = \overline{\frac{E_1^2}{R} + \frac{E_2^2}{R}}$$

$$= \overline{\frac{(1-\alpha)^2}{R}\cos^2\omega t} + \overline{\frac{4\alpha^2}{R}\cos^2\left(\frac{2\omega+\delta}{2}t\right)\cos^2\frac{\delta}{2}t}$$

$$= \frac{(1-\alpha)^2}{R}\left(\frac{1-\cos 2\omega t}{2}\right)$$

$$+ \frac{4\alpha^2}{R}\left(\frac{1-\cos(2\omega+\delta)t}{2}\right)\left(\frac{1-\cos\delta t}{2}\right).$$

Averaging at the signal frequency yields

$$P(t) = \frac{(1-\alpha)^2}{2R} + \frac{\alpha^2}{R} - \frac{\alpha^2}{R}\cos\delta t. \quad (2)$$

In the case of equal signal and LO powers, this expression simplifies to

$$P(t) = (\alpha^2/R)(1-\cos\delta t). \quad (3)$$

This formula shows that the absorbed power varies from zero to a certain maximum and the average power is the sum of two equal initial components: $P = \alpha^2/2R + \alpha^2/2R = \alpha^2/R$.

If the LO power is greater than the signal power, the interference term is automatically formed with only a fraction of the former power corresponding to equal amplitudes. This implies that the LO power can be as small as the signal power. In order to provide for a sufficiently large dynamical range, the LO power should be increased to a level corresponding to the maximum expected signal.

CEB as a mixer. In the case of operation at millikelvin temperatures, the CEB response is rather slow (not exceeding several megahertz), since the electron–phonon interaction sharply drops with decreasing temperature. As the phonon temperature grows to a helium level, the interaction increases and its power can be expressed as

$$P_{ep} = \Sigma v(T_e^5 - T_{ph}^5), \quad (4)$$

where Σ is the material constant, v is the volume, T_{ph} is the phonon temperature, and T_e is the electron temperature. The effective thermal relaxation time constant in the absence of an electrothermal feedback (electron cooling) can be evaluated using the simple relations $\tau_0 = C_v/G_{ep}$, $\tau = \tau_0/(L+1)$, where $C_v = v\gamma T_e$ is the absorber heat capacity, $G_{ep} = 5\Sigma T_e^4$ is the electron–phonon thermal conductivity, and $L = G_{cool}/G_{ep}$ is the electrothermal feedback gain. Numerical estimates of the thermal time constant yield 10 μs at 100 mK and 150 ps at 4.2 K. The electron cooling reduces this value by a factor of 10–100, depending on the cooling efficiency. As a result, the intermediate frequency cutoff may reach 10 GHz, which corresponds to the working range of intermediate frequency amplifiers used in coherent receivers for radio astronomy, where the noise temperature does not exceed 10 K.

The power conversion factor can be evaluated proceeding from the basic relations for the electron cooling: if the input heating signal power P_{sig} is fully compensated by electron cooling, the electron temperature remains constant. Since the energy removed per electron of the cooling current is kT, we have $P_{cool} = P_{sig} = kT\Delta I/e$ or $I = eP_{sig}/kT$. The power dissipated in the intermediate frequency load can be estimated as resulting from simple Joule heating by the intermediate-frequency current at a bias voltage close to the gap voltage. For two SIN junctions connected in series, this yields $P_{IF} = 2V_\Delta\Delta I = 2eV_\Delta P_{sig}/(kT)$. Then, the power gain is

$$G = P_{IF}/P_{sig} = 2eV_\Delta/kT = eV_{2\Delta}/kT. \quad (5)$$

This value differs from a minimum of 3 dB for the losses of any classical mixer. Then, the possible power gain for a niobium bolometer at helium temperatures can be estimated at $G_{mix} = 9$. Thus, we have a significant gain rather than a loss.

The noise characteristics can be evaluated to the first approximation proceeding from a shot noise of the SIN junction at the mixer output. We take the input noise to be zero (for the input signal, the mixer is simply a matched resistive load equal to the resistance of the metallic thin-film absorber). The current noise at the output, $I_n^2 = 2eI\Delta f$, can be converted into power as $P_n \approx 2eIR\Delta f \approx 2eV_\Delta\Delta f = kT_n\Delta f$. This yields the output noise temperature $T_n^{out} = eV_{2\Delta}/k$, which is about 30 K for niobium. Taking into account the above power gain of $G = eV_{2\Delta}/kT$, we obtain a very optimistic estimate for the noise temperature converted to input: $T_n = T$.

