JETP Letters, Vol. 78, No. 3, 2003, pp. 101–105. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 131–136.

Pions in Primary Cosmic Rays of Ultrahigh Energies

L. G. Dedenko[1,*], T. M. Roganova[2], G. F. Fedorova[2], and E. Yu. Fedunin[1]

[1] *Faculty of Physics, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*

** e-mail: ddn@dec1.sinp.msu.ru*

[2] *Institute of Nuclear Physics, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*

Received April 24, 2003; in final form, June 25, 2003

In the framework of the Coleman–Glashow hypothesis of an extremely weak violation of Lorentz invariance, neutral and charged pions can be stable for energies above 10^{19} eV and enter into the composition of primary cosmic rays of ultrahigh energies. The kinematic exclusion of reactions of pions with relic photons is particularly important, because it allows the Greisen–Zatsepin–Kuzmin paradox to be resolved. The parameters of extensive air showers induced by primary pions calculated within the model of quark–gluon strings with allowance for the Landau–Pomeranchuk–Migdal effect and interactions of neutral pions of ultrahigh energies are not contradictory to the available data of observations. It has been shown that observations of production heights of muons with energies above 10 GeV will make it possible to distinguish between primary nuclei, protons, and pions; to verify Lorentz invariance for energies above 10^{20} eV; and to obtain a new limit on the difference between the maximum possible velocities of muons and pions $(c_\mu - c_\pi) < 4 \times 10^{-26}$. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 13.85.Tp; 96.40.Pq

Based on the first observations of giant air showers (GASs) at Volcano Ranch Array [1] and further investigations at Haverah Park [2], Yakutsk [3], AGASA [4], SUGAR [5], Fly's Eye [6], and Hires [7] arrays, a very important discovery was made, namely, GASs with energies above 10^{20} eV were detected. This discovery is difficult to reconcile with the famous Greisen–Zatsepin–Kuzmin (GZK) effect [8, 9]. The accuracies of the estimates of both the energy E_0 of the particles that induce GASs and their arrival directions on the celestial sphere are of primary importance. The energy of the inclined GAS detected at the Yakutsk array was first estimated at $\sim 1.2 \times 10^{20}$ eV [3]. A second analysis of the parameters of this shower with allowance for the deviations of muons in the Earth's magnetic field provided an estimate of $\sim 3 \times 10^{20}$ eV [10]. The revision of the Haverah Park data reduced the estimated energy by about 30% [11]. As a result, none of the four showers with energies above 10^{20} eV remained in these data. We note that the magnetic field effect was disregarded. The revision of Fly's Eye and Hires data [12] reduced the number of the detected showers with energies above 10^{20} eV from eight to one. However, this revision cannot be considered final because of uncertainties both in the atmospheric state and in the intensity of 391-nm fluorescence light and also because of the use of the data in monomode rather than stereomode. Thorough analysis of all factors that can lead to both an overestimation and an underestimation of the energy showed that eleven showers with energies above 10^{20} eV were detected at the AGASA, where two showers have energies above $(2–3) \times 10^{20}$ eV [13]. Thus, more than 10 (possibly about 20, according to all data) showers with energies above 10^{20} eV were observed, which contradicts the GZK effect, because close sources of particles with such energies have not yet been found. The analysis made in [14] showed that the arrival directions of 20 intense showers detected at the Yakutsk array are uniformly distributed over the celestial sphere. A similar result was also obtained at the AGASA array [15]. The arrival directions of GASs detected at the Fly's Eye and Hires arrays also do not point to particular close sources [16]. Moreover, Farrar and Biermann [17] and Virmani *et al.* [18] found that the arrival directions of, respectively, five and eleven showers with energies above 10^{19} eV correlate with quasars at cosmological distances from the Earth. In addition, a correlation with BL Lacertae also located at enormous distances from the Earth was found in [19]. We point in particular to the observations of doublets and triplets (pairs and triples) of showers arriving at the Earth from the same direction [20]. These observations can be treated as a possible manifestation of pointlike sources [21]. The identification of the arrival directions of showers with remote objects and the observations of doublets and triplets imply that the showers are induced by neutral particles. Indeed, in stochastic intergalactic magnetic fields with intensity B and length D, a relativistic charged particle with energy E passing distance d is deflected at the angle

$$\delta\Theta = 15^\circ \left(\frac{d}{d_0}\right)^{1/2} \left(\frac{D}{D_0}\right)^{1/2} \left(\frac{B}{B_0}\right) \left(\frac{E_0}{E}\right), \qquad (1)$$

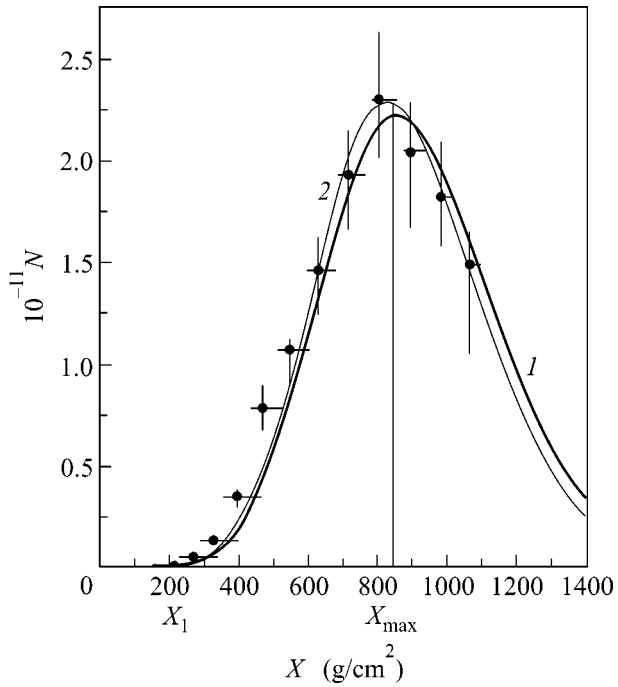


Fig. 1. Cascade curves for pion-induced GASs with $E = 3 \times 10^{20}$ eV. Experimental points with error bars are taken from [6]. Lines *1* and *2* are calculated under assumptions (i) and (ii), respectively [30].

where the numerical factor is calculated for $d_0 = 1$ Mpc, $D_0 = 16$ Mpc, $B_0 = 10^{-9}$ G, and $E_0 = 10^{19}$ eV. Simple estimates show that $\delta\Theta \sim 15^\circ$ for $E = 3 \times 10^{19}$ eV and $D = 160$ Mpc. Therefore, an additional analysis is likely necessary for charged particles. The hypothesis of neutrinos with ultrahigh energies, which interact with relic neutrinos near the Earth [22, 23] and induce "Z-boson" showers, is the simplest assumption that involves neutral primary particles and can resolve both the problems of identification with remote objects and the GZK paradox. However, because of certain restrictions, the sources of such neutrinos must have very specific properties [24]. Since the shower observed in [3] consists of only muons and the showers detected in [4] also contain muons, the hypothesis of primary gamma rays must be analyzed with caution, even though it does not contradict the Fly's Eye, HiRes, and, in part, AGASA data [25]. At the same time, beginning with [26], the concept of violation of Lorentz invariance was involved to explain the GZK paradox (see review [27]). In the framework of the hypothesis that Lorentz invariance is very weakly violated, Coleman and Glashow [28] showed that neutral pions and neutrons can be neutral particles of primary cosmic rays. The parameters of this hypothesis for primary neutrons were refined in [29]. In [30], using the data of observations [6], we obtained limits on the parameters violating Lorentz invariance. In this work, in the framework of the Coleman–Glashow hypothesis of a very weak violation of Lorentz invariance, we analyze the observable consequences of the presence of pions in primary cosmic rays of ultrahigh energies. Calculations are carried out in the model of quark–gluon strings (QGS) [31] and include the Landau–Pomeranchuk–Migdal effect [32] and interactions of neutral pions of ultrahigh energies. As was shown in [28], neutral pions are stable for energies above the threshold

$$E = m_{\pi^\circ}/\sqrt{c_\gamma^2 - c_{\pi^\circ}^2}, \tag{2}$$

where c_γ and c_{π° are the maximum possible velocities of gamma rays and neutral pions, respectively. As in [30], we analyze the assumptions that (i) $E > 10^{19}$ eV and $c_\gamma - c_{\pi^\circ} = 10^{-22}$ and (ii) $E > 10^{18}$ eV and $c_\gamma - c_{\pi^\circ} = 10^{-20}$. To propagate at cosmological distances, neutral pions must obviously not interact with relic photons. The reaction

$$\pi^\circ + \gamma \longrightarrow \omega(782), \tag{3}$$

whose threshold in the standard theory

$$E_{\rm thr} = (m_{\omega(782)}^2 - m_{\pi^\circ}^2)/4\omega_\gamma, \tag{4}$$

where ω_γ is the photon energy, differs from the threshold for protons by 7%, can be kinematically forbidden under the assumptions accepted in [28] if

$$\begin{aligned} c_{\omega(782)} - c_{\pi^\circ} &> 2\omega_\gamma^2/(m_{\omega(782)}^2 - m_{\pi^\circ}^2) \\ &\approx 1.86 \times 10^{-25}(\omega_\gamma/\omega_0)^2, \end{aligned} \tag{5}$$

where $c_{\omega(782)}$ is the maximum possible velocity of the $\omega(782)$ meson and $\omega_0 = 2.35 \times 10^{-4}$ eV is the characteristic thermal energy of photons. Similarly, under the assumptions accepted in [28], charged pions can be stable for energies above the threshold (disregarding the neutrino mass)

$$E_p = \sqrt{\frac{m_\pi^2 - m_\mu^2}{c_\mu^2 - c_\pi^2}} \approx 6.45 \times 10^{19}\left(\frac{10^{-24}}{c_\mu - c_\pi}\right)^{1/2}, \tag{6}$$

where c_π and c_μ are the maximum possible velocities of pions and muons, respectively; E_p is measured in electron volts; and we take $c_\mu = c_\nu$ for simplicity. Charged pions will be able to propagate at cosmological distances if the reaction

$$\pi^\pm + \gamma \longrightarrow \omega^\pm(770) \tag{7}$$

is kinematically forbidden. This is the case under the condition

$$\begin{aligned} c_{\omega(770)} - c_{\pi^\pm} &> 2\omega_\gamma^2/(m_{\omega(770)}^2 - m_\pi^2) \\ &\approx 0.963 \times 10^{-25}(\omega_\gamma/\omega_0)^2 \end{aligned} \tag{8}$$

similar to condition (5).

First, we note that the calculated parameters of pion-induced GASs are consistent with the data of observations. Figure 1 shows cascade curves calculated for pions under hypotheses (i) and (ii) and the data of observations [6]. It is seen that the cascade curve for an individual shower induced at a sufficiently high altitude in the atmosphere agrees closely with the experimental data for one event. The depths X_{max} and X_1 of the shower maximum and the first interaction, respectively, are also indicated in this figure. Figure 2 shows (lines *1* and *2*) the primary-pion energy dependence of the average depth of the shower maximum and the data for the event observed in [6]. As is seen, the parameters of the shower induced by a primary pion are consistent with the data of observations within the accuracy. The dashed and dash–dotted lines in Fig. 2 are calculated for primary protons and iron, respectively, in the QGS model and standard variant. Since the data [6] agree with the calculated depth of the shower maximum induced by a primary proton within one standard deviation, an important limit on the parameters violating Lorentz invariance can be obtained. To exclude showers induced by primary neutral pions from relatively close sources, it is necessary to impose the limit $(c_\gamma - c_{\pi^\circ}) < 10^{-22}$ [30]. To exclude showers induced by primary charged pions, the new limit $(c_\mu - c_\pi) < 4 \times 10^{-26}$ is necessary, because the shower energy is equal to $E_0 \approx 3 \times 10^{20}$ eV. If sources are at cosmological distances, it is necessary to impose additional limits $(c_{\omega(782)} - c_{\pi^\circ}) < 2 \times 10^{-23}$ and $(c_{\omega(770)} - c_\pi) < 10^{-23}$ for neutral and charged pions, respectively.

Parameters that can distinguish between showers induced by pions, protons, and nuclei are obviously of interest. Since the mean free paths for a given energy are different for different particles, it is reasonable to compare the distributions of these mean free paths. These distributions can be obtained as follows. First, the distribution of the depth X_{max} of shower maxima can be obtained from both observations and calculations. Figure 3 shows these distributions for (a) protons and (b) pions under assumptions (*1*) (i) and (*2*) (ii). The dotted histogram in Fig. 3a is the calculation for protons in the standard variant of the model. If X_1 is the depth at which the first interaction occurs, we can determine the quantity $\Delta X = X_{max} - X_1$. This relation obviously enables one to determine the depth X_1 from the known X_{max} and ΔX as

$$X_1 = X_{max} - \Delta X. \quad (9)$$

As was mentioned above, the depth X_{max} is determined from observations for each shower. The quantity ΔX can be taken from calculations. Figure 4 shows the calculated distributions of ΔX for primary (a) protons and (b) pions (notation is the same as in Fig. 3). As is seen, the standard deviations are equal to 1–1.5 and 10 g/cm^2 for pions and protons, respectively. Therefore, in the

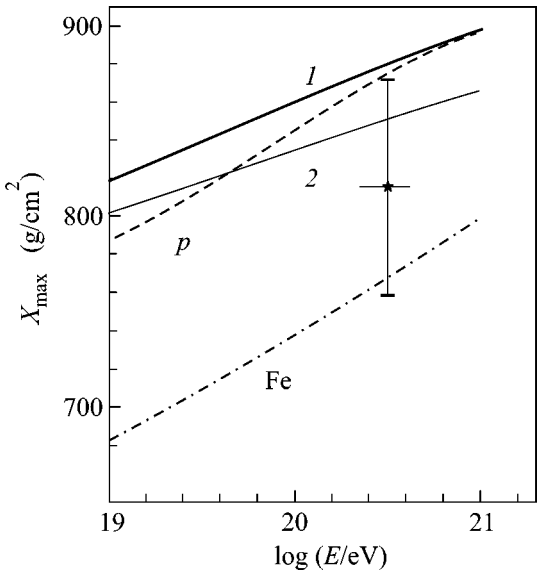


Fig. 2. Energy dependence of the depth X_{max} for (dashed and dash–dotted lines for primary protons and iron, respectively) standard variant and under the assumptions (*1*) (i) and (*2*) (ii) for pions [30].

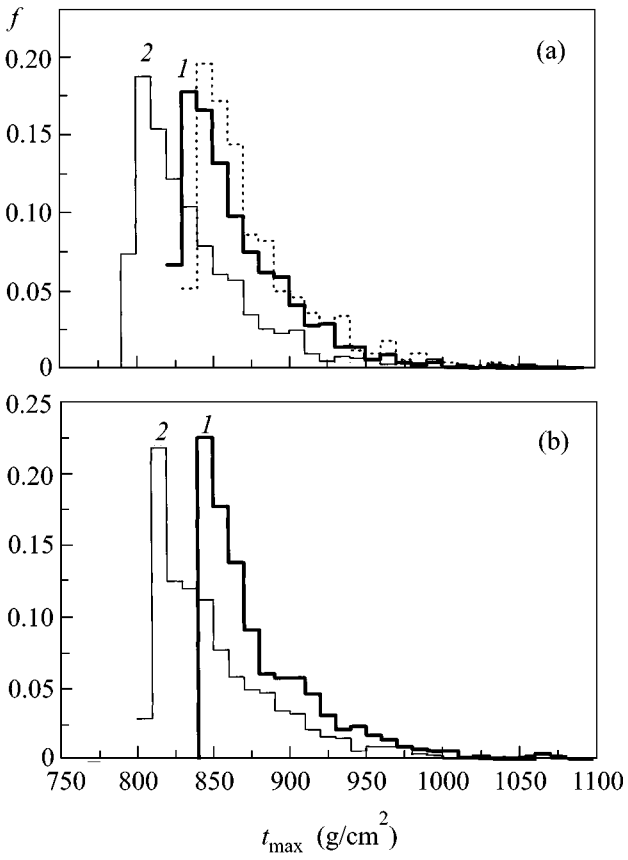


Fig. 3. Distributions of the depth X_{max} for (a) protons and (b) pions; the dotted histogram corresponds to the standard variant for protons, and histograms *1* and *2* are for assumptions (i) and (ii), respectively [30].

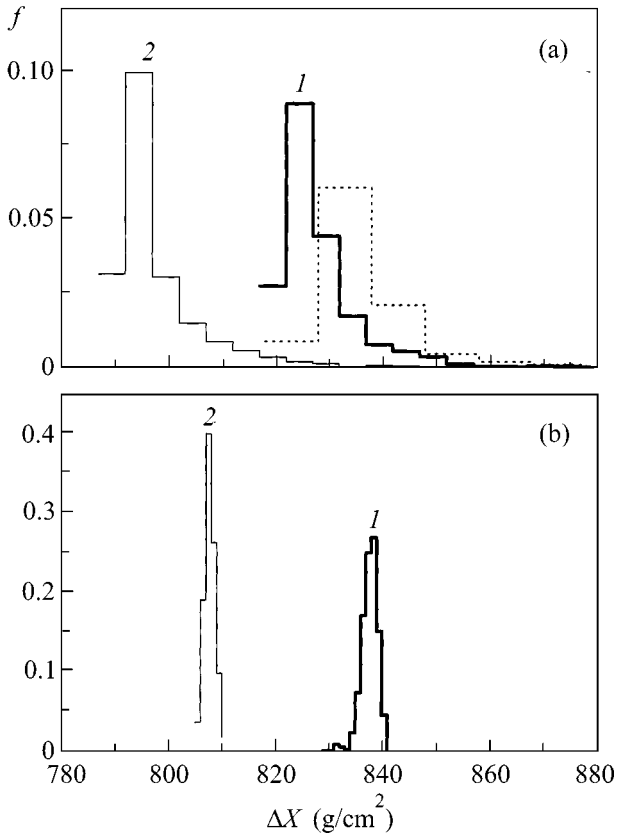


Fig. 4. Same as Fig. 3, but for the depth ΔX.

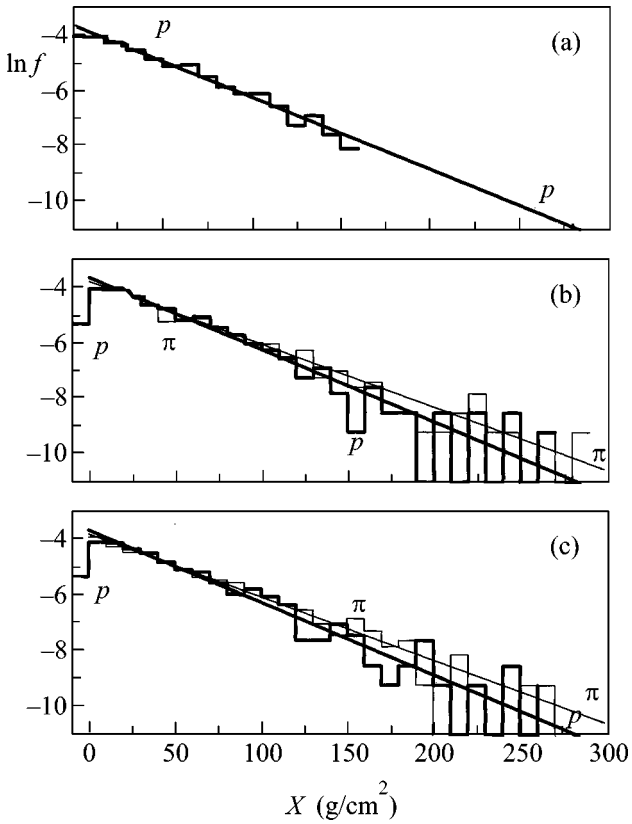


Fig. 5. Distributions of the depth X_1 for (thick histograms and lines) protons and (thin histograms and lines) pions as obtained with the (a) standard variant and under assumptions (b), (c) (i) and (ii) [30].

framework of the model under consideration, ΔX can be taken a constant equal to, e.g., its average value $\Delta\bar{X}$. Figure 5 shows calculated distributions of the depth X_1 estimated by Eq. (9) and the $e^{-x/\lambda}$ distribution (a) for primary protons in the standard variant and (b), (c) for protons and pions under assumptions (i) and (ii). Since the mean free paths of protons $\lambda_p = 38$ g/cm^2 and pions $\lambda_\pi = 40.4$ g/cm^2 for energy $E_0 \approx 3 \times 10^{20}$ eV are close to each other, the distributions for pions and protons in Figs. 5b and 5c virtually coincide with each other. However, although this figure does not discriminate between protons and pions, the proposed method can be used to test the QGS model by estimating the interaction cross section for ultrahigh energies.

Finally, the distribution of the generation altitudes of high-energy muons ($E > 10$ GeV) is particularly interesting and distinguishes between different primary particles. Figure 6 shows these distributions for the standard model with primary (dotted line) iron and (dashed line) protons and under assumptions (i) and (ii) with (lines *1p* and *2p*, respectively) protons and (lines 1π and 2π, respectively) pions. As is seen (the calculation error is as small as 10 g/cm^2), the peak in the distribution for protons is located deeper in the atmosphere than the peak for iron by about 60 g/cm^2. This circumstance makes it possible to distinguish between the primary nuclei and protons and to treat more reliably the decrease in the depth $X_{\max}$ of proton-induced showers as the manifestation of a violation of Lorentz invariance [30]. In turn, the peak in the distribution for protons is located higher in the atmosphere than the peak for pions by about 90 g/cm^2, which enables pions to be identified in primary cosmic rays. Methods for determining the distributions of the generation altitudes of high-energy muons become of special importance [33].

In conclusion, we emphasize again that the assumption of the presence of pions in primary cosmic rays is consistent with the available experimental data. The 90-g/cm^2 increase in the depth of the peak in the distribution of the generation altitudes of high-energy muons as compared to that in the standard variant for protons is a characteristic marker of pions in primary cosmic rays. At the same time, if the experimental data are interpreted within the standard model the following new limits on the parameters violating Lorentz invariance can be obtained: $(c_\gamma - c_{\pi^\circ}) < 10^{-22}$ and $(c_\mu - c_\pi) < 4 \times 10^{-26}$, as well as extra limits $(c_{\omega(770)} - c_\pi) < 10^{-23}$ and $(c_{\omega(778)} - c_{\pi^\circ}) < 2 \times 10^{-23}$ for the case of remote sources.

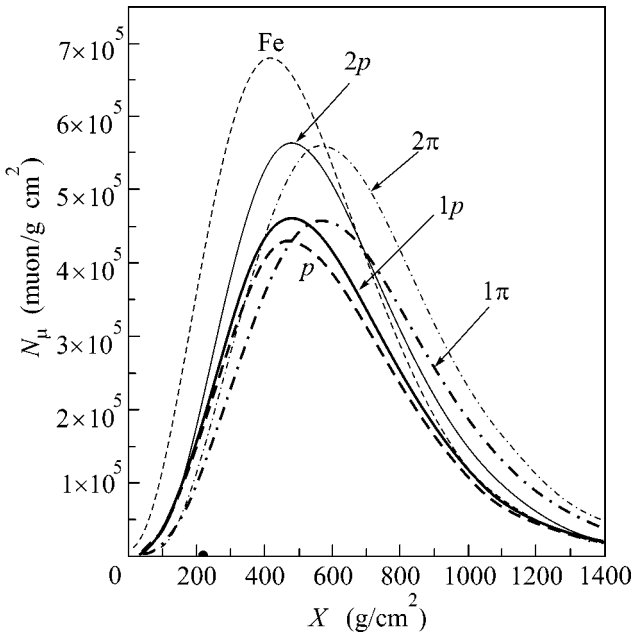


Fig. 6. Distributions of the generation altitudes of high-energy muons: standard variant with primary (dashed line) protons and (dotted line) iron; assumptions (i) and (ii) with (lines *1p* and *2p*, respectively) protons and (lines 1π and 2π, respectively) pions.

We are deeply grateful to G.T. Zatsepin for valuable remarks. This work was supported by the Russian Foundation for Basic Research (project nos. 1782.2003.2 and 03-02-16290; the former is for the support of the Leading Scientific School by Zatsepin).

REFERENCES

1. J. Linsley, Phys. Rev. Lett. **10**, 146 (1963).
2. G. Brooke, G. Cunningham, P. I. Eames, *et al.*, in *Proceedings of the 19th ICRC, La Jolla, California* (1985), Vol. 2, p. 150.
3. N. N. Efimov, T. A. Egorov, A. V. Glushkov, *et al.*, in *Proceedings of International Workshop on Asrtophysical Aspects of the Most Energetic Cosmic Rays, Kofu* (1990), p. 20.
4. N. Hayashida, K. Honda, M. Honda, *et al.*, Phys. Rev. Lett. **73**, 3491 (1994).
5. M. M. Winn, J. Ulrichs, L. Horton, *et al.*, in *Proceedings of the 19th ICRC, La Jolla, California* (1985), Vol. 9, p. 499.
6. D. J. Bird, S. C. Corbato, H. U. Dai, *et al.*, Astrophys. J. **424**, 491 (1994).
7. T. Abu-Zayyad, G. Archbold, K. Belov, *et al.*, in *Proceedings of 26th ICRC, Salt Lake City* (1999), Vol. 3, p. 264.
8. K. Greisen, Phys. Rev. Lett. **16**, 748 (1966).
9. G. T. Zatsepin and V. A. Kuzmin, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 78 (1966) [JETP Lett. **4**, 53 (1966)].
10. E. E. Antonov, A. V. Glushkov, L. G. Dedenko, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 614 (1999) [JETP Lett. **69**, 650 (1999)].
11. M. Ave, J. A. Hinton, J. Knapp, *et al.*, in *Proceedings of the 27th ICRC, Gamburg* (2001), Vol. 1, p. 381.
12. T. Abu-Zayyad, G. C. Archbold, J. A. Bellido, *et al.*, astro-ph/0208301 (2002).
13. M. Takeda, N. Sakaki, K. Honda, *et al.*, astro-ph/0209422 (2002).
14. E. E. Antonov, I. L. Buylova, L. G. Dedenko, *et al.*, Comput. Phys. Commun. **147**, 480 (2002).
15. N. Hayashida, K. Honda, N. Inoue, *et al.*, astro-ph/0008102 (2000).
16. J. W. Elbert and P. Sommers, Astrophys. J. **441**, 151 (1995).
17. G. R. Farrar and P. L. Biermann, Phys. Rev. Lett. **81**, 3579 (1998).
18. A. Virmani, S. Bhattacharya, P. Jain, *et al.*, Astropart. Phys. **17**, 489 (2002).
19. P. G. Tinyakov and I. I. Tkachev, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 499 (2001) [JETP Lett. **74**, 445 (2001)]; A. V. Uryson, astro-ph/0303347.
20. M. Takeda, M. Chikawa, M. Fukushima, *et al.*, in *Proceedings of the 27th ICRC, Gamburg* (2001), Vol. 1, p. 341.
21. P. G. Tinyakov and I. I. Tkachev, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 3 (2001) [JETP Lett. **74**, 1 (2001)].
22. D. Fargion, B. Mele, and A. Salis, Astrophys. J. **517**, 725 (1999).
23. T. J. Weiler, Astropart. Phys. **11**, 303 (1999).
24. O. E. Kalashev, V. A. Kuzmin, D. V. Semikoz, *et al.*, Phys. Rev. D **65**, 103003 (2002).
25. L. G. Dedenko, G. T. Zatsepin, T. M. Roganova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **61**, 233 (1995) [JETP Lett. **61**, 241 (1995)].
26. D. A. Kirzhnits and V. A. Chechin, Pis'ma Zh. Éksp. Teor. Fiz. **14**, 261 (1971) [JETP Lett. **14**, 172 (1971)].
27. G. Amelino-Camelia, J. Ellis, N. E. Mavromatos, *et al.*, Nature **393**, 763 (1998).
28. S. Coleman and S. L. Glashow, Phys. Rev. D **59**, 116008 (1999).
29. S. L. Dubovsky and P. G. Tinyakov, Astropart. Phys. **18**, 89 (2002).
30. E. E. Antonov, L. G. Dedenko, A. A. Kirillov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 506 (2001) [JETP Lett. **73**, 446 (2001)].
31. A. B. Kaĭdalov, K. A. Ter-Martirosyan, and Yu. M. Shabel'skiĭ, Yad. Fiz. **43**, 1282 (1986) [Sov. J. Nucl. Phys. **43**, 822 (1986)].
32. A. B. Migdal, Phys. Rev. **103**, 1811 (1956).
33. G. Agnetta, M. Ambrosio, C. Aramo, *et al.*, Astropart. Phys. **6**, 301 (1997).

Translated by R. Tyapaev

JETP Letters, Vol. 78, No. 3, 2003, pp. 106–109. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 137–140.

Phase Conjugation at the Surfaces of Optically Excited CuI Films

A. N. Gruzintsev

Institute of Problems of Technology of Microelectronics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia

e-mail: gran@ipmt-hpm.ac.ru

Received June 20, 2003

The possibility of realizing an optical phase conjugation in an excited semiconductor medium is shown theoretically and experimentally. A phase conjugation is revealed for the photon energy equal to half the energy of the radiative recombination of excitons in CuI films pumped by a nitrogen laser at room temperature. The dependences of the phase-conjugation signal intensity on its spectral composition are investigated. The quadratic interaction of light and exciton electromagnetic oscillations in the semiconductor medium is suggested as an explanation of this effect. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 42.65.Hw; 78.66.Jg

A wide-gap CuI semiconductor material has the highest exciton binding energy (40 meV) for its class of compounds [1], which allows exciton luminescence to be observed at room temperature. Though different crystalline modifications can occur in polycrystalline CuI films (zinc blende, wurtzite, and hexagonal [2]), all of them have almost the same band gap. An intense exciton peak in the violet region with a maximum at 410 nm (its FWHM is ~10 nm) is almost always observed in the photoluminescence (PL) spectra under excitation by a nitrogen laser [3]. This makes CuI a model material in studies of the physics of excitons and a promising material for creating high-power semiconductor laser sources in the violet spectral region. One of the most important elements in a semiconductor laser is the mirrors of the resonator. They determine its quality factor, the lasing threshold, the laser-radiation directivity, and its spectral half-width.

The use of phase-conjugating mirrors (PCMs) in gas-laser resonators was considered in several works [4–6]. An advantage of this type of mirrors is that the resonator can be self-tuned and PCMs can correct distortions caused by the amplifying medium or by the optical elements in the resonator. Since, in this case, the incident wave is always reflected in the strictly opposite direction, it is unnecessary to observe a stringent parallelism of the resonator mirrors. In these experiments [6] on phase conjugation, the pump power was approximately 1.6×10^7 W/cm^2 and the power of the reflected wave was 2×10^{-3} of the signal-wave power (the mirror reflectivity). In all of the aforementioned versions, phase conjugation is achieved using a three-wave mixing process in a nonlinear crystal or a gaseous medium. Due to this fact, PCMs have a large volume, a low reflectivity, and cannot be used in rather small semiconductor lasers.

In this work, an effect of electromagnetic-wave phase conjugation on excitons in CuI films highly excited by a UV laser was discovered and investigated. In this case, only the process of two-wave mixing of electromagnetic oscillations takes place, whose probability is rather high in noncentrosymmetric electrooptic semiconductor materials with the wurtzite structure.

The expression for the polarization **P** of a semiconductor medium exposed to electromagnetic oscillations E_1 and E_2 with different frequencies Ω and ω can be written in tensor form [7]:

$$P_i = \chi^1_{ij}E_j + \chi^2_{ijk}E_jE_k + \chi^3_{ijkl}E_jE_kE_l + \ldots, \qquad (1)$$

where the terms on the right-hand side are summed over all identical indices (j, k, l = 1, 2, 3) or over three components of the total electric field; and χ^1, χ^2, and χ^3 are the linear, quadratic, and cubic optical susceptibilities of the medium, which, in the general case, are second-, third-, and fourth-rank tensors, respectively. Let us consider in more detail the term with the quadratic susceptibility in Eq. (1). This term is characteristic of crystals without a center of symmetry in the crystal lattice and is responsible for the generation of the second harmonic and an electrooptic effect in materials (the latter has a significant magnitude in CuI with the hexagonal wurtzite lattice) [1]). To simplify our consideration, we assume that the polarization vector **P** and both the electric fields E_1 and E_2 are directed along the crystal-growth axis coinciding with the spatial Z axis. The light wave propagates along the X axis and has the form $E_2 = \{E_{20}\exp(-i\omega t + ikx) + E^*_{20}\exp(i\omega t + ikx)\}$. In this case, the other electromagnetic oscillation in the semiconductor corresponds to the point of an exciton at rest (**K** = 0) in the polariton dispersion curve and has the

0021-3640/03/7803-0106$24.00

form $E_1 = \{E_{10}\exp(-i\Omega t) + E_{10}^*\exp(i\Omega t)\}$. The second term in Eq. (1) then has the form

$$\chi^2(E_1 + E_2)(E_1 + E_2) = 2\chi^2 E_1 E_2 + \chi^2 E_1^2 + \chi^2 E_2^2. \quad (2)$$

Let us examine in detail the first term on the right-hand side of Eq. (2) and expand it assuming that the frequency of the light wave is equal to half the exciton frequency; i.e., $\omega = \Omega/2$:

$$2\chi^2 E_1 E_2 = 2\chi^2\{E_{10}E_{20}\exp(-i3\omega t + ikx) + E_{10}E_{20}^*\exp(-i\omega t - ikx) + \text{c.c.}\}, \quad (3)$$

where c.c. means the complex conjugate terms with respect to the first two terms in Eq. (3). Expression (3) for the semiconductor polarization yields two electromagnetic waves. The first wave does not conform to the crystal dispersion curve, whereas the second one is the phase-conjugated optical wave propagating in the counter direction with respect to the incident wave E_2 regardless of the angle of light incidence onto the semiconductor surface. Its intensity is proportional to the incident-wave intensity, the exciton oscillation intensity, and the quadratic nonlinear optical susceptibility of the medium. If excitons are excited in CuI by UV radiation of a pulsed nitrogen laser, whose absorption depth is ~0.1–0.2 μm, it is easy to obtain the phase-conjugation conditions in a thin near-surface semiconductor layer.

In this work, we used CuI films deposited on quartz substrates by a high-vacuum thermal evaporation technique. The film thickness was 1.5 μm. These CuI films had a polycrystalline structure including the cubic, wurtzite, and hexagonal phases. An exciton peak with a maximum at 410 nm prevailed in their luminescence spectrum at room temperature (see Fig. 2). The PL spectra of films were investigated under excitation by the pulsed radiation of an LGI-505 nitrogen laser with a pulse duration of 10 ns, a peak power of 15 kW, and a wavelength of 337.1 nm. Figure 1 shows a schematic of the setup used in the experiments on optical phase conjugation in a CuI film. It is important to note here that the light wave incident on the sample was produced by a Narva-100 incandescent lamp powered by a dc current source. The angle of incidence was > 45° in order to prevent penetration of the geometrically reflected light into the recording MDR-b monochromator. After being transmitted through a high-luminosity MDR-12 monochromator, the light from the incandescent lamp was continuous in time and had a spectral half-width of <5 nm. The spectrum recording system was tuned to the frequency of the exciting LGI-505 laser and consisted of a Unipan-237 amplifier of alternating signals that were fed to a computer. A CuI sample was kept at room temperature and oriented so that electromagnetic oscillations of laser and light waves had components along the film-growth axis. The laser beam was focused to the film surface into a spot with a diameter of <100 μm, thus allowing us to achieve a pump-power density of ~10^8 W/cm^2. The coincidence of the light and laser spots on the sample was monitored using an optical microscope. An ordinary glass plate was used as a beamsplitter (Fig. 1). When the spectra of the phase-conjugated signal and the film luminescence were recorded, KS-15 and SZS-23 light filters were used, respectively, for eliminating the second-order parasitic signals of the MDR-b double monochromator in the recorded spectra.

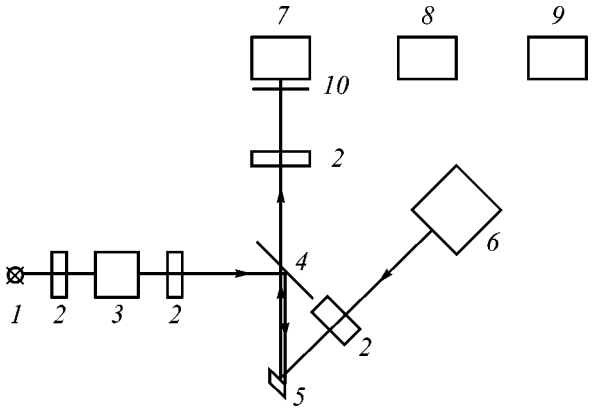


Fig. 1. Schematic of the setup used in experiments on optical phase conjugation on a cuprous iodide film: (*1*) Narva-100 light source; (*2*) lenses; (*3*) MDR-12 monochromator; (*4*) beamsplitter; (*5*) CuI sample; (*6*) LGI-505 laser; (*7*) MDR-6 monochromator; (*8*) FEU-79 photomultiplier tube; (*9*) spectrum-recording system; and (*10*) light filter.

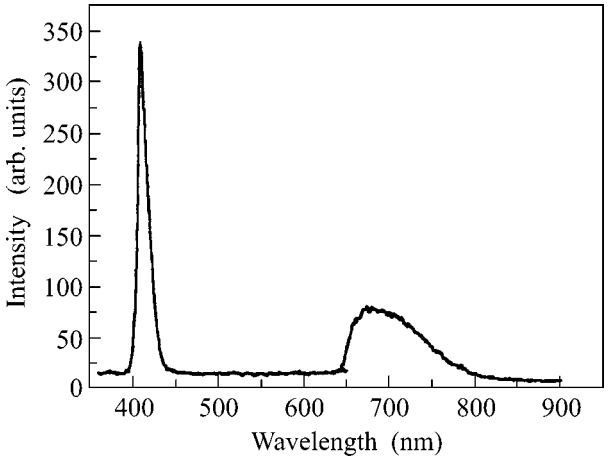


Fig. 2. Photoluminescence spectrum of a CuI film on quartz. The temperature of measurements is T = 300 K.

Figure 2 shows a PL spectrum (curve *1*) of the CuI films under study measured under the laser excitation without additional illumination. The spectrum exhibits an exciton line peaked at 410 nm and a peak at 680 nm in the long-wavelength impurity region. This feature of the spectrum can be associated with radiative recombination at complexes of intrinsic and impurity defects in CuI [2, 3].

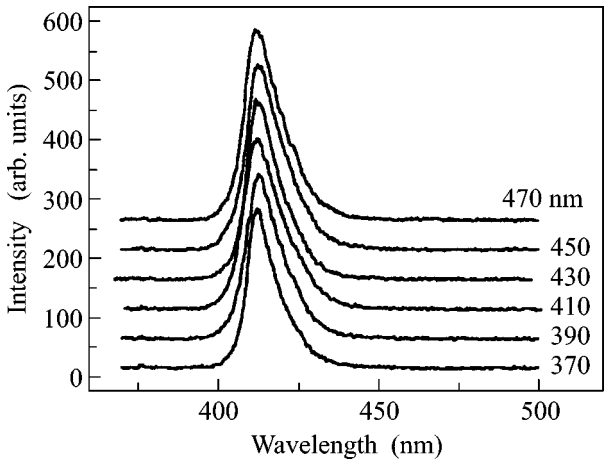


Fig. 3. Spectra of the luminescence and phase-conjugation signal for a CuI film on quartz in the blue–violet spectral region for various wavelengths of illuminating photons. The temperature of measurements is $T = 300$ K.

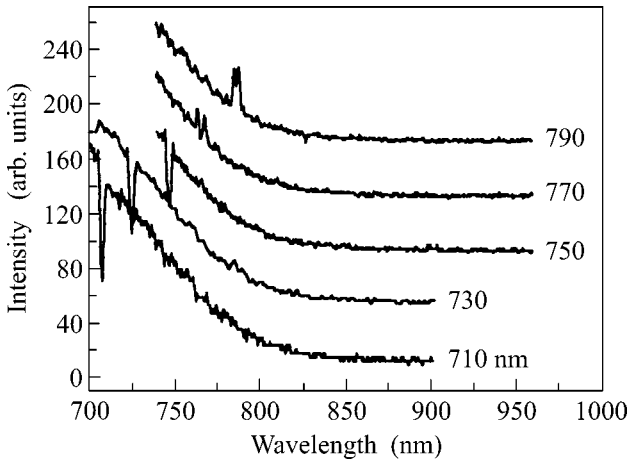


Fig. 4. Spectra of the luminescence and phase-conjugation signal for a CuI film on quartz in the red spectral region for various wavelengths of illuminating photons. The temperature of measurements is $T = 300$ K.

Let us now, simultaneously with pulsed excitation by a nitrogen laser, illuminate the sample with a temporally continuous light flux from the blue–violet spectral region with photon wavelengths varying between 370 and 470 nm (Fig. 3). Figure 3 shows that there is only a film photoluminescence signal peaked at 410 nm in the recorded signal in the blue–violet spectral region. Note that a change in the spectral position of the continuous illumination has no effect on the shape and intensity of the spectral lines.

The pattern of the signal recorded in the red spectral region with the energy of illuminating photons equal to half the energy of the radiative recombination of exciton in CuI (Figs. 4 and 5) is much more interesting. In this case, a KS-15 infrared light filter, which cut off not only the scattered 337.1-nm laser light but also the signal of exciton luminescence from the CuI film, was installed in front of the recording MDR-6 monochromator. This was done in order to cut off the probable second order of the 410-nm luminescence line in the monochromator. We see that, under illumination with 710-nm photons in the red spectral region (curve *1* in Fig. 4), a dip with a minimum at 710 nm is observed against the PL-signal background. The position of this dip depends on the wavelength of the continuous illumination (curves *2* and *3* in Fig. 4). However, as the wavelength of the incident photon approaches the PL band edge (curves *4* and *5*), a narrow peak of the phase-conjugated wave arises instead of the dip. The energy position of this peak coincides with the energy of the photons illuminating the sample. The method for signal recording also shows that, in contrast to the incident wave, this signal has a pulsed nature with the pulse repetition rate of the exciting laser. The spectral half-width of the phase-conjugated signal coincides with the half-width of the incident light wave, and the intensity of the phase-conjugated wave depends on its spectral position. The maximum intensity of the phase-conjugated signal is observed at a 810-nm wavelength of the incident light (curve *1* in Fig. 5). This narrow peak virtually vanishes for wavelengths exceeding 870 nm (curves *5* and *6*).

Figures 4 and 5 show that the intensity of the phase-conjugated signal reaches a maximum at wavelengths of incident photons of 810–830 nm, which almost coincides with the double wavelength of the PL exciton peak (Fig. 2). Note that, according to the theoretical calculations presented above, the phase-conjugation effect can appear for media with a quadratic optical nonlinearity for the energy of incident photons equal to half the energy of exciton oscillations (see Eq. (3)). In this case, the intensity of the phase-conjugated wave is proportional to the intensity of the incident light and the

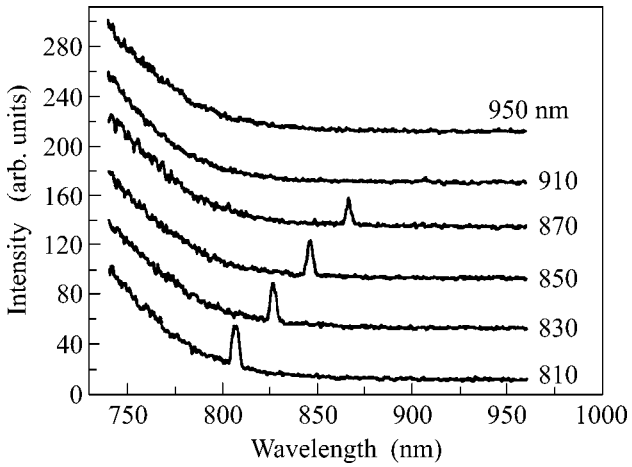


Fig. 5. Spectra of the luminescence and phase-conjugation signal for a CuI film on quartz in the red spectral region for various energies of illuminating photons. The temperature of measurements is $T = 300$ K.

intensity of the exciton wave. Since the radiative-recombination time of excitons in CuI films at room temperature is ~1 ps, exciton oscillations exist only at the moment of the action of a laser pump pulse on the sample. Consequently, the phase conjugation has a pulsed character, which is confirmed by the alternating signal recorded at the laser-pulse repetition rate in our experiments. Because the peak intensity of the phase-conjugated signal exhibits a clearly pronounced spectral dependence (curve *2* in Fig. 2), the process observed in our study cannot be the scattering of the continuous light by polycrystalline grains of the CuI film, whereas the signal of purely geometrical light reflection from the sample surface did not fall within the aperture of the recording system.