Another important advantage of the CEB mixer is that matching of the input signal power is achieved much simpler than that in SIS mixers. Indeed, in our case, there is no need to compensate for the intrinsic capacitance of the tunnel junctions because the impedance due to this capacitance is much lower than the real impedance (resistance) of the metal absorber, which can be readily made equal to the real impedance of the complementary planar antenna on a dielectric substrate (amounting approximately to 70 Ω). Thus, the exact impedance matching is provided in a broad frequency range. The matching at an intermediate frequency is much like that in SIS structures and has been well developed for SIS mixers.

In order to evaluate the resistances and capacitances of niobium SINIS mixers, we can use the standard char-

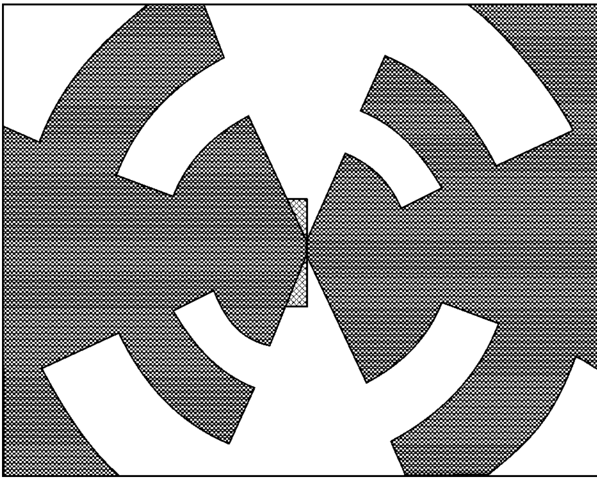

Fig. 1. General view of the SINIS bolometer integrated with a log-periodic planar antenna (bright field corresponds to metallization layers).

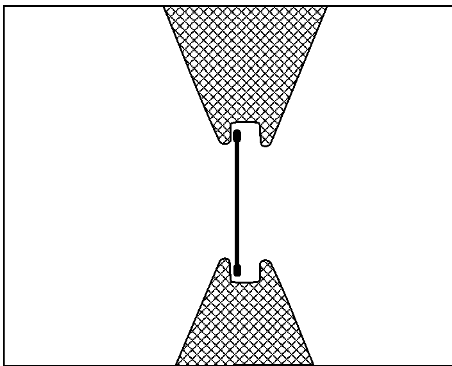

Fig. 2. Central part of a planar antenna with the SINIS bolometer.

acteristics of SIS junctions with aluminum oxide barriers [5]. These structures are characterized by the product R_nA (R_n is the normal resistance and A is the area of the junction) within 25–30 Ω μm^2, which corresponds to a current density of 7–8 kA/cm^2 and a specific capacitance of 70 fF/μm^2. Thus, for a junction area of 1 μm^2, the resistance can be about 30 Ω and the capacitance, about 70 fF. At an intermediate frequency of 4.5 GHz, the capacitive impedance amounts to about 500 Ω and the normal resistance, to 30 Ω. The absorber resistance to be matched with a planar antenna at a given signal frequency is about 70 Ω. As a result, the matching at the intermediate frequency has to be provided between a 130-Ω impedance of the SINIS structure and a 50-Ω impedance of the standard cooled intermediate frequency amplifier, which is readily achieved using a quarter-wave coplanar matching transformer with a characteristic impedance of 80 Ω. By implementing the tunnel junction technology [6] and using an AlN barrier with R_nA = 1 Ω μm^2, one can completely eliminate the problem of matching at the intermediate frequency. Indeed, a contribution due to the serial resistance of the tunnel junction in this case decreases down to several ohms, and the impedance at the intermediate frequency is determined by the absorber resistance that can be taken equal to 70 Ω for matching to the planar antenna at the signal frequency.

CEB development and investigation. In order to study the limiting characteristics of CEBs, we have developed the first generation of such devices using a standard technology for the thermal deposition of tunnel junctions via a suspended two-layer resist mask. An aluminum film was deposited at an angle to the substrate and then oxidized so as to obtain a tunneling barrier. A copper absorber stripe with a length of 10 μm was deposited perpendicularly to the substrate. The sample structures were formed on oxidized silicon substrates by means of electron-beam lithography followed by thermal deposition via a mask. In the first step, a 160-nm-thick Cr–Au–Pd layer was deposited in order to obtain contact pads and planar antennas. Then, the bolometer structure proper was formed via a two-layer PMMA copolymer mask in a single technological cycle. A bottom 70-nm-thick superconducting aluminum layer was obtained by thermal deposition at a rate of 0.3 nm/s. The tunneling barrier was formed by thermal oxidation for 2 min in oxygen at a pressure of 10 Pa. Finally, a 70-nm-thick copper absorber film was deposited on top of the structure. As a result, the tunnel junctions were formed in the regions of overlap between the oxidized aluminum electrodes and the absorber film. The junction areas were 0.5×0.7 and 0.2×0.3 μm^2.