The dips in the curve of the red PL band (curves *1–3* in Fig. 4) for the wavelengths of the continuous illumination in the range 710–750 nm, which lie within this band, have an interesting nature. In this case, a process of stimulated luminescence takes place, and the greater part of the energy of the excited states of complexes of impurity and intrinsic defects is not radiated isotropically into a solid angle of 4π but propagates along the illumination direction, avoiding our recording system.

Hence, a phase-conjugation process for an optical wave, which is incident on the surface of a CuI film highly excited by a nitrogen laser, has been theoretically predicted (see Eq. (3)) and experimentally investigated. It is shown that phase conjugation occurs at energies of the incident photons equal to half the energy of exciton radiative recombination. Thus, it can be asserted in this case that phase conjugation occurs on the exciton states of the semiconductor film. If we take into account that laser radiation in the case of band-to-band excitation is absorbed in the surface layer of a submicron thickness (0.1 μm), it is easy to obtain a high exciton density in the semiconductor pumped by a pulsed nitrogen laser. In this case, it is precisely thin CuI films (which contain the hexagonal phase and exhibit high luminescent properties) that can be used for optical phase conjugation.

The mechanism of phase conjugation in the semiconductor medium that we observed substantially differs from the process of four-wave mixing on counterpropagating beams in an optically nonlinear medium, in which the phase conjugation of the signal wave can also occur, but the medium must be transparent for all light beams and possess a cubic optical nonlinearity. Since the value of the cubic optical nonlinearity is small, the optical phase-conjugation efficiency is low and a large volume of the nonlinear medium is necessary. In our case, the semiconductor is nontransparent for the pump wave and thin layers can be used. In order to completely restore the phase in the phase-conjugated wave, the phase-conjugating mirrors must be optically thin, i.e., on the order of the light wavelength. This can be accomplished only in our case.

The author is grateful to A.L. Despotuli for CuI films put at our disposal.

This work was supported by INTAS, grant no. 2002-0796.

REFERENCES

1. A. Kamilli, W. Send, D. Gerthen, *et al.*, Phys. Rev. B **62**, 13053 (2000).
2. M. Cardona, Phys. Rev. **152**, 262 (1963).
3. B. Honerlage, R. Levy, J. B. Grun, *et al.*, Phys. Rep. **124**, 161 (1985).
4. P. A. Belanger, A. Hardy, and A. E. Seigman, Appl. Opt. **19**, 602 (1980).
5. J. F. Lam and W. P. Brown, Opt. Lett. **5**, 61 (1980).
6. D. M. Bloom and G. C. Bjoklund, Appl. Phys. Lett. **31**, 592 (1977).
7. F. Kaczmarek, *Wstep do Fizyki Laserow* (PWN, Warszawa, 1978; Mir, Moscow, 1981).

Translated by A. Seferov

JETP Letters, Vol. 78, No. 3, 2003, pp. 110–114. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 141–145.

Double *K*-Shell Ionization of Atoms by a Single Photon¶

A. I. Mikhailov[1], I. A. Mikhailov[1], A. V. Nefiodov[1], G. Plunien[2], and G. Soff[2]

[1] *St. Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg, 188300 Russia*

[2] *Institut für Theoretische Physik, Technische Universität Dresden, D-01062 Dresden, Germany*

Received April 14, 2003; in final form, June 26, 2003

We investigate the double *K*-shell ionization of heliumlike ions by a single photon. A fast convergence of QED perturbation theory with respect to the parameter $1/Z$ is demonstrated in the entire nonrelativistic domain for moderate nuclear charge numbers $Z \geq 2$. The ratio of double-to-single photoionization cross sections is calculated for light heliumlike ions, taking into account the leading orders of $1/Z$ and αZ expansions. A comparison of our results with the available experimental data for a number of neutral atoms is presented. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 32.80.Fb; 12.20.Ds; 31.25.-v

In studies of electron correlations in atoms, the most attractive processes are those in which the electron–electron interaction plays the crucial role. One of such a fundamental phenomenon is the double photoionization of an atom caused by the absorption of a single photon, the so-called double photoeffect, which has been investigated for more than 30 years [1–3]. Since a photon interacts only with one electron, the simultaneous ejection of two electrons is exclusively caused by the electron–electron interaction. Accordingly, electron correlations show up here most clearly. So far, the experiments have been mainly performed with helium, the simplest many-electron atomic system. The majority of investigations concern the energy dependence of the ratio R of double-to-single photoionization cross sections [4–6]. With increasing photon energy ω, the ratio R grows rapidly beyond the ionization threshold. Then, after having a maximum near the threshold, it declines slowly, approaching the constant limit of 1.72(12)% [7] in the asymptotic domain of nonrelativistic photon energies much larger than the threshold energy I_{2K} for double ionization from the K-shell, that is, $I_{2K} \ll \omega \ll m$, where m is the electron mass ($\hbar = c = 1$). One of the most frequent problems in theoretical descriptions of double photoionization is the gauge dependence of numerical results.

Thanks to recent developments of novel synchrotron radiation sources, experiments with intense collimated beams of tunable monochromatic x-rays in the keV regime have become feasible [7–9]. This represents a challenge to theoretical investigations of double ionization of atomic inner-shell electrons in the entire nonrelativistic domain, both for photon energies $\omega \ll m$ and for targets with moderate values of nuclear charge numbers Z. In [10, 11], a Z-scaling law was suggested for the ratio R of double-to-single photoionization cross sections in the asymptotic energy regime, $I_{2K} \ll \omega \ll m$. For the energy domain near the threshold, $\omega \simeq I_{2K}$, *ab initio* calculations are presently not available. Nevertheless, one can mention here a model estimate of the two-electron photoejection cross section σ^{++} obtained in [12], which, however, strongly disagrees with the existing experimental data. Another numerical calculation of $Z^4\sigma^{++}$ has been performed within the framework of the convergent close-coupling model for He, Li^+, and O^{6+} [13].

In this letter, we have investigated the double photoionization of the ground state of heliumlike ions. The perturbation theory is developed with respect to the electron–electron interaction. As a zero approximation, Coulomb wave functions and Coulomb Green's functions are utilized. The study is performed for photon energies much smaller than the electron rest energy. Accordingly, all electrons involved in the ionization process are considered as being nonrelativistic. This implies the smallness of the Coulomb parameter, that is, $\alpha Z \ll 1$. However, the nuclear charge number Z is supposed to be high enough to utilize $1/Z$ as an expansion parameter. A similar approach has already been used in the asymptotic part of the nonrelativistic domain, $I_{2K} \ll \omega \ll m$, where all formulas can be significantly simplified [10]. In contrast to [10], we consider the entire nonrelativistic domain of incident photon energies with special emphasis on the threshold region. We have analyzed the limits of applicability of the approximations employed. Since K-shell electrons are essentially separated from the other electrons in an atom, it turns out that our formulas also describe fairly well the double K-shell ionization in the case of light neutral atoms.

In the leading order of nonrelativistic perturbation theory, double photoionization is described by the gauge-invariant set of Feynman diagrams shown in

¶This article was submitted by the authors in English.

Fig. 1. To ensure gauge invariance, the energy conservation law is defined to zeroth order as follows [14]:

$$E_{p_1} + E_{p_2} = \omega - I_{2K}. \quad (1)$$

Here, E_{p_1} and E_{p_2} are the one-electron energies in the continuum final state, $I_{2K} = 2I$ with $I = \eta^2/2m$ being the Coulomb potential for single ionization, and $\eta = m\alpha Z$ is the average momentum of a *K*-shell electron. Since the photon energy ω is distributed between two electrons, an escaping electron can have any energy within the range between 0 and $\omega - I_{2K}$.

The domain of photon energies near the threshold of double photoionization corresponds to the dipole regime characterized by $\omega \ll \eta$. Under the latter condition, the photon momentum $\boldsymbol{k}$ can be neglected, while the recoil momentum $\boldsymbol{q}$ transferred to the nucleus and the momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ of both outgoing electrons are of the order of a characteristic atomic momentum, $q \sim p_1 \sim p_2 \sim \eta$. Correspondingly, the process occurs at atomic distances of the order of the *K*-shell radius. This implies that the interaction with the Coulomb field of the nucleus has to be included in the initial, intermediate, and final electron states; that is, Coulomb wave and Coulomb Green's functions should be used already in the zeroth approximation [15–17]. In addition, the electron–electron interaction has to be taken into account in both initial and final states. All Feynman graphs depicted in Fig. 1 are expected to give comparable contributions to the total cross section for double photoionization.

Except for the parameter αZ, there are further Coulomb parameters involved in the problem, which characterize the interaction of intermediate and both outgoing electrons with the Coulomb field of the nucleus. They are given by $\xi = \eta/p$, where p being the momentum of the virtual electron, $\xi_1 = \eta/p_1$, and $\xi_2 = \eta/p_2$. Formally, these ξ parameters are values of the order of 1. Accordingly, dependences upon them in Coulomb Green's and wave functions have been taken into account exactly, but including terms of the order of αZ only. Terms of the order of $(\alpha Z)^2$ have been omitted in the present consideration. In the vicinity of the threshold, it can happen that the momenta of the ejected electrons, p_1 and/or p_2, become extremely small or, equivalently, ξ_1 and/or ξ_2 can reach rather large values compared to 1. This situation corresponds to the infrared catastrophe or to the quasiclassical limit, known also as the Wannier regime [18].

In the following, it is convenient to introduce dimensionless quantities, such as energies and momenta. We shall express all momenta in units of $\eta = m\alpha Z$. For example, the dimensionless momentum of a photon is just $k = k/\eta$. The energies are calibrated in units of $I = \eta^2/2m$. Then, the dimensionless energy of the incoming photon is given by $\varepsilon_\gamma = \omega/I$, while $\varepsilon_i = E_{p_i}/I$ ($i = 1, 2$) represent the dimensionless energies of the ejected

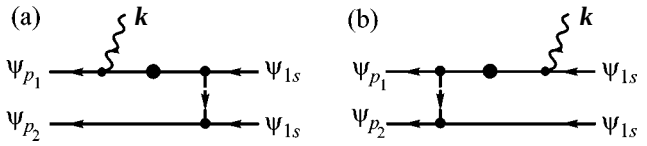


Fig. 1. Feynman diagrams for the double ionization of the atomic *K* shell by a single photon. Solid lines denote electrons in the Coulomb field of the nucleus, the dashed line denotes the electron–electron Coulomb interaction, and the wavy line denotes an incident photon. The line with a heavy dot corresponds to the Coulomb Green's function. For this line, only the energy is conserved, while the momentum is violated due to the interaction with a nucleus. Diagram (a) takes into account the electron–electron interaction in the initial state, while diagram (b) accounts for it in the final state. In addition to these graphs, one also has to take into account the corresponding exchange diagrams.

electrons. The energy conservation law is $\varepsilon_1 + \varepsilon_2 = \varepsilon_\gamma - 2$. The double-photoionization threshold corresponds to the photon energy $\varepsilon_\gamma^{\text{th}} = 2$. The amplitudes of the process being expressed in terms of dimensionless quantities become dependent on the nuclear charge number Z via the photon momentum $k = \alpha Z\varepsilon_\gamma/2$.

The total cross section of the double photoionization can be written in the following form:

$$\sigma^{++}(\varepsilon_\gamma; k) = \sigma_0 \frac{2^{10}\pi}{3Z^4} Q(\varepsilon_\gamma; k), \quad (2)$$

where $\sigma_0 = \alpha\pi a_0^2$ and $a_0 = 1/m\alpha$ is the Bohr radius. Due to a complicated dependence on the variables ε_γ and k, the function $Q(\varepsilon_\gamma; k)$ can be obtained only by numerical integration. The dipole regime corresponds to the approximation $k = 0$. The smallness of the photon momentum is due to the mutual interplay between two input quantities, ε_γ and Z. If one sets $k = 0$, the function $Q(\varepsilon_\gamma)$ becomes independent of the value of Z. The function holds the same for the entire helium isoelectronic sequence. Accordingly, the ratio of double-to-single ionization cross sections, which is usually measured experimentally, is given by

$$R(\varepsilon_\gamma) = \frac{\sigma^{++}(\varepsilon_\gamma)}{\sigma^{+}(\varepsilon_\gamma)} = \frac{1}{Z^2}\frac{Q(\varepsilon_\gamma)}{H(\varepsilon_\gamma)}. \quad (3)$$

The energy factor $H(\varepsilon_\gamma)$, which enters in the expression for the cross section of the single *K*-shell photoionization in the leading order of perturbation theory, reads as follows [19]:

$$H(\varepsilon_\gamma) = \frac{1}{\varepsilon_\gamma^4}\frac{\exp(-4\tau\cot^{-1}\tau)}{[1 - \exp(-2\pi\tau)]}, \quad (4)$$

where $\tau = 1/\sqrt{\varepsilon_\gamma - 1}$.

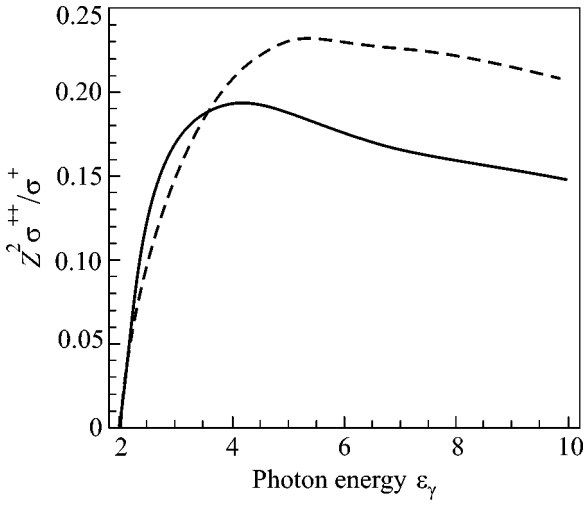


Fig. 2. Different contributions to the universal ratio $Z^2R(\varepsilon_\gamma)$ of double-to-single photoionization cross sections calculated in the Coulomb gauge for k = 0 according to Eq. (3). Dashed line: contribution due to the electron–electron interaction in the initial state only; solid line: total contribution of all diagrams.

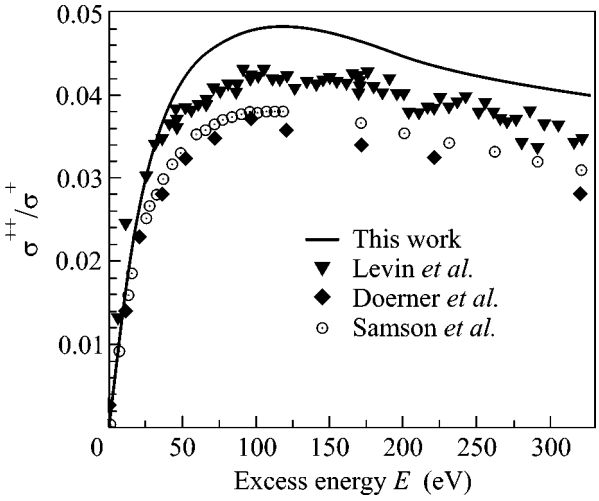


Fig. 3. The ratio of double-to-single photoionization cross sections of the ground state of helium is given as a function of the excess energy $E = E_{p_1} + E_{p_2}$. Experimental data taken from Levin *et al.* [4], Dörner *et al.* [5], and Samson *et al.* [6] account also for the contribution from the cross section $\sigma^{+}*$ of single photoionization with excitation, while the theoretical calculation is performed according to Eq. (3).

All our numerical calculations were performed using the Coulomb gauge, since it easily allows one to separate the leading contributions in the nonrelativistic limit. Corrections due to spin-dependent terms turn out to be suppressed by a factor of the order of $(\alpha Z)^2$ in transition amplitudes of the process and, therefore, have been neglected. We compared the universal quantity $Z^4\sigma^{++}$ calculated in the dipole approximation according to Eq. (2) with that obtained in [13]. The curves lie quite close to each other, although they are not identical. The contributions to σ^{++} due to two-photon exchange graphs are relatively small. In Fig. 2, the universal ratio $Z^2R(\varepsilon_\gamma)$ is shown, which is valid for moderate Z values. Accounting for the electron–electron interaction to the final state results in significant corrections to the total cross section of the double photoionization. Note, however, that the individual contributions of each diagram are gauge-dependent.

In Fig. 3, we compare our numerical results for the ratio σ^{++}/σ^+ with the most recent measurements for helium [4–6]. The comparison of theoretical and experimental data is given in dependence on the excess energy, since the experimental threshold energy I_{2K}^{exp} = 79.0 eV and the theoretical one I_{2K} = 108.85 eV are different. However, this is an apparent problem. In fact, the expansion with respect to $1/Z$ converges by an order of magnitude in the entire nonrelativistic domain, starting from helium. The contribution to the ionization potential due to the electron–electron interaction treated by the exact one-photon exchange leads to the correction $\Delta I_{2K}^{(1)} = -34.02$ eV; that is, the threshold energy becomes equal to 74.83 eV. The correction due to two-photon exchange diagrams, which is known to be $\Delta I_{2K}^{(2)} = 4.29$ eV [20], yields a threshold energy of about 79.12 eV. The contribution, which results from the three-photon exchange, has been calculated in the ladder approximation [21]. It yields $\Delta I_{2K}^{(3)} = -0.12$ eV [22] and brings the theoretical threshold energy into agreement with the experimental value. Independent measurements of the ratio of double-to-single photoionization cross sections performed by Dörner *et al.* [5] and by Samson *et al.* [6] seem to be more reliable than the experimental data by Levin *et al.* [4] (see discussion in [6]). The disagreement between our results and those obtained in [5, 6] is due to next-to-leading order terms of perturbation theory in $1/Z$, which have been omitted in the present evaluation of single photoionization cross section.

The fast convergence of the expansion over $1/Z$ in the entire nonrelativistic domain, including the threshold area, can be understood by a careful distinction between formal and real characteristic scales. The characteristic spatial distances involved in the problem turn out to be somewhat larger than the K-shell radius. This is confirmed, for example, by calculations of the double-electron ionization in Compton scattering, where perturbation theory with respect to $1/Z$ is proven to be satisfactory for helium in the asymptotic domain of nonrelativistic photon energies [23, 24]. The distinction between formal and real scales also explains the fact that, for heliumlike ions with $Z > 2$, the double photoionization from the excited 2^1S state turns out to be

For various neutral atoms, the nuclear charge numbers Z, the experimental energies ω of an incident photon, the experimental potentials I^{exp} for single K-shell ionization, dimensionless photon energies ε_γ, effective values Z_{eff} for the nuclear charge, and the theoretical and experimental ratios $R(\varepsilon_\gamma)$ of double-to-single K-shell ionization cross sections are tabulated. The photon energies ω are calibrated in units of the experimental ionization potentials I^{exp} taken from [26]. The theoretical ratios $R(\varepsilon_\gamma)$ are calculated using the effective values Z_{eff} according to Eq. (3)

Neutral atom	Z	ω (keV)	I^{exp} (keV)	ε_γ	Z_{eff}	$R(\varepsilon_\gamma)$		
						this work	experiment	
Ne	10	3	0.87	3.45	8.0	0.29×10^{-2}	0.3×10^{-2}	[27]
Ti	22	17.4	4.97	3.50	19.11	0.51×10^{-3}	0.53×10^{-3}	[28]
Cr	24	17.4	5.99	2.90	20.98	0.37×10^{-3}	0.38×10^{-3}	[28]
Fe	26	17.4	7.12	2.44	22.88	0.23×10^{-3}	0.24×10^{-3}	[28]
Ni	28	17.4	8.34	2.09	24.76	0.51×10^{-4}	1.1×10^{-4}	[28]
Cu	29	20	8.99	2.22	25.70	1.1×10^{-4}	$1.3(3) \times 10^{-4}$	[9]
Mo	42	50	20.01	2.50	38.35	0.87×10^{-4}	$3.4(6) \times 10^{-4}$	[8]

more probable than from the K shell [11, 25]. It should be noted that, in the next-to-leading order of perturbation theory in $1/Z$, there are vertex contributions due to the crossed diagrams, which entangle correlation effects in the initial and final electron states.

Considering double K-shell photoionization in neutral atoms, the wave functions and Green's functions possess an essentially non-Coulomb behavior. Accordingly, numerical calculations require the use of the Hartree–Fock method already in the zeroth approximation. Formally, the screening effect can be simulated by replacing the true nuclear charge number Z in Eq. (3) with an effective value Z_{eff}. The latter can be defined by equating the experimental potential I^{exp} for single K-shell ionization and the effective one, that is, $I^{exp} = m(\alpha Z_{eff})^2/2$. Apart from the nickel and molybdenum systems, good overall agreement between our predictions for neutral atoms with available experimental data is achieved (see the table). The significant disagreement for the nickel atom seems to be due to high uncertainties of the results, both theoretical and experimental. The ratio $R(\varepsilon_\gamma)$ here is extremely sensitive to the photon energy, because the latter is very close to the threshold energy. In the case of molybdenum, a possible explanation for the deviation may be connected with relativistic effects, for example, with the spin–orbit interaction, which were neglected in the present consideration. Another reason might be higher error bars rather than those quoted in [8].

In conclusion, we demonstrate a fast convergence of QED perturbation theory with respect to the parameter $1/Z$ in the entire nonrelativistic domain for $Z \geq 2$. Double K-shell photoionization has been investigated for heliumlike ions and neutral atoms with moderate Z values, taking into account the leading orders of $1/Z$ and αZ expansions. Going beyond the leading-order consideration requires rigorous QED description.

We thank E.P. Kanter and R.W. Dunford for providing experimental data to us prior to publication. A.M. is grateful to the Dresden University of Technology for the hospitality and for financial support from DFG. G.S. and G.P. acknowledge financial support from BMBF, DFG, and GSI. A.N. is supported by the Alexander von Humboldt Foundation. A.M. and I.M. acknowledge support from the Russian Foundation for Basic Research (grant nos. 01-02-17246 and 00-15-96610).

REFERENCES

1. M. Y. Amusia, in *Atomic Photoeffect*, Ed. by K. T. Taylor (Plenum, New York, 1990).
2. J. H. McGuire, N. Berrah, R. J. Bartlett, *et al.*, J. Phys. B **28**, 913 (1995).
3. J. S. Briggs and V. Schmidt, J. Phys. B **33**, R1 (2000).
4. J. C. Levin, G. B. Armen, and I. A. Sellin, Phys. Rev. Lett. **76**, 1220 (1996).
5. R. Dörner, T. Vogt, V. Mergel, *et al.*, Phys. Rev. Lett. **76**, 2654 (1996).
6. J. A. R. Samson, W. C. Stolte, Z.-X. He, *et al.*, Phys. Rev. A **57**, 1906 (1998).
7. L. Spielberger, O. Jagutzki, R. Dörner, *et al.*, Phys. Rev. Lett. **74**, 4615 (1995).
8. E. P. Kanter, R. W. Dunford, B. Krässig, *et al.*, Phys. Rev. Lett. **83**, 508 (1999).
9. R. Diamant, S. Huotari, K. Hämäläinen, *et al.*, Phys. Rev. A **62**, 052519 (2000).
10. M. Y. Amusia, E. G. Drukarev, V. G. Gorshkov, *et al.*, J. Phys. B **8**, 1248 (1975).
11. R. C. Forrey, H. R. Sadeghpour, J. D. Baker, *et al.*, Phys. Rev. A **51**, 2112 (1995).
12. M. A. Kornberg and J. E. Miraglia, Phys. Rev. A **49**, 5120 (1994).
13. A. S. Kheifets and I. Bray, Phys. Rev. A **58**, 4501 (1998).

14. A. Y. Istomin, N. L. Manakov, and A. F. Starace, J. Phys. B **35**, L543 (2002).
15. V. G. Gorshkov, A. I. Mikhailov, and V. S. Polikanov, Nucl. Phys. **55**, 273 (1964).
16. V. G. Gorshkov, Zh. Éksp. Teor. Fiz. **47**, 352 (1964) [Sov. Phys. JETP **20**, 234 (1965)].
17. V. G. Gorshkov and V. S. Polikanov, Pis'ma Zh. Éksp. Teor. Fiz. **9**, 464 (1969) [JETP Lett. **9**, 279 (1969)].
18. G. H. Wannier, Phys. Rev. **90**, 817 (1953).
19. A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, 3rd ed. (Nauka, Moscow, 1969; Wiley, New York, 1974).
20. I. Lindgren, H. Persson, S. Salomonson, *et al.*, Phys. Rev. A **51**, 1167 (1995).
21. O. Y. Andreev, L. N. Labzowsky, G. Plunien, *et al.*, Phys. Rev. A **67**, 012503 (2003).
22. O. Y. Andreev, private communication.
23. M. Y. Amusia and A. I. Mikhailov, J. Phys. B **28**, 1723 (1995).
24. T. Surić, K. Pisk, and R. H. Pratt, Phys. Lett. A **211**, 289 (1996).
25. M. Y. Amusia, A. I. Mikhailov, and I. A. Mikhailov, J. Phys. B **32**, 4885 (1999).
26. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer, Berlin, 1985).
27. E. P. Kanter, private communication.
28. J. Ahopelto, E. Rantavuori, and O. Keski-Rahkonen, Phys. Scr. **20**, 71 (1979).

JETP Letters, Vol. 78, No. 3, 2003, pp. 115–118. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 146–149.

X-ray Study of Microdroplet Plasma Formation under the Action of Superintense Laser Radiation

Y. Fukuda[1], K. Yamakawa[1], Y. Akahane[1], M. Aoyama[1], N. Inoue[1], H. Ueda[1], J. Abdallah, Jr.[2], G. Csanak[2], A. Ya. Faenov[3,*], A. I. Magunov[3,*], T. A. Pikuz[3], I. Yu. Skobelev[3], A. S. Boldarev[4], and V. A. Gasilov[4]

[1]*Advanced Photon Research Center, Japan Atomic Energy Research Institute, 619-0215 Kyoto, Japan*

[2]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

[3]*Center for Data on Multicharged Ion Spectra, All-Russia Research Institute of Physicotechnical and Radioelectronic Measurements, Mendeleevo, Moscow region, 141570 Russia*

[4]*Institute of Mathematical Modeling, Russian Academy of Sciences, Moscow, 125047 Russia*

**e-mail: misdc@vniiftri.ru*

Received June 10, 2003

We have measured the X-ray emission spectra of a plasma generated by laser radiation with an intensity above 10^{19} W/cm^2 and a pulse duration of 30 fs acting upon an argon jet target with a large abundance of micron-sized clusters. The time variation of the X-ray yield from ions of various multiplicities, calculated within the framework of a nonstationary kinetic model, shows a good qualitative agreement with the experimental time-integrated spectrum. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 52.38.Ph; 52.25.Os; 52.25.Dg; 52.50.Jm

1. In recent years, there has been rapid progress in the spectroscopy of X-ray emission from microplasma sources formed as a result of the interaction of ultrashort multiterawatt laser pulses with cluster targets. One of the goals consists in elucidation of the mechanisms of interaction between radiation and matter under such extreme conditions, which would provide for the creation and development of high-intensity coherent and incoherent X-ray sources for various applications. Experiments using high-intensity laser radiation (10^{17}–10^{18} W/cm^2) were performed mostly with small clusters 10–100 nm in diameter [1–7]. The spectra of X-rays generated by plasma formed in the case of large (~1 μm) clusters were studied previously using femtosecond laser pulses of lower intensity [8]. Hard X-ray and hot electron production by microdroplets of water under the action of femtosecond laser pulses was studied in [9]. The results of these investigations showed that irradiation of a gas jet target enriched with clusters of a medium size (100–200 nm) leads to soft X-ray emission with a considerable yield.

The aim of this work was to study the high-resolution X-ray emission spectra of plasma formed from large argon clusters (containing over 10^{10} atoms) under the action of superintense laser radiation.

2. The experiments with argon clusters were performed at JAERI (Kyoto, Japan) using a 100-TW Ti:sapphire laser system based on the chirped pulse amplification technique. The system generates 20-fs pulses at a repetition rate of 10 Hz and is capable of collimating output radiation with intensities after focusing up to 10^{20} W/cm^2 [10, 11]. The initial pulses (λ_{las} = 800 nm; τ_{las} = 10 fs; repetition rate, 82.7 MHz) were generated by a Ti:sapphire oscillator. Stretched to 10 ns and frequency modulated, the pulses were amplified by a regenerative amplifier and two multipass amplifiers. The amplified pulses were compressed to 20 fs by a vacuum compressor providing a maximum output energy of 1.9 J. The compressed pulses were directed to a vacuum target chamber by two gold-coated plane mirrors and focused by an *f*/3 gold-coated off-axis parabolic mirror. The focal spot size was 11 μm at $1/e^2$, which was only 10% greater than the diffraction limit. A Gaussian spot accounts for 64% of the total laser output energy. For a laser pulse duration of 30 fs and an energy of 300 mJ used in our experiments, the peak intensity amounted to 1.2×10^{19} W/cm^2. A prepulse arriving 1 ns before the main pulse was suppressed by passing the output signal through two double Pockels cells. The contrast ratio of the main pulse intensity to that of the prepulse was greater than 10^5.

3. Argon clusters were produced by expanding a high-pressure (up to p = 60 bar) gas jet into vacuum via a specially designed pulsed supersonic conical nozzle with the input and output diameters of 0.5 and 2.0 mm, respectively, and a length of 75 mm. The nozzle shape and dimensions were optimized by numerical simulation of a two-phase jet with a maximum yield of clusters of required size (d_{cl} ~ 1 μm) [12, 13].

0021-3640/03/7803-0115$24.00

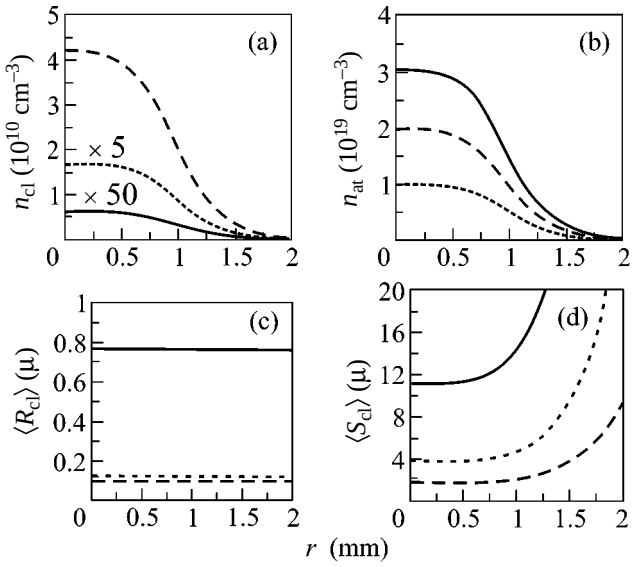


Fig. 1. Calculated radial profiles of the parameters of a two-phase argon jet target at a distance of 1.5 mm from the nozzle output: (a) total cluster concentration n_{cl}; (b) concentration of atomic species n_{at}; (c) average cluster radius $R_{cl}\rangle$; (d) average distance between clusters $\langle S_{cl}\rangle$. The input gas pressure p = 60 bar (solid curves), 40 bar (dashed curves), and 20 bar (dotted curves).

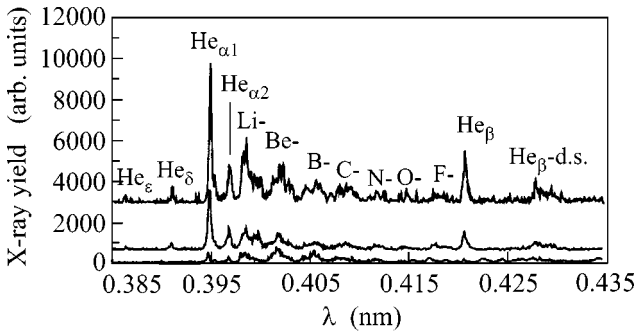


Fig. 2. The experimentally measured time-integrated X-ray emission spectra of a microdroplet plasma generated by 30-fs laser pulses of various intensities and contrasts: I = 1.2×10^{19} W/cm^2, C = 2.2×10^5 (top curve); I = 3×10^{18} W/cm^2, C = 2×10^3 (middle curve); I = 3×10^{18} W/cm^2, C = 50 (bottom curve).

Figure 1 presents some results of two-dimensional hydrodynamic calculations of a two-phase argon jet for various input pressures. At a maximum pressure of 60 bar, the average cluster size exhibits a sharp increases (reaching about 1.5 μm) and is virtually constant in the jet cross section. The total concentrations of clusters and atomic species also remain almost constant in the axial region with a radius of 1.5 mm. This homogeneity is important for a cluster target to be used for obtaining an active radiating plasma medium.

4. In the experiment, the laser beam was focused on a jet at a distance of 1.5 mm from the nozzle output. Spatially resolved X-ray spectra of the laser-generated plasma were measured using a focusing spectrometer with spatial resolution (FSSR-1D) [14–17]. The spectrometer is equipped with a spherically bent mica crystal (R = 150 mm) and a vacuum-compatible X-ray CCD camera (DF420-BN, ANDOR). The crystal was placed at a distance of 381.2 mm from the plasma source and oriented at the Bragg angle (θ = 54.3°) relative to the center of a spherical surface (the fourth-order reflection for a wavelength of λ = 0.405 nm). The reflection plane of the spectrometer was oriented in the laser beam propagation direction in order to provide for the one-dimensional spatial resolution in the transverse direction. The size of the active plasma zone (emitting in the spectral region of the He-like argon ion and the corresponding dielectronic satellites) in this direction was below 200 μm.

Figure 2 shows the typical densitometer traces of the plasma X-ray emission spectra measured at an input gas pressure of p = 60 bar and various values of the intensity and contrast of a 30-fs laser pulse. The upper spectrum corresponds to the maximum intensity of the laser pulse (1.2×10^{19} W/cm^2). In order to reduce the prepulse effect on the clusters, the contrast of the main pulse was set at a level of C = 2.2×10^5. In the fourth order of reflection, the spectrum exhibits the resonant ($He_{α1}$) and intercombination ($He_{α2}$) lines of the $1s2p$–$1s^2$ transition in the He-like argon ion. In addition, there is a clearly resolved structure of dielectric satellites in the Li-like ion and the lines of radiative transitions from autoionization states with the configurations $1s2s^m2p^n$ in ions of lower charge (from Be- to F-like). The fifth order of reflection displays the lines of $1snp$–$1s^2$ (n = 3–6) transitions in the He-like ion and the dielectric satellites of the $1s3p$–$1s^2$ ($He_β$) line. The n = 4 line ($He_γ$) falls between satellite lines of the Li-like ion. The middle curve in Fig. 2 shows the spectrum measured for the main peak intensity of 3×10^{18} W/cm^2 and a contrast of C = 2×10^3. Here, the X-ray yield decreases almost in proportion to the laser pulse intensity despite a significantly reduced contrast. Further decrease in the contrast (down to C = 50) leads to a dramatic change in the spectrum (see the bottom curve in Fig. 2), whereby the X-ray yield in the lines of the He-like ion drops sharply relative to the satellite emission intensity. This can be explained by the dominating role of the prepulse, the intensity of which at this contrast reaches about 10^{17} W/cm^2. This power density is sufficient to ionize even large clusters, which decay before the arrival of the main pulse.

It should be noted that neither the lines of the He-like ions nor their closest satellites were observed at an input gas pressure of 40 bar. According to the results of calculations presented in Fig. 1, the average size of clusters formed under these conditions is almost ten times smaller than that for p = 60 bar. Thus, at an input gas pressure of p = 40 bar, the prepulse destroys clusters at all values of the contrast used in these experiments.

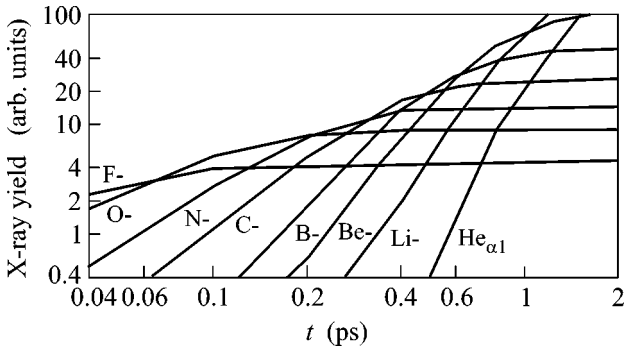


Fig. 3. Time variation of the X-ray yield calculated for differently charged argon ions.

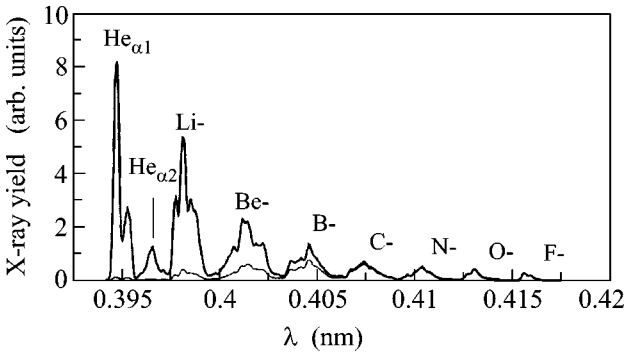


Fig. 4. Time-integrated X-ray emission spectra calculated for the time moments of 0.6 ps (thin line) and 1.4 ps (thick line) after the laser pulse end.

5. The experimental X-ray emission spectra are time-integrated and spatially integrated in the Bragg reflection plane. In order to elucidate the plasma radiation dynamics in more detail, we have performed model calculations of the line emission spectrum with allowance for the time relaxation of the nonequilibrium electron energy distribution established in the plasma immediately after the main femtosecond laser pulse. For the sake of simplicity, the consideration was restricted to a homogeneous nonstationary kinetic model of a microdroplet plasma with frozen ions. The model is described by the equation

$$\frac{\partial f(E,t)}{\partial t} = \mathrm{St} f, \qquad (1)$$

where the collision integral on the right-hand side accounts for the elastic collisions between electrons and for the inelastic electron–ion collisions. The assumption of spatial homogeneity is justified by a comparison of the radiative lifetimes of the autoionization states observed for argon ions (10^{-1}–10^{-2} ps) to the characteristic time of cluster destruction (size doubling) estimated by the formula

$$\tau_{\mathrm{dec}} = 1.2 \times 10^{-12} D_0 \sqrt{A/T_i}, \qquad (2)$$

where A is the atomic number, T_i is the effective ion temperature (expressed in kiloelectronvolts), and D_0 is the initial cluster diameter (expressed in microns). The ion temperature can be estimated from the observed Doppler width of the resonance line of the He-like argon ion, which yields T_i = 6.7 keV. Then, for 1-μm argon clusters (A = 18), formula (2) gives $\tau_{\mathrm{dec}} \approx 2$ ps, which is much greater than the radiative lifetime of the autoionization states in the satellite lines.

Equation (1) was numerically solved in the interval of electron energies from 0 to 10 keV with a uniform step of 100 eV. At the initial moment corresponding to the laser pulse end, all argon atoms were considered as ionized up to the state of Ne-like ions. The threshold laser intensity for the tunneling ionization of the Na-like argon atom ($I_{\mathrm{th}} = 3.6 \times 10^{16}$ W/cm^2 [18]) can be achieved even with allowance of the field weakening inside a microdroplet. The ion density was taken equal to 10^{21} cm^{-3} and the residual electron energy was distributed with a narrow interval at 5 keV. At each time step (in interval from 0 to 2 ps), the populations of more than 3000 levels of ions (from Ne- to He-like) were calculated by solving a set of balance equations. The collision rates were calculated using the corresponding electron energy distribution function. The populations of levels were used to calculate the spectral dependence of the X-ray yield in the energy range studied.

The time dependence of the total radiation yield from different ions obtained in these calculations are shown in Fig. 3. It is seen that the emission from lower charged ions reaches the saturation level earlier. This is explained by a rapid shift of the ion composition towards higher charge numbers as a result of ionization by hot electrons. The radiation yield of the He-like ion is still increasing after $t > 1$ ps, because electrons are thermalized at the mean energy below 1 keV and the ion composition goes to the stationary state. Of course, this yield growth will be limited by the decay of clusters.

Figure 4 presents the time-integrated spectra calculated for the time moments 0.6 and 1.4 ps after the end of the laser pulse. As can be seen, the X-ray emission spectrum of the ions from F- to C-like is completely formed by 0.6 ps, whereas the emission from He- and Li-like ions is accumulated during the subsequent period of time. This feature of the radiation dynamics can be verified by direct time-resolved spectral measurements.

6. To summarize, we have studied for the first time the X-ray emission spectra of a microdroplet argon plasma formed under the action of 30-fs laser pulses with a peak intensity above 10^{19} W/cm^2. For this purpose, a special nozzle was designed and manufactured based on the results of numerical modeling of the process of cluster formation in a gas jet target. The nozzle allowed large argon clusters to be obtained without additional cooling of the valve. A comparatively large

decay time of such clusters makes it possible to decrease the negative effect of a prepulse, which (under conditions of ultrahigh laser intensity) produces destruction of small clusters even at the main pulse contrast above 10^5. The calculated pattern of relaxation of the electron energy distribution function and the corresponding ion level populations after the laser action qualitatively agree with the experimentally observed time-integrated X-ray emission spectra of a microdroplet plasma. Predicted temporal features of the X-ray yield can be verified by direct time-resolved spectral measurements.

This study was supported by the Russian Foundation for Basic Research (project no. 02-01-00708), INTAS (grant no. 01-0233), and the U.S. Department of Energy (J.A. and G.C., contract no. W-7405-ENG-36). The research described in this publication was made possible in part by the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF Award no. RP1-2328-ME-02).

REFERENCES

1. A. McPherson, T. S. Luk, B. D. Thompson, *et al.*, Phys. Rev. Lett. **72**, 1810 (1994).
2. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).
3. T. Ditmire, J. Zwelback, V. P. Yanovsky, *et al.*, Nature **398**, 489 (1999).
4. V. Kumarappan, M. Krishnamurthy, D. Mathur, and L. C. Tribedi, Phys. Rev. A **63**, 023203 (2001).
5. G. C. Junkel-Vives, J. Abdallah, Jr., F. Blasco, *et al.*, Phys. Rev. A **64**, 021201R (2001).
6. J. Abdallah, Jr., A. Ya. Faenov, I. Yu. Skobelev, *et al.*, Phys. Rev. A **63**, 032706 (2001).
7. T. Shiraishi, M. Mori, and K. Kondo, Phys. Rev. A **65**, 045201 (2002).
8. E. Parra, T. Alexeev, J. Fan, *et al.*, Phys. Rev. E **62**, R5931 (2000).
9. T. D. Donnelly, M. Rust, I. Weiner, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **34**, L313 (2001).
10. K. Yamakawa, M. Aoyama, S. Matsuoka, *et al.*, Opt. Lett. **23**, 1468 (1998).
11. K. Yamakawa and C. P. J. Barty, IEEE J. Sel. Top. Quantum Electron. **6**, 658 (2000).
12. A. S. Boldarev, V. A. Gasilov, F. Blasco, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 583 (2001) [JETP Lett. **73**, 514 (2001)].
13. G. C. Junkel-Vives, J. Abdallah, Jr., T. Auguste, *et al.*, Phys. Rev. E **65**, 036410 (2002).
14. A. Ya. Faenov, S. A. Pikuz, A. I. Erko, *et al.*, Phys. Scr. **50**, 333 (1994).
15. T. A. Pikuz, A. Ya. Faenov, S. A. Pikuz, *et al.*, J. X-Ray Sci. Technol. **5**, 323 (1995).
16. I. Yu. Skobelev, A. Ya. Faenov, B. A. Bryunetkin, *et al.*, Zh. Éksp. Teor. Fiz. **108**, 1263 (1995) [JETP **81**, 692 (1995)].
17. B. K. Young, A. L. Osterheld, D. F. Price, *et al.*, Rev. Sci. Instrum. **69**, 4049 (1998).
18. M. V. Ammosov, N. B. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1986)].