The bolometers were integrated with log-periodic and double dipole antennas. The general view of the typical SINIS structure with a log-periodic antenna is presented in Fig. 1. Figure 2 shows a magnified image of the tunneling junction. A limiting NEP value of 10^{-18} W/Hz$^{1/2}$ was determined by dc measurements in a cryostat at 50 mK. The current, voltage, and spectral responses were measured at a temperature of 280 mK using a cryostat with an optical window. As can be seen from the spectrum presented in Fig. 3, the proposed SINIS bolometer is capable of detecting radiation in the submillimeter wavelength range with frequencies up to about 2 THz, which is significantly higher than the frequency corresponding to the energy gap for aluminum. The NEP at 280 mK was 2×10^{-17} W/Hz$^{1/2}$. The sample topology and the procedure of measurements are described in more detail elsewhere [7]. The time constant of aluminum samples measured using the input signal modulation technique was on the order of a microsecond at 280 mK. The characteristic time of the electron–phonon relaxation in copper is usually estimated using the relation $\tau_{ep} = 20/T^3$, according to which

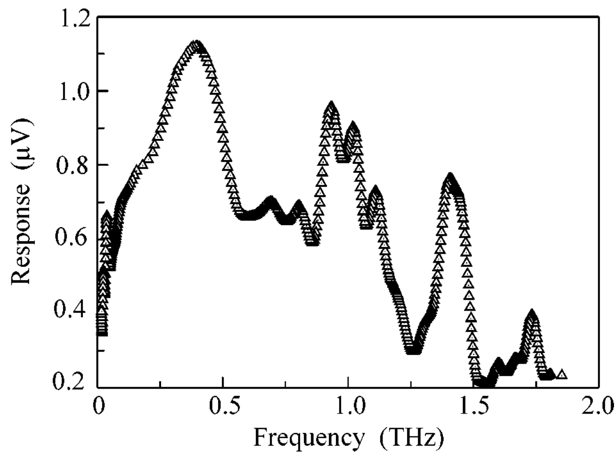

Fig. 3. Spectral response of the bolometer in the terahertz frequency range.

τ_{ep} is about 1 μs at 300 mK and decreases to 0.36 ns at 4.2 K. In order to increase the intermediate frequency in CEB mixers of the second generation, we intend to switch from Al–AlOx–Cu to Nb–AlOx–Pd structures and measure their characteristics at liquid helium temperature (4.2 K). These changes will allow us to increase the intermediate frequency above 1 GHz. According to our estimates, by selecting the absorber material, changing the film thickness, and introducing electrothermal feedback, it will be possible to further increase the intermediate frequency up to 10 GHz.

In conclusion, we have experimentally demonstrated the implementation of a SINIS structure as an incoherent detector. Based on an analysis of the bolometer characteristics, we propose to use CEBs as coherent detectors of radiation in the terahertz frequency range. These bolometers reduce the level of requirements to a cryogenic system and offer advantages over the existing SIS and HEB mixers with respect to the signal frequency, response intensity, and noise level.

This work was supported in part by INTAS (grant no. 01-686), Vetenskapsradet (Swedish Research Council), and Kungl Vetenskapsakademien (Royal Swedish Academy of Sciences).

REFERENCES

1. A. Karpov, D. Miller, F. Rice, *et al.*, Proc. SPIE **5498**, 616 (2004).
2. E. M. Gershenzon, M. E. Gershenzon, G. N. Gol'tsman, *et al.*, Zh. Tekh. Fiz. **59** (2), 111 (1989) [Sov. Phys. Tech. Phys. **34**, 195 (1989)].
3. T. M. Klapwijk, R. Barends, J. R. Gao, *et al.*, Proc. SPIE **5498**, 129 (2004).
4. L. Kuzmin, I. Devyatov, and D. Golubev, Proc. SPIE **3465**, 193 (1998).
5. L. Filippenko, S. Shitov, P. Dmitriev, *et al.*, IEEE Trans. Appl. Supercond. **11**, 816 (2001).
6. H. Dmitriev, I. Lapitskaya, L. Filippenko, *et al.*, IEEE Trans. Appl. Supercond. **13**, 107 (2003).
7. M. Tarasov, L. Kuz'min, E. Stepantsov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **79**, 356 (2004) [JETP Lett. **79**, 298 (2004)].

Translated by P. Pozdeev