Translated by P. Pozdeev

JETP Letters, Vol. 78, No. 3, 2003, pp. 119–122. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 150–153.

Single-Mode Nonlinear Regime of Weibel Instability in a Plasma with Anisotropic Temperature¶

V. Yu. Bychenkov*, **, W. Rozmus, and C. E. Capjack*****

** Lebedev Physics Institute, Russian Academy of Sciences, Moscow, 119991 Russia*

*** Department of Physics, University of Alberta, Edmonton T6G 2J1, Alberta, Canada*

**** Department of Electrical and Computer Engineering, University of Alberta, Edmonton T6G 2J1, Alberta, Canada*

Received June 16, 2003

An analytical solution is found to the vortex electron anisotropic hydrodynamic equations that describe the nonlinear evolution of the long-wavelength Weibel instability. The presented analytical approach shows that the long-wavelength Weibel instability saturates without a decrease in the temperature anisotropy in the single-mode regime due to the rotation of the anisotropy axes. The generated magnetic field is circular-polarized, and its amplitude varies periodically in time. © *2003 MAIK "Nauka/Interperiodica"*.

PACS numbers: 52.35.Hr

It is well known that a plasma with an electron energy anisotropy is unstable [1]. The Weibel instability is a pure electromagnetic mode that describes magnetic field excitation with a characteristic time scale much longer than the electron plasma wave period. Such an instability likely appears in the interaction of a short-duration X-ray pulse with a gas media where weak collisions are unable to rapidly remove the electron energy anisotropy arising from photoionization. In the past, the only known artificial source of short intense X-rays was a nuclear explosion. However, recent developments in intense ultrashort X-ray laser technology have opened a new regime of X-ray–matter interaction, in which intense X-ray laser pulses penetrate deep into the targets and produce nonequilibrium plasma of macroscopic size. An example is the XFEL project [2], which is generating interest in the need to understand extreme plasma states that can be encountered in the interaction of X-rays with matter.

The most advanced studies dealing with the nonlinear stage of a plasma with an anisotropic temperature have been performed with numerical simulations [3–8]. It is also known that the vortex electron anisotropic hydrodynamics (VEAH) model [9] is well suited for analytical study of this problem. However, the analytical results obtained to date mainly demonstrate the existence of explosionlike or self-similar solutions to the VEAH equations [9, 10], while the conditions of their approach with an initial value problem remain unclear. Recent three-dimensional (3D) particle-in-cell simulations (PIC) have shown that the Weibel instability for high temperature anisotropy ($T_\perp/T_\parallel \gg 1$) evolves into a single-mode 1D regime with a finite saturated averaged anisotropy ($T_\perp/T_\parallel \sim 1$) and a circularly polarized long-scale-length magnetic field [8]. The initial conditions for these simulations corresponded to a small-amplitude stochastic initial magnetic field, which is different from the 1D VEAH numerical model for a one-component magnetic field [7], where the single-mode regime was found to exist for a short-wavelength Weibel instability growing from the initial harmonic perturbations. The single-mode regime helical polarization of the magnetic field, which saturates due to periodic energy exchange with thermal electrons, was predicted in [11].

For a plasma with anisotropic temperature, we examine the nonlinear evolution of a long-wavelength Weibel instability, $ck < \omega_p$, where k is the wave number, ω_p is the electron plasma frequency, and c is the speed of light. We have found an exact single-mode solution of the VEAH equations that can be applied to the initial value problem for the Weibel instability itself and to provide analytic confirmation of the existence of the 1D saturation regime displayed in recent 3D PIC simulations [8].

The VEAH model [9] is formulated in terms of the coupled equations for the temperature tensor $\mathbf{T}$ and the quasistatic magnetic field $\mathbf{B}$. In the long-wavelength limit, these equations are

$$\frac{\partial \mathbf{\Omega}}{\partial t} = -\frac{1}{m_e}\nabla \times \nabla \cdot \mathbf{T}, \quad \frac{\partial \mathbf{T}}{\partial t} = \{\mathbf{T} \times \mathbf{\Omega}\}, \qquad (1)$$

where $\mathbf{\Omega} = eB/m_e c$ is the electron hyrofrequency, e and m_e are the electron charge and mass, and $\{\ldots\}$ signifies a symmetrization of the tensor, $\{A_{ij}\} = A_{ij} + A_{ji}$. Equation (1) conserves the temperature anisotropy. For a sin-

¶This article was submitted by the authors in English.

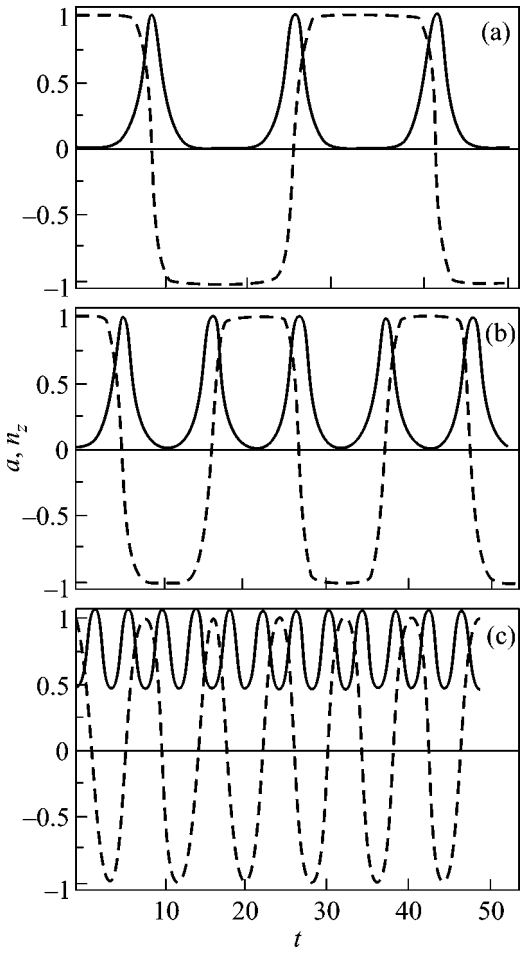


Fig. 1. Evolution of the magnetic field (solid lines) and the anisotropy axis (dashed lines) for $\Phi_0 = 0$ and a_0 = (a) 0.001, (b) 0.02, and (c) 0.5.

gle-axis temperature anisotropy,

$$\mathbf{T} = T_\parallel \mathbf{nn} + T_\perp(\mathbf{I} - \mathbf{nn}), \tag{2}$$

the longitudinal and transversal temperatures, $T_\parallel$ and $T_\perp$, do not change in time. Here, $\mathbf{I}$ is the absolute unit tensor and $\mathbf{n}(\mathbf{r}, t)$ is the unit vector in the direction of the anisotropy axis. We assume homogeneous $T_\parallel$ and $T_\perp$ in space, which corresponds to the standard Weibel instability [1].

The hypothesis that the nonlinear regime of the long-wavelength Weibel instability can be described in terms of $\mathbf{\Omega}$ and $\mathbf{n}$ was put forward in [12]. However, the corresponding equations,

$$\frac{\partial \mathbf{\Omega}}{\partial t} = \frac{T_\parallel - T_\perp}{m_e}\nabla \times (\mathbf{n} \times \nabla \times \mathbf{n} - \mathbf{n}\nabla \cdot \mathbf{n}),$$
$$\frac{\partial \mathbf{n}}{\partial \mathbf{t}} = \{\mathbf{n} \times \mathbf{\Omega}\}, \quad |\mathbf{n}| = 1, \tag{3}$$

have not been solved for the initial value problem. In order to solve the initial value problem for the one-dimensional case ($\partial/\partial z$), we take advantage of the following parameterization:

$$\mathbf{\Omega} = \{\gamma_0 a(t)\sin kz, -\gamma_0 a(t)\cos kz, 0\},$$
$$\mathbf{n} = \{\sin\Phi(\mathbf{t})\cos\mathbf{kz}, \sin\Phi(\mathbf{t})\sin\mathbf{kz}, \cos\Phi(\mathbf{t})\}. \tag{4}$$

This corresponds to a circularly polarized single-mode magnetic field. Here, for convenience, we have introduced the growth rate of the long-wavelength Weibel instability, $\gamma_0 = k\sqrt{(T_\perp - T_\parallel)/m_e}$.

One may confirm through the direct substitution of Eqs. (4) into Eqs. (3) that the assumed structure satisfies the VEAH model. The corresponding equations for the dimensionless amplitude of the magnetic field a and the angle Φ, which defines the evolution of the anisotropy axis, are

$$2\ddot{\Phi} = \sin 2\Phi, \quad a = \dot{\Phi}. \tag{5}$$

The solutions to these equations are given in quadratures as follows:

$$\gamma_0 t = \int_{\Phi_0}^{\Phi} d\varphi[a_0^2 + \sin(\varphi + \Phi_0)\sin(\varphi - \Phi_0)]^{-1/2},$$
$$a^2 = a_0^2 + \sin(\Phi + \Phi_0)\sin(\Phi - \Phi_0), \tag{6}$$

where the two constants of integration $a(0) = a_0$ and $\Phi(0) = \Phi_0$ can be related to the initial amplitude of the magnetic field and the initial direction of the anisotropy axis. This nonlinear solution may be expressed in terms of an elliptic integral of the first kind.

For the standard definition of the problem for Weibel instability, one assumes $\Phi_0 = 0$ and $a_0 \ll 1$. Initially, the amplitude of the magnetic field increases exponentially with the growth rate γ_0. After that, the growth of the magnetic field slows, and the magnetic field reaches a maximum and then decreases to the initial value. This process is then repeated in a periodic fashion. During one magnetic field cycle, the anisotropy direction changes to the opposite direction but returns to the starting direction during the next cycle. The period of the magnetic field pulsations slowly decreases with a_0. Figure 1 shows the time dependences of the dimensionless amplitude of the magnetic field and the z component of the anisotropy vector, n_z, for $n_z(0) = 1$, where the time scale is given in γ_0^{-1} units. The solutions for $a(t)$ and n_z are the multi-instanton solution and multikink solution, respectively.

Considerable initial amplitudes a_0 (e.g., Fig. 1c) are also of interest for the evaluation of the nonlinear behavior of the Weibel instability, which is initially excited in the short-wave-length domain and then evolves to long scale lengths with the temperature anisotropy saturation and circular-polarized magnetic field formation. Such behavior of the Weibel instability evolving to the 1D case was recently observed in 3D

PIC simulations [8]. For these simulations, the analytic approach may be of particular interest, because, for the enlargement of the spacial scales, the computations become very demanding. Corresponding comparison with the theory also requires one to account for $\Phi(0) \neq 0$. In accordance with this, two qualitatively different situations are possible: $|a_0| \lesseqgtr |\sin\Phi_0|$. The VEAH solutions for $|a_0| > |\sin\Phi_0|$ are well represented by Fig. 1, while the opposite case is illustrated by Fig. 2. The specific case where $|a_0| = |\sin\Phi_0|$ is also presented (Fig. 2c). The case where $|a_0| < |\sin\Phi_0|$ is represented by the bipolar multi-instanton solution for the magnetic field amplitude. For the initial conditions $|a_0| = |\sin\Phi_0|$, the solutions for $a(t)$ and n_z become the solitary instanton and kink modes, respectively. However, as $t \longrightarrow \infty$, the magnetic field and the inhomogeneity of the anisotropy for this solution disappear, i.e., it is unstable, and the Weibel instability must arise again.

The period t_0 of the nonlinear solution (6) also depends on the sign of the expression $a_0^2 - \sin\Phi_0^2$. Denoting $p^2 \equiv a_0^2 - \sin\Phi_0^2$, one finds

$$t_0 = (4/\gamma_0\sqrt{1+p^2})K(1/\sqrt{1+p^2}) \qquad (7)$$

for the time required for a complete cycle of anisotropy rotation in the case $|a_0| > |\sin\Phi_0|$. Similarly, for $-p^2 \equiv a_0^2 - \sin\Phi_0^2 < 0$ this period reads

$$t_0 = (4/\gamma_0)K(\sqrt{1-p^2}), \qquad (8)$$

where $0 < p < 1$. In Eqs. (7) and (8), $K(u)$ is the complete elliptic integral of the first kind with the modulus u. In accordance with Figs. 1 and 2, the period decreases with p (with the magnetic field amplitude in $\Phi_0 \longrightarrow 0$ case), as shown in Fig. 3, where the period is in units of γ_0. The soliton solution corresponds to $p = 0$. As $p \longrightarrow 1$, Eq. (8) gives the period $t_0 = 2\pi$. For example, in the case of the standard problem definition for the Weibel instability, $n_z(0) = 1$, the period of nonlinear pulsations is $t_0 \simeq 8\pi$ for $a_0 = 0.01$ and $t_0 \simeq 4\pi$ for $a_0 = 0.25$.

The averaged (over the period) magnetic energy is small, $\langle B^2\rangle/8\pi n_e T_\perp \ll 1$, because $ck < \omega_p$ and the duration of nonlinear pulsations is shorter than t_0 providing $\langle a^2\rangle < 1$, i.e.,

$$\frac{\langle B^2\rangle}{8\pi n_e T_\perp} = \frac{\langle a^2\rangle}{2}\left(1 - \frac{T_\parallel}{T_\perp}\right)\frac{c^2k^2}{\omega_p^2}, \qquad (9)$$

where n_e is the electron density. This is illustrated by Fig. 4 for $\Phi_0 = 0$, where the dimensionless magnetic energy $\langle a^2\rangle$ has a very weak dependence on a_0. A small saturated magnetic field energy $\langle B^2\rangle/8\pi n_e T_\perp \sim 0.02$ has also been observed in PIC simulations [8], where in the final state the single-mode regime was observed with

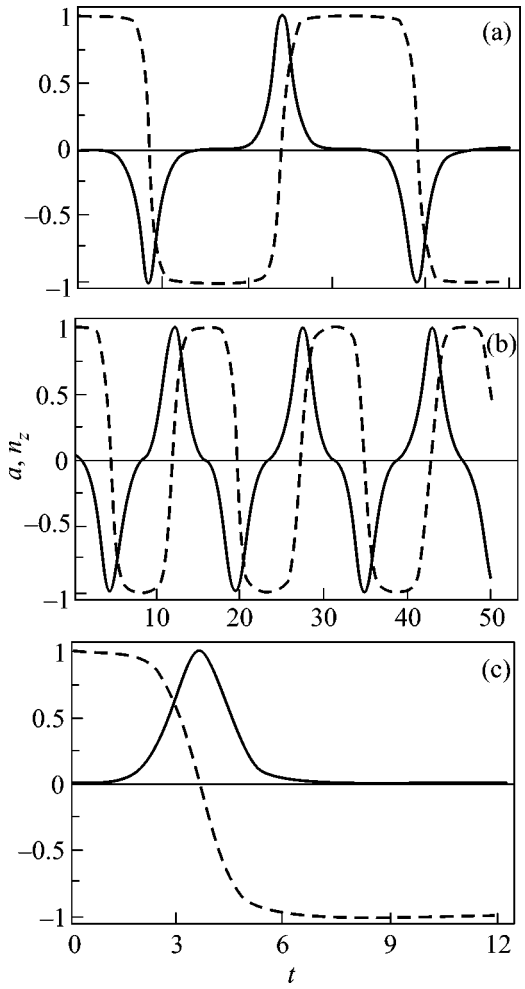


Fig. 2. Evolution of the magnetic field (solid lines) and the anisotropy axis (dashed lines) for $a_0 = 0.001$, $\Phi_0 = 0.002$ (a), $a_0 = 0.05$, $\Phi_0 = 0.1$ (b), and $a_0 = \sin\Phi_0$, $\Phi_0 = 0.05$ (c).

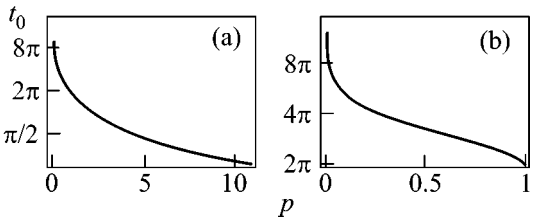


Fig. 3. Period of the nonlinear solutions of VEAH for (a) $a_0^2 - \sin\Phi_0^2 > 0$ (logarithmic scale) and (b) $a_0^2 - \sin\Phi_0^2 < 0$.

$ck/\omega_p \simeq 0.5$ and the saturated anisotropy $T_\parallel/T_\perp \simeq 1/3$. Assuming from Fig. 4 a rank value $\langle a^2\rangle \sim 1/4$, one can estimate from Eq. (9) that $\langle B^2\rangle/8\pi n_e T_\perp$ is very close to the PIC simulation results [8].

In summary, we have obtained an analytical solution to the VEAH equations (3), which describe the relax-

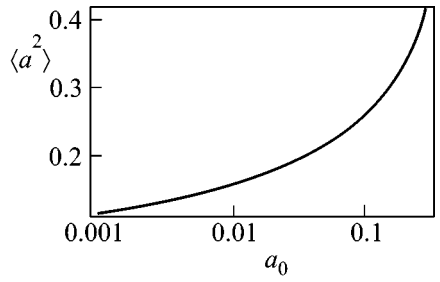


Fig. 4. Averaged dimensionless magnetic field energy.

ation of the long wavelength Weibel instability that originates from the anisotropy of the electron temperature. The mechanism responsible for the saturation of the instability is the rotation of the anisotropy axis rather than the temperature becoming isotropic. Such a rotation leads to dephasing of the generated magnetic field with respect to the source of the anisotropy. The solution that has been found corresponds to a single-mode circular-polarized magnetic field. Our theoretical model explains the formation of single-mode magnetic structures that are observed in PIC simulations [8]. The characteristics of the Weibel plasma have been derived, and this can have a fundamental significance and potential application in the implementation of the forefront XFEL project [13].

This work was partly supported by the Natural Sciences and Engineering Research Council of Canada and the Russian Foundation for Basic Research (grant no. 03-02-16428), and the International Science and Technology Center (grant no. 2104).

REFERENCES

1. E. M. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
2. H. Wabnitz, L. Bittner, A. R. B. de Castro, *et al.*, Nature **420**, 482 (2002).
3. R. L. Morse and C. W. Nielson, Phys. Fluids **14**, 830 (1971).
4. J. M. Wallace and E. M. Epperlein, Phys. Fluids B **3**, 1579 (1991).
5. R. C. Davidson, D. A. Hammer, I. Haber, and C. E. Wagner, Phys. Fluids **15**, 317 (1972).
6. V. Yu. Bychenkov, V. N. Novikov, V. P. Silin, and V. T. Tikhonchuk, Sov. J. Plasma Phys. **17**, 272 (1991); V. Yu. Bychenkov, V. N. Novikov, and V. P. Silin, Sov. J. Plasma Phys. **17**, 485 (1991); V. Yu. Bychenkov and V. N. Novikov, Plasma Phys. Rep. **23**, 670 (1997).
7. V. Yu. Bychenkov, A. I. Golubev, N. A. Ismailova, *et al.*, Plasma Phys. Rep. **26**, 54 (2000).
8. D. V. Romanov, V. Yu. Bychenkov, W. Rozmus, *et al.*, Bull. Am. Phys. Soc. **47**, 87 (2002).
9. V. Yu. Bychenkov, V. P. Silin, and V. T. Tikhonchuk, Sov. J. Plasma Phys. **15**, 407 (1989); Phys. Lett. A **138**, 127 (1989); Theor. Math. Phys. **82**, 11 (1990).
10. V. Yu. Bychenkov, V. F. Kovalev, and V. V. Pustovalov, Plasma Phys. Rep. **22**, 999 (1996); V. A. Terekhin, E. V. Uvarov, V. Yu. Bychenkov, and V. F. Kovalev, Plasma Phys. Rep. **25**, 409 (1999).
11. V. A. Terekhin, V. T. Tikhonchuk, and E. V. Uvarov, Phys. Lett. A **254**, 210 (1999); Plasma Phys. Rep. **26**, 308 (2000).
12. V. Yu. Bychenkov, V. P. Silin, and V. T. Tikhonchuk, Sov. Phys. JETP **71**, 709 (1990).
13. T. Tschentscher, Proc. SPIE **4500**, 1 (2001).

JETP Letters, Vol. 78, No. 3, 2003, pp. 123–128. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 154–159.

Self-Consistent Expansion of a Multicomponent Plasma into a Vacuum

A. A. Ivanov[1] and K. S. Serebrennikov[2]

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

[1]*e-mail: ivanov@ard.kiae.ru*

[2]*e-mail: kss@ard.kiae.ru*

Received March 23, 2003; in final form, June 26, 2003

The problem of the one-dimensional collisionless expansion of a multicomponent plasma into a vacuum is considered. In the hydrodynamic approximation, an approximate analytical solution for an arbitrary set of ion species with masses $M_1, \ldots, M_n$ and charge numbers $Z_1e, Z_2e, \ldots, Z_ne$ is found by using the technique of self-similar variables employed by Gurevich, Pariĭskaya, and Pitaevskiĭ for the case of single-species ions. A numerical iterative algorithm is developed in which the analytical solution is used as a first approximation. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 52.27.Cm; 52.30.-q

The problem of the one-dimensional expansion of plasma into a vacuum in the hydrodynamic formulation is a classical problem of the physics of plasma in which the electron temperature greatly exceeds the ion temperature. In the pioneering works of Gurevich, Pariĭskaya, and Pitaevskiĭ, a self-consistent analytical solution was obtained for single-species ions [1]. The authors used the technique of introducing a self-similar variable (developed in ideal fluid dynamics [2]), which enabled them to reduce the total number of variables. The problem was solved using the approximations of plasma quasi-neutrality and isothermality of electrons. In [3], the technique was generalized to the case of more than one ion species. The acceleration of ions at the expanding plasma front, where the quasi-neutrality approximation is no longer valid, was studied both analytically and numerically by Crow *et al.* [4]. It was shown that the ions accelerate infinitely fast and, behind the front, the self-similar solution from [1] is generally applicable. Mora and Pellat [5] obtained analytical corrections that take into account electron cooling near the front caused by the transfer of energy for the acceleration of ions. Chan *et al.* [6] (see also references therein) verified experimentally that, behind the ion front, the potential is well described by the self-similar solution from [1] until the effects of confined geometry come into play.

Ivanov *et al.* [7–9] showed that the propagation of a spatially localized group of hot electrons into a plasma was also self-similar. Based on the approach [10], Ivanov *et al.* developed a self-consistent kinetic theory of the relaxation of a low-pressure quasineutral negative-ion plasma [11, 12]. The solution was obtained for the case of two kinds of ions (negative and positive) with comparable densities.

Kovalev *et al.* [13] used the group theoretical methods to obtain an analytical solution for the expansion of a plasma bunch in the kinetic formulation. Note that only the quasi-neutrality approximation was employed by the authors.

At present, in connection with progress in the physics and performance of multicharged-ion sources [14], the problem of the expansion of a multicomponent plasma in which each component is multiply ionized [15, 16] has attracted considerable interest.

1. Model and basic equations. Let us consider the problem of the one-dimensional expansion of a multicomponent plasma into a vacuum. The plasma consists of electrons and ions with charge numbers $Z_1e, Z_2e, \ldots, Z_ne$ and masses $M_1, \ldots, M_n$. The plasma is homogeneous and initially resides in the left half-space $x < 0$. Such a situation is characteristic of, say, the plasma of an electron cyclotron discharge in a magnetic mirror at the instant of switching off the external microwave field. Generally, the temperature of the electron component T_e is much higher than that of the ion component, and, hence, one can use the approximation of cold hydrodynamics for ions

$$\frac{\partial N_k}{\partial t} + \frac{\partial}{\partial x}(N_k v_k) = 0, \tag{1}$$

$$\frac{\partial v_k}{\partial t} + v_k\frac{\partial v_k}{\partial x} + \frac{Z_k e}{M_k}\frac{\partial \varphi}{\partial x} = 0. \tag{2}$$

Here, N_k and v_k are the density and the mean velocity of the kth ion species, respectively, and φ is the electro-

static potential induced by charge separation. Electrons move much faster than ions and hence manage to attain a quasisteady thermodynamic equilibrium. To obtain an analytical solution, we assume that the distribution function of electrons is Maxwellian at $x \longrightarrow -\infty$. Then their spatial distribution is described by the Boltzmann distribution

$$N_e = N_0 \exp(e\varphi/T_e). \quad (3)$$

With the exception of the rightmost region (the front), in which the electron density is higher than the ion one, the quasi-neutrality equation is true

$$\sum_{k=1}^{n} Z_k N_k = N_e. \quad (4)$$

As shown in [1, 17], the solution to the set of equations of cold hydrodynamics can be sought in the self-similar form, since the initial conditions do not contain any space scale. In our case, there can be several different ion-acoustic velocities $C_{sk} = \sqrt{T_e/M_k}$; however, the initial and the boundary conditions are identical for all the ion species. Let us introduce a normalized self-similar variable $\xi = \frac{x}{t}\sqrt{\frac{M_1}{T_e}}$ and assume that all the variables depend only on ξ. After introducing the dimensionless variables

$$u_k = \frac{v_k}{\bar{v}_s}; \quad \bar{v}_s = \sqrt{\frac{T_e}{M_1}}; \quad \psi = \frac{e\varphi}{T_e}; \quad n_k = \frac{N_k}{N_0},$$

Eqs. (1)–(4) take the form

$$(u_k - \xi)\frac{dn_k}{d\xi} + n_k \frac{du_k}{d\xi} = 0, \quad (5)$$

$$(u_k - \xi)\frac{du_k}{d\xi} + A_k \frac{d\psi}{d\xi} = 0, \quad (6)$$

$$\psi = \ln\left(\sum Z_k n_k\right), \quad (7)$$

where $A_k = Z_k M_1/M_k$. The set of Eqs. (5)–(7) has a trivial solution that satisfies the boundary conditions in the left half-space ($x < 0$, $t \longrightarrow 0$, $\xi \longrightarrow -\infty$):

$$\psi = 0, \quad n_k = n_k^0, \quad u_k = 0, \quad \sum Z_k n_k^0 = 1. \quad (8)$$

However, this solution does not satisfy the boundary conditions in the right half-space ($x > 0$), in which, at $t = 0$ or $\xi \longrightarrow \infty$, there must be no particles: $n_k = 0$. In the general case, the existence condition for the nontrivial solution is obtained by zeroing the determinant of the matrix composed of the coefficients of the derivatives with respect to ξ in Eqs. (5) and (6). However, this can be done in a simpler way. On substituting $\psi(\xi)$ from Eq. (7) into Eq. (6), we obtain

$$\frac{du_k}{d\xi} = -\frac{A_k}{(u_k - \xi)} \frac{\sum Z_i (dn_i/d\xi)}{\sum Z_i n_i}.$$

Substituting this in Eq. (5), we obtain

$$\frac{dn_k}{d\xi} = -\frac{A_k n_k}{(u_k - \xi)^2} \frac{\sum Z_i (dn_i/d\xi)}{\sum Z_i n_i},$$

which, for the nontrivial case $\sum Z_i (dn_i/d\xi) \neq 0$, finally gives, after summing over k, the following matching condition:

$$\sum \frac{Z_k A_k n_k}{(u_k - \xi)^2} = \sum Z_k n_k. \quad (9)$$

At $t > 0$, the trivial solution exists in the left half-space, which has not yet been affected by the process of plasma expansion. Both solutions are joined at certain ξ_0 to meet all the boundary conditions, and the solution obtained has a derivative discontinuity at the joining point. The set of Eqs. (5)–(7) with additional condition (9) is nonlinear, and finding an exact analytical solution seems to be hardly possible. However, it was shown in [1] that, for a certain simple case, it is easy to obtain an analytical solution. For the single-species ions, when $n_k = n_1\delta_{1,k}$, $u_k = u_1\delta_{1,k}$, and $A_k = Z_1\delta_{1,k}$, condition (9) takes a simple form which itself gives the solution

$$(u_1 - \xi)^2 = Z_1 \quad (10)$$

or

$$u_1 = \xi + \sqrt{Z_1}. \quad (11)$$

Substituting the result obtained into Eq. (5), as written for $k = 1$, we readily have

$$n_1 = \text{const} \cdot \exp(-\xi/\sqrt{Z_1} - 1).$$

Here, the constant, as well as the joining point ξ_0, are to be obtained from the matching condition. Namely, at $\xi = \xi_0$, $u_1(\xi) = 0$ and $n_1(\xi) = 1/Z_1$, and hence $\xi_0 = -\sqrt{Z_1}$ and

$$n_1 = \frac{1}{Z_1}\exp\left(-\frac{1}{\sqrt{Z_1}} - 1\right). \quad (12)$$

At $\xi < \xi_0$, there exists a trivial solution $u_1(\xi) = 0$ and $n_1(\xi) = 1/Z_1 = \text{const}$. Then, finally, from Eq. (7) we have the normalized potential

$$\psi = -\xi/\sqrt{Z_1} - 1. \quad (13)$$

Figure 1 shows the solution for the single-species ions with $Z_1 = 1$. It should be noted that, in fact, a jump of the electric field moves left with the ion-acoustic velocity. The attempt [3] to find a solution for the case

of two kinds of ions was restricted to the linear approximation with the n_2/n_1 ratio being considered as a series expansion parameter.

2. The linear potential approximation. By analogy with the case of single-species ions, we will seek an approximate solution, assuming that the potential depends linearly on ξ when $\xi > \xi_0$ and equals zero when $\xi < \xi_0$. In the nontrivial-solution region, let us write

$$\psi = \tilde{\psi} = -p\xi - 1. \quad (14)$$

Here, parameter p may be dependent on the charges and masses of all the ions. As the potential is zero to the left of ξ_0, the joining point is related to p through simple relation $p = -1/\xi_0$. Note that the joining point is obviously the same for all the ion species with different k. Equations (5) and (6) take the form

$$\frac{d\ln n_k}{d\xi} = -\frac{1}{u_k - \xi}\frac{du_k}{d\xi}; \quad (15)$$

$$(u_k - \xi)\frac{du_k}{d\xi} = A_k p. \quad (16)$$

The latter equation is a particular case of the Abel equation of the second kind and can be solved rigorously [18]. Note first that the set of Eqs. (14)–(16) has an exact integral solution. Indeed, consider the derivative

$$\frac{d\ln(u_k - \xi)}{d\xi} = \frac{1}{(u_k - \xi)}\frac{du_k}{d\xi} - \frac{1}{(u_k - \xi)}.$$

Rewriting the last term using Eq. (16), we obtain

$$\frac{1}{(u_k - \xi)}\frac{du_k}{d\xi} = \frac{d\ln(u_k - \xi)}{d\xi} + \frac{1}{A_k p}\frac{du_k}{d\xi}.$$

Then, by using Eq. (15), rewrite the left-hand side and integrate it to obtain

$$\ln(u_k - \xi) + \frac{1}{A_k p}u_k + \ln n_k = \text{const}. \quad (17)$$

From the boundary conditions $\xi = \xi_0$, $u_k = 0$, and $n_k = n_{k0}$, we find that the constant on the right-hand side of Eq. (17) is const = $\ln(n_{k0}) + \ln(-\xi_0)$.

Then, finally

$$n_k = n_{k0}\frac{-\xi_0}{u_k - \xi}\exp\left(-\frac{u_k}{A_k p}\right). \quad (18)$$

Thus, under the assumption given by Eq. (14), we obtained an exact solution that relates the density and the velocity of the kth ion species. It is necessary now to find the explicit dependence of u_k on ξ. We will do this taking into account that Eq. (16) (the Abel equation of the second kind) admits a particular solution of the kind

$$g_k(\xi) = \xi + A_k p. \quad (19)$$

If we considered here the case of single-species ions, this solution would be the sought one [see Eq. (11)].

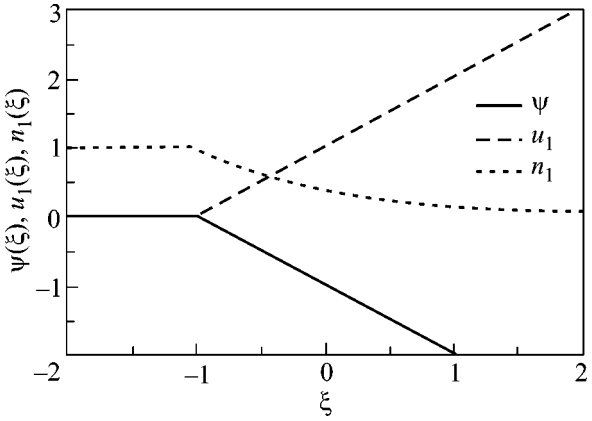


Fig. 1. Functions $u_1(\xi)$, $n_1(\xi)$, and $\psi(\xi)$ for $Z_1 = 1$. The joining point is $\xi_0 = -1$.

However, it is not our case. Moreover, the ξ value for which $g_k(\xi) = 0$ depends on k. Hence, the solution discussed does not satisfy the boundary conditions. Let us seek the solution in the form

$$u_k(\xi) = g_k(\xi) + \tilde{u}_k(\xi) = \xi + A_k p + \tilde{u}_k(\xi); \quad \tilde{u}(\xi_0) = -(\xi_0 + A_k p). \quad (20)$$

Such a choice allows one to meet the boundary conditions at the joining point. Substituting Eq. (20) into Eq. (16), we have

$$(\xi + A_k p + \tilde{u}_k - \xi)\left(1 + \frac{d\tilde{u}_k}{d\xi}\right) = A_k p; \quad \frac{d\tilde{u}_k}{d\xi} = -\frac{\tilde{u}_k}{A_k p + \tilde{u}_k}.$$

Its solution is a transcendental algebraic equation

$$\tilde{u}_k - \tilde{u}_k^0 + A_k p \ln\frac{\tilde{u}_k}{\tilde{u}_k^0} = -(\xi - \xi_0), \quad (21)$$

where $\tilde{u}_k^0 = \tilde{u}_k(\xi_0)$. Now, if we construct a function $\xi = \xi(\tilde{u}_k)$ by using Eq. (21) and swap the axes, we find the sought solution for the velocity of the kth ion species:

$$u_k(\xi) = \xi + A_k p + \tilde{u}_k(\xi) = \tilde{u}_k(\xi) - \tilde{u}_k^0 + \xi - \xi_0. \quad (22)$$

In Eq. (22), only the ξ_0 quantity has not yet been known. It can be easily obtained from Eq. (9) by substituting $\xi = \xi_0$, $u_k(\xi_0) = 0$, and $n_k(\xi_0) = n_{k0}$:

$$\sum\frac{Z_k A_k n_{k0}}{(-\xi_0)^2} = \sum Z_k n_{k0}; \quad (-\xi_0) = \sqrt{\frac{\sum Z_k A_k n_{k0}}{\sum Z_k n_{k0}}}. \quad (23)$$

The solution given by Eq. (21) can also be represented through the function G inverse to xe^x

$$\tilde{u}_k(\xi) = A_k p \cdot G\left[\frac{\tilde{u}_k^0}{A_k p}\cdot\exp\left(-\frac{g_k(\xi)}{A_k p}\right)\right].$$

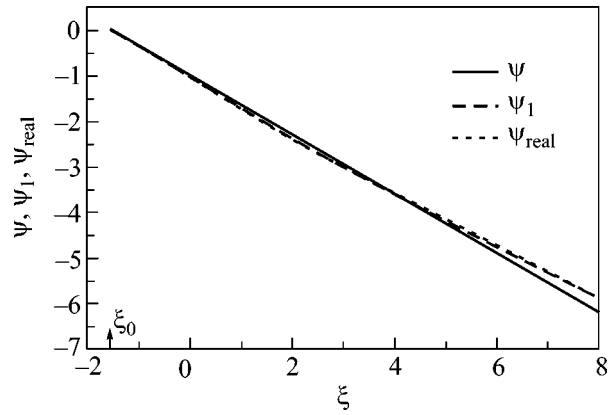


Fig. 2. Linear potential ψ, potential ψ_1 refined by using densities ψ_1, and real potential χ_{real}.

Let us consider the limiting cases. In the vicinity of the point ξ_0, where $\xi - \xi_0 \ll |\xi_0|$, we have, according to Eq. (21),

$$\tilde{u}_k - \tilde{u}_k^0 + A_k p \ln\frac{\tilde{u}_k}{\tilde{u}_k^0} = \Delta\tilde{u}_k + A_k p \ln\left(1 + \frac{\Delta\tilde{u}_k}{\tilde{u}_k^0}\right)$$

$$\approx \Delta\tilde{u}_k\left(1 + \frac{A_k p}{\tilde{u}_k^0}\right) = -\frac{\xi_0}{\tilde{u}_k^0}\Delta\tilde{u}_k \approx -(\xi - \xi_0)$$

or

$$u_k(\xi) = \Delta\tilde{u}_k + \xi - \xi_0 \approx -(\xi - \xi_0)\left(1 + \frac{A_k p}{\xi_0}\right) + (\xi - \xi_0) = \frac{(\xi - \xi_0)}{(-\xi_0)}A_k p, \quad (24)$$

$$n_k = n_{k0}\frac{-\xi_0}{u_k - \xi}\exp\left(-\frac{u_k}{A_k p}\right)$$

$$\approx n_{k0}\frac{-\xi_0}{\Delta\tilde{u}_k - \xi_0}\exp\left(-\frac{\xi - \xi_0}{(-\xi_0)}\right) \quad (25)$$

$$\approx n_{k0}\left(1 + \frac{\Delta\tilde{u}_k}{\xi_0}\right)\left(1 - \frac{(\xi - \xi_0)}{(-\xi_0)}\right) \approx n_{k0}\left(1 - A_k p\frac{\xi - \xi_0}{\xi_0^2}\right).$$

Consider the case of large positive ξ. Let us first rewrite Eq. (21) in a more appropriate form by using Eq. (22):

$$A_k p \ln\frac{\tilde{u}_k}{\tilde{u}_k^0} = -(\tilde{u}_k - \tilde{u}_k^0) - (\xi - \xi_0),$$

$$\tilde{u}_k = \tilde{u}_k^0 \exp\left(-\frac{u_k}{A_k p}\right).$$

Suppose that, at $\xi \gg 1$, $u_k = g_k + \tilde{u}_k \approx g_k = \xi + A_k p \gg \tilde{u}_k$. Indeed, for large ξ, we then have

$$\tilde{u}_k = \tilde{u}_k^0 \exp\left(-\frac{g_k}{A_k p}\right) = \tilde{u}_k^0 \exp\left(-\frac{\xi}{A_k p} - 1\right) \ll \tilde{u}_k^0$$

and our assumption that $\tilde{u}_k$ can be neglected is thus justified:

$$u_k = g_k + \tilde{u}_k \approx g_k = \xi + A_k p. \quad (26)$$

Using (18), we obtain for the densities

$$n_k = n_{k0}\frac{-\xi_0}{u_k - \xi}\exp\left(-\frac{u_k}{A_k p}\right) \approx n_{k0}\frac{1}{A_k p^2}\exp\left(-\frac{\xi}{A_k p} - 1\right). \quad (27)$$

It follows that, at large ξ, the largest contribution to the density is made by the term with the highest A_k value (i.e., with the highest charge-to-mass ratio). Let us denote the corresponding k as k'.

3. On the accuracy of the linear potential approximation. Remind that we expressed parameter p through ξ_0 (whose exact value is given by Eq. (23)) using Eq. (14) as follows:

$$p = -\frac{1}{\xi_0} = \sqrt{\frac{\sum Z_k n_{k0}}{\sum Z_k A_k n_{k0}}}. \quad (28)$$

If we derive, however, parameter p from the condition given by Eq. (9) using asymptotic expressions (26) and (27) and retaining only the terms with $k = k'$, we will have

$$\frac{Z_{k'}A_{k'}n_{k'}}{(A_{k'}p)^2} = Z_{k'}n_{k'}; \quad p = \sqrt{\frac{1}{A_{k'}}}, \quad (28')$$

which differs from Eq. (28). Of course, this is because approximation (14) is generally speaking not valid for the case of many kinds of ions. At large ξ, the deviations of the potential from linearity come into play. Note, however, that Eq. (28) adequately describes the transition to the case of single-species ions described in Section 1.

Let us illustrate the above by the example of three ion species with equal masses and charges $Z_{1,2,3} = 1$, 2, and 3. The slight deviation of the potential from linearity is illustrated in Fig. 2, which shows the straight line from Eq. (14) and the refined potential ψ_1 calculated by the equation $\psi_1 = \ln(\sum Z_k n_k(\xi))$, in which n_k is taken from Eq. (18). Figure 2 also shows the real potential ψ_{real}, which is to be calculated in the next section. The velocities and densities of the ion components calculated by Eqs. (18), (21), and (22) are shown in Figs. 3 and 4.

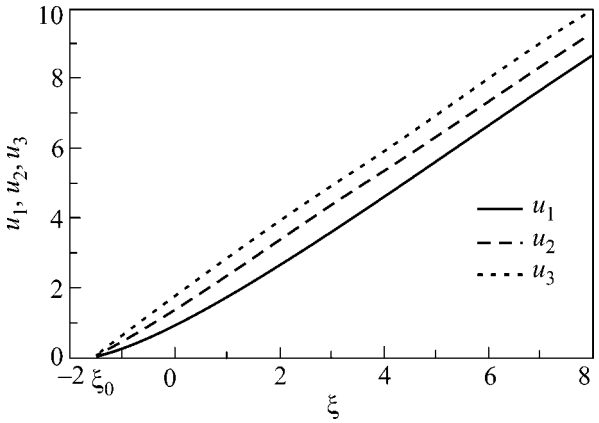


Fig. 3. Normalized ion velocities in the linear potential approximation.

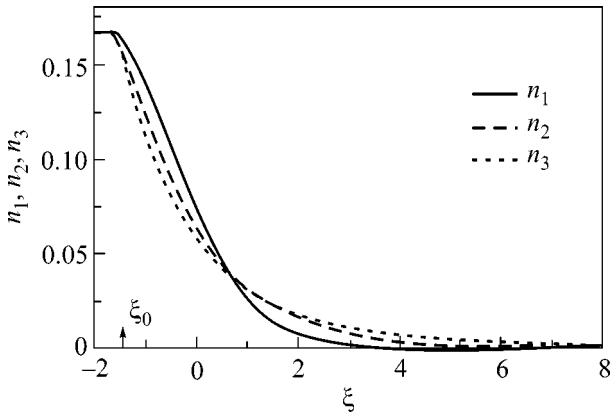


Fig. 4. Normalized ion densities in the linear potential approximation.

4. Iterative calculations of the real potential. To numerically simulate the potential in the case of many kinds of ions, it is suitable to use approximate expressions (18), (21), and (22) as a zero approximation. Then, as a first approximation for the potential, we have from Eq. (7)

$$\psi_1 = \ln\left(\sum Z_k n_k^0(\xi)\right), \tag{29}$$

where $n_k^0(\xi)$ stand for the functions obtained in Section 2. Now we will follow the scheme described in Section 2. To derive the refined ion velocities, we substitute the refined potential from Eq. (29) into the equation of motion (6) and have

$$(u_k^1 - \xi)\frac{du_k^1}{d\xi} = -A_k\frac{d\psi_1}{d\xi}. \tag{30}$$

This is the Abel equation of the second kind [18], though with a complicated function $\psi_1(\xi)$ on the right-hand side. It is not possible now to obtain an analytical solution; however, the equation can be easily solved numerically by employing net methods. To calculate the densities, it is reasonable first to integrate Eq. (15) from ξ_0 to ξ to yield the formula

$$n_k^1 = n_{k0}\frac{-\xi_0}{u_k^1 - \xi}\exp\left(-\int_{\xi_0}^{\xi}\frac{d\xi}{u_k^1 - \xi}\right), \tag{31}$$

which is similar to Eq. (18). To further refine the potential, the densities obtained by Eq. (31) are to be entered in $\psi_2 = \ln\left(\sum Z_k n_k^1(\xi)\right)$, and so on. The iterations rapidly converge to the exact solution ψ_{real}, which is shown in Fig. 2 for the case of three kinds of ions. It is seen that the linear potential provides a good approximation to the calculated ψ_{real}, whereas the first refinement of the potential, ψ_1, virtually coincides with ψ_{real}.

Note that numerical simulations (not presented in this paper) by the Vlasov kinetic equation for the case when the initial electron distribution function is Maxwellian so that the electron temperature is much higher than the ion temperature give results very close to the above hydrodynamic calculations.

We intend to continue studying the problem concerned by using the kinetic approach for arbitrary ion temperatures and a non-Maxwellian initial electron distribution.

We are grateful to A.A. Luk'yanov for helpful discussions. This work was supported by INTAS, grant no. 01-0373.

REFERENCES

1. A. V. Gurevich, L. V. Pariĭskaya, and L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **49**, 647 (1965) [Sov. Phys. JETP **22**, 449 (1965)]; Zh. Éksp. Teor. Fiz. **54**, 891 (1968) [Sov. Phys. JETP **27**, 476 (1968)]; A. V. Gurevich and L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **56**, 1778 (1969) [Sov. Phys. JETP **29**, 954 (1969)].
2. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Gostekhizdat, Moscow, 1964).
3. A. V. Gurevich, L. V. Pariĭskaya, and L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **63**, 516 (1972) [Sov. Phys. JETP **36**, 274 (1972)].
4. J. E. Crow, P. L. Auer, and J. E. Allen, J. Plasma Phys. **14**, 65 (1975).
5. P. Mora and R. Pellat, Phys. Fluids **22**, 2300 (1979).
6. C. Chan, N. Hershkowitz, A. Ferreira, *et al.*, Phys. Fluids **27**, 266 (1984).
7. A. A. Ivanov, L. L. Kozorovitskiĭ, and V. D. Rusanov, Dokl. Akad. Nauk SSSR **184**, 811 (1969) [Sov. Phys. Dokl. **14**, 126 (1969)].
8. A. A. Ivanov, Ya. N. Istomin, L. L. Kozorovitskiĭ, and V. D. Rusanov, Prikl. Mekh. Tekh. Fiz. **1**, 51 (1971).

9. A. A. Ivanov, V. D. Rusanov, and R. Z. Sagdeev, Pis'ma Zh. Éksp. Teor. Fiz. **12**, 29 (1970) [JETP Lett. **12**, 20 (1970)].
10. C. E. Hill and K. Langbein, Rev. Sci. Instrum. **69**, 643 (1998).
11. A. A. Ivanov, L. I. Elizarov, A. B. Sionov, and M. Bacal, Phys. Rev. E **52**, 6679 (1995).
12. A. A. Ivanov, A. B. Sionov, F. El. Balghiti-Sube, and M. Bacal, Phys. Rev. E **55**, 956 (1997).
13. V. F. Kovalev, V. Yu. Bychenkov, and V. T. Tikhonchuk, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 12 (2001) [JETP Lett. **74**, 10 (2001)].
14. R. Geler, *Electron Cyclotron Resonance Ion Sources and ECR Plasmas* (Inst. of Physics, Bristol, 1996).
15. A. Girard and G. Melin, Nucl. Instrum. Methods Phys. Res. A **382**, 252 (1996).
16. M. Kidera, M. Lamoureux, V. Mironov, *et al.*, Rev. Sci. Instrum. **70**, 4234 (1999).
17. A. V. Gurevich and L. P. Pitaevskiĭ, *Problems of the Plasma Theory* (Atomizdat, Moscow, 1980).
18. E. Kamke, *Gewöhnliche Differentialgleichungen* (Academie, Leipzig, 1959; Nauka, Moscow, 1976).

Translated by N. Ustinovskiĭ

JETP Letters, Vol. 78, No. 3, 2003, pp. 129–133. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 160–164.

Surface Ferroelectric Phase Transition in Multilayer Polymer Langmuir Films

T. V. Murzina[1], T. V. Misyuryaev[1], Yu. G. Fokin[1], S. P. Palto[2], S. G. Yudin[2], and O. A. Aktsipetrov[1,*]

[1]*Department of Physics, Moscow State University, Moscow, 117234 Russia*

[2]*Shubnikov Institute of Crystallography, Russian Academy of Sciences, Moscow, 117333 Russia*

**e-mail: aktsip@shg.ru*

Received June 16, 2003

The ferroelectric phase transition on the free surface of a polymer ferroelectric Langmuir–Blodgett film was studied by the optical second harmonic generation (SHG) technique. A hysteresis in the temperature dependence of the SHG intensity observed for a multilayer film of a poly(vinylidene fluoride)–trifluoroethylene copolymer in the vicinity of $T \approx 15$°C is a manifestation of the first-order ferroelectric phase transition in the topmost surface monolayer of the film. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 75.70.Cn; 77.84.Jd; 75.30.Kz; 68.47.Pe; 426.5.Ky

The surface contributions to the ferroelectric and magnetic phenomena, as well as the surface ferroelectric and ferromagnetic phase transitions, are of special interest in the physics of low-dimensional systems because dimensionality is a factor of basic importance in these phenomena. In recent decades, a considerable progress was achieved in experimental investigations of the surface magnetism and the two-dimensional (2D) phenomena in magnetic monolayers. This progress was related primarily to the development of molecular beam epitaxy (MBE) techniques, which made possible the obtaining of low-dimensional magnetic systems such as epitaxial monolayers on the surface of single crystals [1], epitaxial hetertostructures [2], and superlattices [3]. In addition, the discovery of the magnetic-field-induced optical second harmonic generation (SHG) [4] effect characterized by a high selective sensitivity with respect to the surface magnetic properties of centrosymmetric magnetic materials provided a basis for the development of a very effective optical method for the investigation of surface magnetic phenomena [5].

A different situation takes place in the field of observation and investigation of the ferroelectric phenomena on the surface and in 2D systems. On the one hand, thin ferroelectric films with a thickness below 20–30 nm cannot be synthesized even using the powerful MBE method. On the other hand, attempts to observe structural phase transitions on the surface of single crystals, not occurring in a high vacuum, by various methods (including nonlinear optical techniques [6]) should be rather referred to as studying the features of phase transitions in thick layers of micron thicknesses with the crystal structure damaged by polishing.

This situation changed in 1995, when thin ferroelectric Langmuir–Blodgett (LB) films were obtained based on poly(vinylidene fluoride) (PVDF) [7]. PVDF is a teflonlike polymer with a structural formula of $(CH_aF_{2-a}CH_bF_{2-b})_n$ (a, b = 0, 1, 2), possessing ferroelectric properties. Of special interest are the ferroelectric copolymers of PVDF with trifluoroethylene, denoted below as P(VDF–TrFE). These copolymers, in contrast to the pure PVDF, exhibit the phase transition from ferroelectric to paraelectric state at a temperature below the melting point [8]. Using the LB method, it is possible to obtain films of controlled thicknesses, even representing a single monomolecular layer (0.5 nm). The existence of 2D ferroelectric systems in the bulk of multilayer P(VDF–TrFE) films synthesized by the LB method was established by ferroelectric techniques (capacitive measurements of the permittivity) [9] and confirmed by the optical SHG method [10] possessing a unique sensitivity with respect to the break of the inverse symmetry of the sample structure upon the ferroelectric–paraelectric phase transition.

The results of dielectric measurements (requiring metal electrodes deposited onto the sample surface) showed that the LB-grown P(VDF–TrFE) films possess ferroelectric properties even at a thickness of several monolayers [11, 12]. The ferroelectric phase transition in the bulk of a multilayer film takes place at a temperature of ~80°C. Using dielectric techniques, the ferroelectric phase transition was also observed at an interface between the LB film and a metal electrode [13]. The temperature dependence of the permittivity of such films exhibits a hysteresis at ≈20°C, which was not observed previously in the bulk crystalline P(VDF–TrFE) samples. It is believed that this very transition is

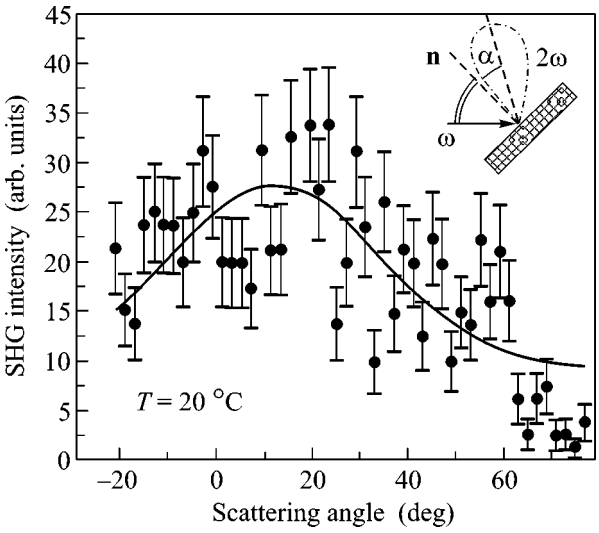


Fig. 1. SHG scattering indicatrix for a 20-ML Langmuir–Blodgett film of P(VDF–TrFE) at 20°C. The inset shows the experimental geometry (α is the polar scattering angle of the SHG radiation).

related to a ferroelectric ordering in the topmost layer of the LB film. However, we can hardly speak of observing a surface ferroelectric phase transition in this case, because the inavoidable metal electrode introduces strong perturbation into the structure and electron properties of the surface layer of a P(VDF–TrFE) film in such experimental systems. It is necessary to use a nonperturbative method providing information about a polar state of the free surface of a ferroelectric material.

Such an opportunity is offered by the optical SHG method not requiring electrodes on the sample surface. Using this method, Choi *et al.* [8] studied the samples of LB-grown films of a P(VDF–TrFE) copolymer without an upper metal electrode and determined the temperature-induced changes in the Brillouin zone structure in the vicinity of the "surface" phase transition. Variation of the sample temperature in the vicinity of this phase transition was accompanied by doubling of the Brillouin zone (indicative of a modification of the electron structure of the topmost surface layer) and by reorientation of the surface dipoles. However, no direct evidence of the ferroelectric nature of the observed transition was available so far.

In this study, we employed the optical SHG method to monitor the temperature-dependent variations in the polar state of the free surface (not perturbed by a metal electrode) of a multilayer ferroelectric LB-grown P(VDF–TrFE) copolymer film. The temperature-induced variations and hysteresis in the SHG response observed in the vicinity of $T \approx 15$°C are manifestations of the first-order ferroelectric phase transition on the free surface of a P(VDF–TrFE) film.

The experiments were performed with a series of multilayer ferroelectric LB films of a 70% PVDF–30% TrFE copolymer. The films were prepared [12] using an 0.01% solution of the copolymer in aqueous DMSO. The LB films were transferred onto fused quartz substrates by the Langmuir–Schaefer technique (a horizontal lift modification of the LB method). The given series included films with thicknesses ranging from 2 to 60 polymer monolayers (MLs).

The quadratic response of the ferroelectric LB films was excited by the radiation of a pulsed YAG:Nd^{3+} laser generating 15-ns pulses at a wavelength of λ_ω = 1064 nm. The incident beam intensity could be varied within 10–50 MW/cm^2. The SHG response from a film studied was separated by an interference filter, detected by a photoelectron multiplier, and measured by a gated electronics. The samples were placed in an optical thermostat capable of maintaining a preset sample temperature in the interval from –10°C to 100°C at a pressure of 10^{-2} Torr. The sample heating rate in our experiments was 1–3 K/min. The temperature was measured to within 1 K by a copper–constantan thermocouple fixed on the sample surface. The diffuse (nonspecular) component of the SHG radiation from an LB film could be measured by rotating the detection system about the vertical axis lying in the film plane.

In order to reveal peculiarities in the structure and symmetry of the nonlinear optical properties of the LB films at T = 20°C, we measured the azimuthal angular dependences of the SHG intensity for various combinations of polarizations of the fundamental radiation and the second harmonic wave. We have also measured the SHG intensity as a function of the polar scattering angle, representing the SHG scattering indicatrix.

It was found that the azimuthal angular dependence of the SHG response for the *s*–*s* combination of polarizations of the fundamental radiation and the second harmonic wave is isotropic, which contradicts the basic prohibition rule for the *s*-polarized SHG on a smooth isotropic surface (the so-called *s*–*s* prohibition) [14].

The SHG intensity distribution with respect to the polar angle does not contain a pronounced specular peak and shows scattering within a broad angular range (Fig. 1). A maximum in the SHG intensity is observed along the normal to the sample surface. Thus, the SHG is diffuse and depolarized (i.e., incoherent), representing the second-order hyper-Rayleigh scattering (HRS). The diffuse and depolarized character of the quadratic nonlinear optical response are indicative of the small-scale spatial inhomogeneity of the LB films studied, with a correlation length of this inhomogeneity much smaller as compared to the fundamental wavelength (1.06 μm) and the second harmonic wavelength (0.53 μm). A correlation length determined from the angular width of the HRS indicatrix [15] in the LB films studied is approximately 100 nm.

Figure 2 shows the temperature variation of the SHG intensity in a 60-ML film of P(VDF–TrFE) measured in the temperature interval from 0 to 110°C. As can be seen, the curve exhibits two regions of pronounced hysteresis. In the first of these, heating the sample above 60°C leads to an increase in the SHG intensity, followed by saturation at temperatures about ~100°C. On cooling the sample, the SHG intensity keeps increasing and drops at temperatures below ~60°C. The second region of hysteresis corresponds to a peak in the SHG intensity at 10–20°C. Figure 3 presents the temperature dependence of the SHG intensity for a 15-ML film, which also shows a pronounced hysteresis in the region of 15°C.

The SHG intensity variation with the temperature in the LB-grown P(VDF–TrFE) films is qualitatively similar to the pattern observed for the permittivity measured by the dielectric technique (see the inset in Fig. 2). This analogy allows us to ascertain that peculiarities observed in the behavior of the SHG intensity are also related to phase transitions in the P(VDF–TrFE) film and that the quadratic response reflects the ferroelectric properties of the polymer film. As was mentioned above, the temperature profile of the permittivity is determined by two ferroelectric phase transitions, occurring in the bulk and at the interface. The LB film regions responsible for these phase transitions are different. The phase transition in the region of 80°C is related to the ferroelectric properties of the bulk of a multilayer LB film and represents the ferroelectric phase transition known in bulk crystalline P(VDF–TrFE) samples. The phase transition at the interface, previously observed in the region of 20–30°C [16], was determined by the ferroelectric properties of the topmost layer of the film occurring in contact with a metal electrode. In our case of an LB film with the free surface, this phase transition is a purely surface ferroelectric phenomenon.

Note that our LB films with the free surface exhibit hysteresis in the SHG intensity at a lower temperature than that reported previously for the film–metal interface: the hysteresis loop is shifted by 5–10°C toward lower temperatures. This difference in the temperatures at which the phase transition is observed on the free surface and at the polymer film–metal electrode interface is probably related to a strong influence of the metal coating on the ferroelectric ordering in the topmost layer of the LB film.

It should also be noted that the free surface of the LB films of a P(VDF–TrFE) copolymer in our experiments is covered by a layer of adsorbed molecules and, in this sense, does not represent a clean surface. The role of such adsorbed layers in establishing a long-range ferroelectric order requires special investigations.

As was mentioned above, a special feature of SHG in the LB-grown P(VDF–TrFE) films is the incoherent and diffuse character of this phenomenon, which allows

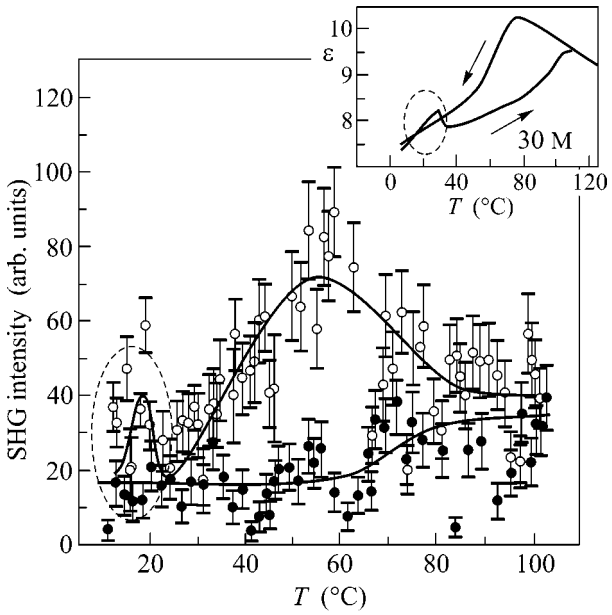


Fig. 2. The temperature variation of the SHG intensity for a 60-ML film of P(VDF–TrFE). The inset shows the analogous dependence of the permittivity of a 30-ML film of P(VDF–TrFE) measured by the dielectric method [11].

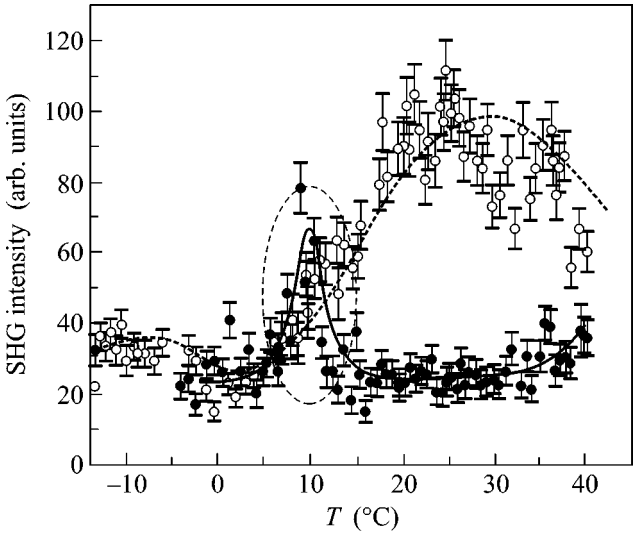


Fig. 3. The temperature variation of the SHG intensity for a 15-ML film of P(VDF–TrFE).

the process to be described in terms of the second-order hyper-Rayleigh scattering. In this case, the SHG intensity is determined by fluctuations of the quadratic polarization and, for ferroelectrics, of the spontaneous polarization as well. The temperature dependence of the HRS intensity may differ from that of the coherent SHG observed for ferroelectric crystals with a regular structure [17]. The incoherent and diffuse character of the SHG response is related to the inhomogeneity of the

film structure, primarily to the small-scale polydomain structure of the ferroelectric phase of the LB films studied. The characteristic size of inhomogeneity in the LB-grown P(VDF–TrFE) films is (according to the HRS indicatrix) on the order of 100 nm, which is much smaller than the fundamental wavelength and the second harmonic wavelength [15, 18].

The SHG intensity scatter per unit solid angle in the direction of the unit vector $\boldsymbol{\xi}$ is determined by averaging correlations of the nonlinear polarization fluctuations:

$$I_d \propto B_{ij}\int \langle \Delta_i(\mathbf{r})\Delta_j^*(0)\rangle \exp(-i\mathbf{qr})d^3r, \qquad (1)$$

where $\mathbf{q} = 2\mathbf{k}_\omega + \mathbf{k}_{2\omega}$; $\mathbf{k}_\omega$ and $\mathbf{k}_{2\omega}$ are the wavevectors of the fundamental and the second harmonic radiation, respectively; $B_{ij} = \sigma_{ij} + \xi_i\xi_j$; $\Delta_i(\mathbf{r}) = P_i(\mathbf{r}) - \langle P_i(\mathbf{r})\rangle$ is the correlation function of the nonlinear polarization P_i; and σ_{ij} is the Kronecker symbol. The absence of a specular contribution in our experiments implies that the ensemble-average of the nonlinear polarization is zero and that the source of diffuse incoherent SHG is fluctuations of the nonlinear polarization at the second harmonic wavelength.

This approach, whereby the SHG process in the LB-grown P(VDF–TrFE) films is described in terms of the second-order HRS, can explain the small value of the SHG response at temperatures below 10–20°C (where the films represent a polar, microscopically noncentrosymmetric ferroelectric phase). If the size of domains with oppositely oriented spontaneous polarization vectors is smaller than the light wavelength, the ferroelectric phase can be considered as a macroscopically centrosymmetric structure: the polar structure is masked by the spatial averaging over an area with a characteristic scale on the order of the light wavelength. In this phase, the SHG signal is incoherent and, hence, small and diffuse, being determined by spatial fluctuations $\Delta(\mathbf{r})$ of the nonlinear polarization. The small-scale character of the domain structure in our LB films is related to the fact that molecules of a P(VDF–TrFE) copolymer are not amphiphilic and, hence, they exhibit no spontaneous orientation on the water surface and are not oriented upon the transfer onto a solid substrate.

The presence of a temperature hysteresis, observed by both dielectric and nonlinear optical methods, is evidence that the corresponding phase transitions are of the first order and admit the coexistence of two phases (ferroelectric and paraelectric) in certain temperature intervals. This explains an increase in the intensity of incoherent SHG with the temperature in the regions of both phase transitions. As the temperature increases, the regions of paraelectric phase appear and give rise to spatial fluctuations of the spontaneous polarization. These fluctuations are related to the fluctuations of dimensions and orientations of ferroelectric domains in the temperature interval of the coexistence of two phases. Therefore, the incoherent SHG intensity will also increase in this temperature interval, because the spatial fluctuations of the nonlinear polarization are determined by increasing fluctuations of the spontaneous polarization.

Thus, we have observed the ferroelectric phase transition on the free surface of multilayer LB-grown P(VDF–TrFE) films, which was related to a ferroelectric ordering of the dipole moments of polymer molecules in the topmost layer of the LB films. A decrease in the temperature of the ferroelectric phase transition on the free surface as compared to that at a film–metal interface is explained by the strong influence of the metal electrode on the ferroelectric order in the topmost layer of a P(VDF–TrFE) film. The possible role of a layer of adsorbed molecules (primarily of water) present on the free surface of an LB film in establishing a long-range ferroelectric order in the topmost layer of the film requires special experimental investigations.

We have proposed a mechanism explaining the incoherent (diffuse and depolarized) SHG process in the regions of the surface and bulk phase transitions in ferroelectric P(VDF–TrFE) films. The incoherent character of the SHG response of these polymer LB films is related to fluctuations of the nonlinear polarization, which, in turn, are determined by fluctuations of the spontaneous polarization. The latter fluctuations exhibit a significant growth due to the coexistence of two phases in the vicinity of the first-order phase transition.

This study was supported by the Russian Foundation for Basic Research (project no. 01-02-04018), the Presidential Program “Leading Scientific Schools of Russia” (project no. 1604.2003.2), the Foundation for Promotion of National Science, and INTAS (grant no. 2002-113F1/b).

REFERENCES

1. N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
2. S. S. P. Parkin, N. More, and K. P. Roche, Phys. Rev. Lett. **64**, 2304 (1990).
3. M. Baibich, J. Broto, and A. Fert, Phys. Rev. Lett. **61**, 2472 (1988).
4. O. A. Aktsipetrov, O. V. Braginskiĭ, and D. A. Esikov, Kvantovaya Élektron. (Moscow) **17**, 320 (1990) [Sov. J. Quantum Electron. **20**, 259 (1990)].
5. J. Reif, J. C. Zink, C.-M. Schneider, and J. Kirschner, Phys. Rev. Lett. **67**, 2878 (1991).
6. E. D. Mishina, T. V. Misuryaev, N. E. Sherstyuk, *et al.*, Phys. Rev. Lett. **85**, 3664 (2000).
7. S. P. Palto, L. M. Blinov, A. Bune, *et al.*, Ferroelectr. Lett. Sect. **19**, 65 (1995).

8. J. Choi, P. A. Dowben, S. Ducharme, *et al.*, Phys. Lett. A **249**, 505 (1998).
9. A. Bune, S. Ducharme, V. M. Fridkin, *et al.*, Appl. Phys. Lett. **67**, 3975 (1995).
10. O. A. Aktsipetrov, T. V. Misuryaev, T. V. Murzina, *et al.*, Opt. Lett. **25**, 411 (2000).
11. S. Ducharme, A. V. Bune, L. M. Blinov, *et al.*, Phys. Rev. B **57**, 25 (1998).
12. A. V. Bune, V. M. Fridkin, S. Ducharme, *et al.*, Nature **391**, 874 (1998).
13. J. Choi, C. N. Borca, P. A. Dowben, *et al.*, Phys. Rev. B **61**, 5760 (2000).
14. O. A. Aktsipetrov, I. M. Baranova, and Yu. A. Il'inskiĭ, Zh. Éksp. Teor. Fiz. **91**, 287 (1986) [Sov. Phys. JETP **64**, 167 (1986)].
15. O. A. Aktsipetrov, A. A. Fedyanin, D. A. Klimkin, *et al.*, Ferroelectrics **183**, 215 (1996).
16. J. Choi, P. A. Dowben, S. Ducharme, *et al.*, Phys. Lett. A **249**, 505 (1999).
17. J. F. Legrand, Ferroelectrics **91**, 303 (1989).
18. T. V. Murzina, G. B. Khomutov, A. A. Nikulin, *et al.*, J. Opt. Soc. Am. B **17**, 63 (2000).

Translated by P. Pozdeev

JETP Letters, Vol. 78, No. 3, 2003, pp. 134–137. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 165–169.

Semiclassical Negative Magnetoresistance of a 2D Electron Gas Caused by Scattering by Short-Range and Long-Range Potentials

A. A. Bykov*, A. K. Bakarov, A. V. Goran, N. D. Aksenova, A. V. Popova, and A. I. Toropov

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

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

Received June 19, 2003

The magnetoresistance of a 2D electron gas confined in narrow GaAs quantum wells with AlAs/GaAs superlattice barriers is studied in the classical range of magnetic fields. It is shown that the negative magnetoresistance observed in this kind of structures with nonplanar heterointerfaces is semiclassical and qualitatively agrees with the model of negative magnetoresistance due to the scattering of charge carriers by two types of random potential, namely, the short-range and long-range ones. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 73.50.Jt; 73.63.Hs; 75.47.-m; 75.75.+a

The magnetoresistance (MR) of a 2D electron gas has been much studied, both theoretically and experimentally, in weak and quantizing magnetic fields. For the region of classical magnetic fields B, which lies between weak and quantizing magnetic fields, the theoretical analysis of MR in terms of the Boltzmann kinetic equation leads to the field-independent Drude expression: $\rho_{xx}(B) = \rho_0 = m/e^2 n_e \tau_{tr}$, where n_e is the concentration of charge carriers, m is the effective mass, and τ_{tr} is the transport relaxation time. However, as was first shown in [1], by including memory effects in the magnetotransport of two-dimensional charge carriers beyond the framework of the kinetic equation, one obtains purely classical reasons for the deviation of $\rho_{xx}(B)$ from constant. Depending on the character of the random scattering potential, the semiclassical MR of a 2D electron gas may be negative [2, 3] or positive [4]. The predicted negative semiclassical MR caused by the scattering of a 2D electron gas by a random inhomogeneous magnetic field [2] was recently observed experimentally [5]. However, the semiclassical magnetotransport properties of a 2D electron gas that are associated with the scattering of charge carriers by a long-range random electrostatic potential remain experimentally little investigated.

This paper is concerned with studying the MR of 2D electron gas in classical magnetic fields on narrow GaAs quantum wells with self-organized nonplanar heterointerfaces. The nonplanar configuration of heterointerfaces in such structures leads not only to a spatial modulation of the 2D electron gas [6] but also to the appearance of a long-range scattering potential correlated with the relief of the growth surfaces [7]. The character of the long-range potential is determined by the morphology of the growth surfaces and can be controlled by changing the conditions of the growth process. Thanks to the latter fact, narrow GaAs quantum wells with self-organized corrugated heterointerfaces are convenient model objects for an experimental study of the effect of the long-range scattering potential on the transport properties of a 2D electron gas.

The structures studied in our experiment were selectively doped GaAs quantum wells, 10 nm in thickness, with AlAs/GaAs superlattice barriers [8]. They were grown by molecular beam epitaxy (MBE) on (100) GaAs substrates with deviations from the (100) plane no greater than 0.02°. The surface morphology of the MBE structures was examined by atomic force microscopy (AFM). The transport properties of the 2D electron gas were studied for three types of MBE structures, which differed in the rms height w of the growth surface roughness: $w \sim 0.3$ (a smooth surface), 0.6 (an intermediate surface roughness), and 4 nm (a corrugated surface). The height of the growth surface roughness depended on the As_4 flux used in the growth process. The magnetotransport experiments were performed at a temperature of 4.2 K in magnetic fields up to 2 T. The Hall bars were fabricated by optical lithography and liquid etching. They were 50 μm wide and 100 μm long and were supplied with a planar Shottky gate, which allowed one to vary the concentration of the 2D electron gas in wide limits. The Hall bars were oriented so that the measuring current occurred along the [110] direction. At $T = 4.2$ K, the 2D electron gas in these structures was characterized by the following equilibrium parameters: the electron concentration $n_e = (1.2–1.7) \times 10^{12}$ cm^{-2} and the mobility $\mu_{xx} = (50–450) \times 10^3$ cm^2/(V s).

0021-3640/03/7803-0134$24.00 © 2003 MAIK "Nauka/Interperiodica"

Figure 1a shows the dependences $\rho_{xx}(B)$ obtained for the MBE structures with different heights of the growth surface roughness. The curve corresponding to the corrugated MBE structure (curve *1*) exhibits a pronounced negative MR in magnetic fields up to 1.2 T. In magnetic fields higher than 1.2 T, the MR becomes positive and performs Shubnikov–de Haas oscillations with increasing field. The curves obtained for the intermediate and smooth MBE structures (curves *2* and *3*) exhibit minor variation of MR in magnetic fields below 0.7 T. The Shubnikov–de Haas oscillations are observed in this case in smaller fields, as compared to the corrugated structure, and have higher amplitudes because of the higher mobility of 2D electrons in these structures. Figure 1b shows the behavior of the relative MR of 2D electron gas for the same MBE structures in magnetic fields up to 0.4 T. The dependences $\rho_{xx}(B)/\rho_0$ are different for different structures. For the corrugated MBE structure, the dependence of the MR on magnetic field is close to parabolic (in this field interval). For the intermediate structure, the curve $\rho_{xx}(B)/\rho_0$ has a maximum, which points to a quasi-one-dimensional periodic potential modulation of the 2D electron gas in this structure [7]. For the GaAs quantum well with smooth heterointerfaces, the MR of the 2D electron gas exhibits a linear behavior with a minor anomaly near zero magnetic field. A qualitatively similar behavior was theoretically predicted in [9].

Figure 2a presents the dependences $\rho_{xx}(B)$ for different concentrations of the 2D electron gas in a corrugated MBE structure. The negative MR clearly manifests itself for this structure in the whole concentration range under study. For the 2D electron concentration $n_e = 1.45 \times 10^{12}$ cm^{-2} near zero magnetic field, $\rho_{xx}(B)$ exhibits a maximum whose nature is the same as that in the dependences obtained for the MBE structures with the intermediate height of the growth surface relief. For $n_e = 1.21 \times 10^{12}$ cm^{-2}, the curve $\rho_{xx}(B)$ has a minimum at $B = B_{\min}$. The magnetic field $B_{\min}$ is determined by the condition $1/R_c^2 \sim n_s$ [10], where n_s is the concentration of short-range scatterers and R_c is the Larmor radius of the charge carriers. This field corresponds to the beginning of the classical localization of charge carriers by the short-range scattering potential in increasing magnetic field, while the resistance at $B = B_{\min}$ is determined by the scattering by the long-range potential.

Figure 2b shows the experimental and theoretical dependences $\Delta\rho_{xx}(B)/\rho_0$ for the 2D electron concentration $n_e = 1.21 \times 10^{12}$ cm^{-2}. Theoretical dependence *2* corresponds to the model of scattering by the short-range potential [11]: $\Delta\rho_{xx}/\rho_0 \sim -\exp(-2\pi/\beta)$, where $\beta = \omega_c\tau_{tr}$; theoretical dependence *3* corresponds to the scattering by both short-range and long-range potentials [3]: $\Delta\rho_{xx}/\rho_0 = -(\omega_c/\omega_0)^2$, where $\omega_0 \sim 1.67\nu_F n_s^{1/2}(l_s/l_L)^{1/4}$, ν_F is the Fermi velocity, $l_s = \nu_F\tau_s$ is the mean free path associated with the scattering from the short-range potential, and $l_L = \nu_F\tau_L$ is the mean free path associated with the scattering from the long-range potential. In estimating the quantity ω_0, we assumed that $\tau_{tr}^{-1} = \tau_s^{-1} + \tau_L^{-1}$. The quantity τ_L was calculated from the value of the MR at $B = B_{\min}$, τ_{tr} was calculated from $\rho_{xx}(B = 0)$, and n_s was found from $R_c(B_{\min})$. From Fig. 2b one can see that model *3* describes the behavior of the negative MR much better than model *2*.

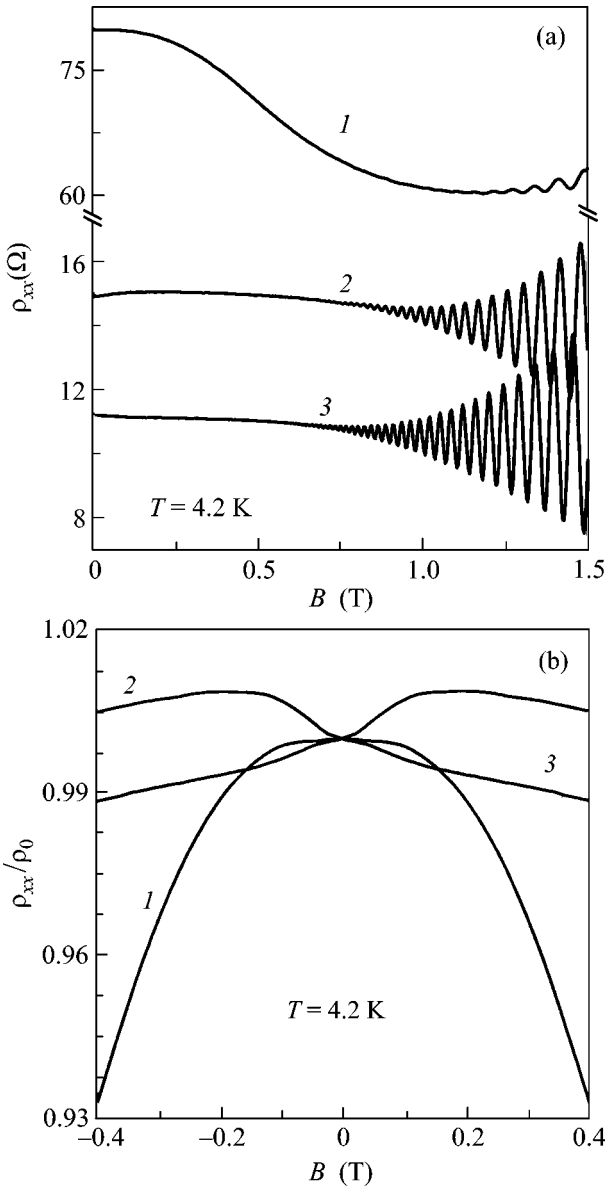


Fig. 1. (a) Dependences $\rho_{xx}(B)$ at $T = 4.2$ K for MBE structures with different rms height of the growth surface relief and (b) the relative MR for the same MBE structures: (*1*) corrugated, (*2*) intermediate, and (*3*) smooth structures.

The characteristic AFM image representing the surface morphology of a corrugated MBE structure is shown in Fig. 3a. The correlation analysis of the relief

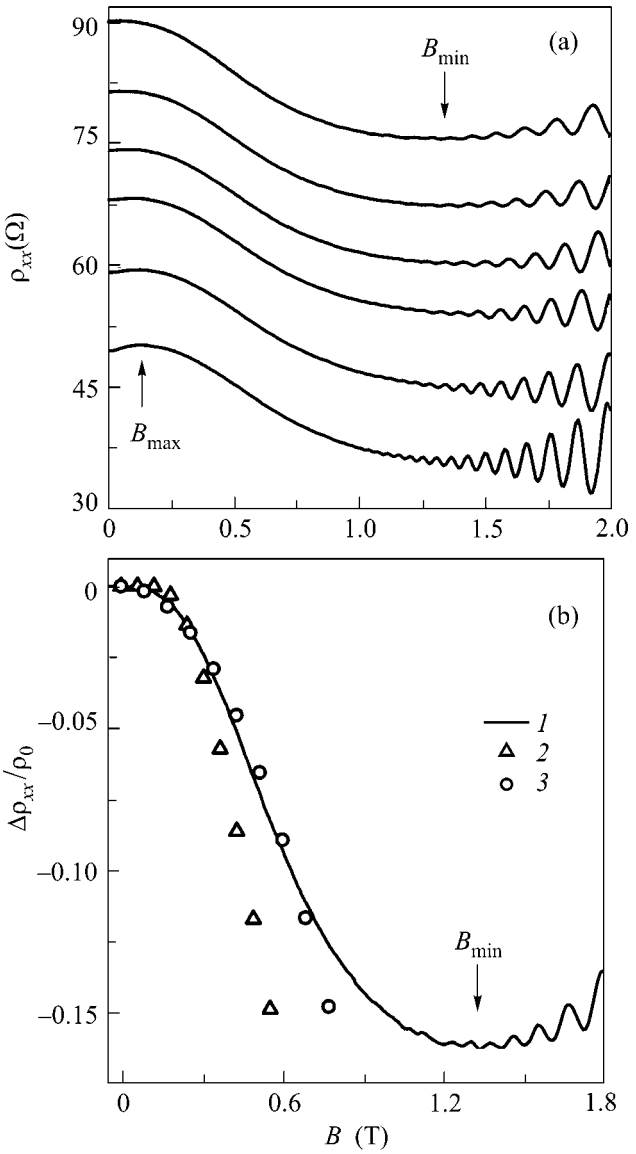


Fig. 2. (a) Longitudinal MR of the 2D electron gas in a corrugated epitaxial structure at $T = 4.2$ K for electron concentrations within $n_e = (1.21–1.45) \times 10^{12}$ cm^{-2} and mobilities within $\mu_{xx} = (57–87) \times 10^3$ cm^2/(V s) (from bottom to top). (b) Dependences $\Delta\rho_{xx}/\rho_0$ at $T = 4.2$ K for a corrugated structure at $n_e = 1.21 \times 10^{12}$ cm^{-2} and $\mu_{xx} = 57 \times 10^3$ cm^2/(V s): (*1*) experimental curve and theoretical dependences from (*2*) [11] and (*3*) [3].

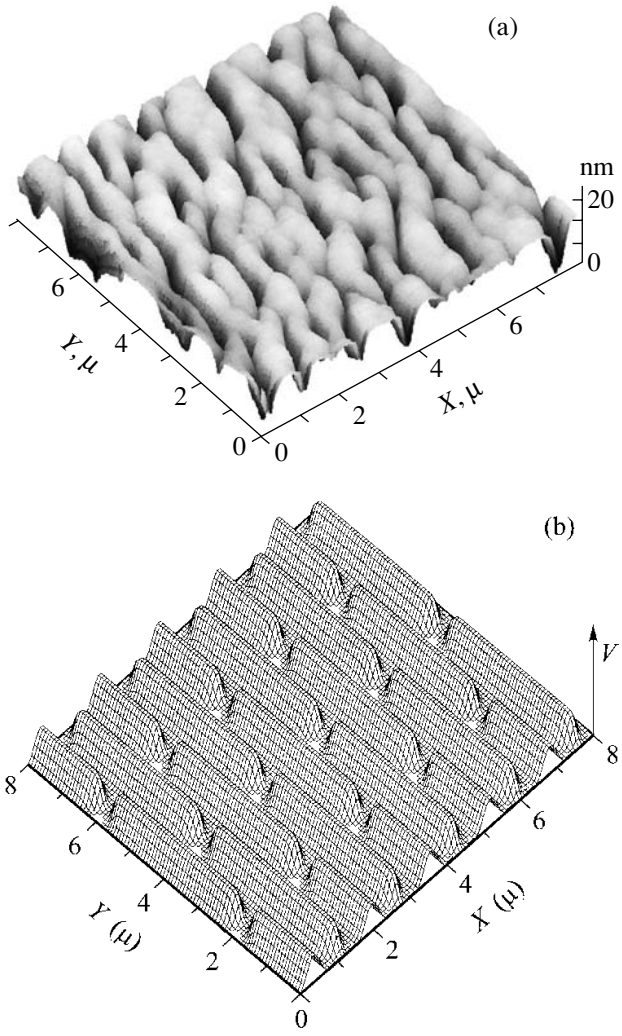


Fig. 3. (a) AFM image of the surface relief of a corrugated MBE structure. (b) Model image of the long-range scattering potential.

of this surface shows that the average period of the relief in the [110] direction is approximately three times smaller than the period in the perpendicular direction. Since, in the structures under study, the long-range scattering potential correlates with the relief of the growth surfaces, this potential should have similar average periods in mutually perpendicular directions. The model image of the long-range potential is shown in Fig. 3b.

Figure 4a presents the behavior of the relative MR for the curves shown in Fig. 2a. One can see that the relative magnitude of the negative MR increases with increasing concentration, and the dependences on magnetic field intersect at a single point corresponding to a particular critical field value $B = B_c$. We interpret the magnetic field B_c as the point corresponding to the beginning of the classical localization by the long-range potential in increasing magnetic field [12]: $B_c = (mE_F/e^2a^2)^{1/2}(V_0/E_F)^{2/3}$, where a is the correlation length of the long-range scattering potential. The estimate of V_0 by this formula yields a value of about 10–15 meV, which agrees well with the value of V_0 determined from the position of B_{max} [7].

We also performed a computer simulation of the motion of a classical particle in the potential shown in Fig. 3b. The initial conditions of the motion were assumed to be arbitrary. The averaging was performed

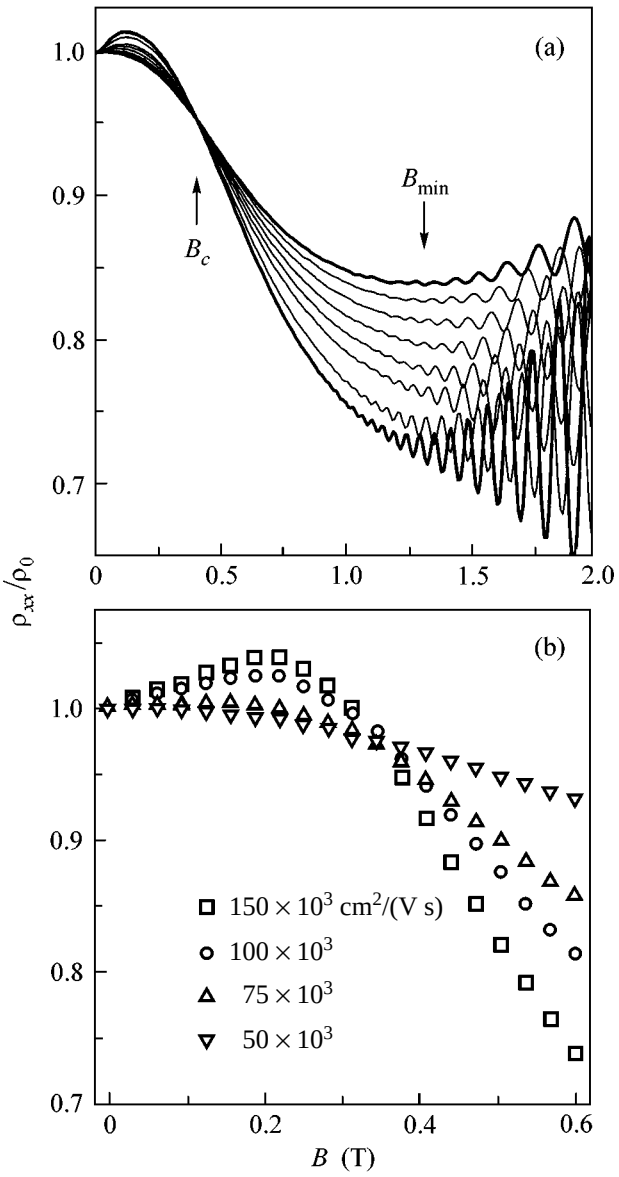


Fig. 4. (a) Relative MR of the 2D electron gas at $T = 4.2$ K for concentrations within $n_e = (1.21–1.45) \times 10^{12}$ cm^{-2} and mobilities $\mu_{xx} = (57–87) \times 10^3$ cm^2/(V s) (from bottom to top). (b) Results of simulating the relative MR of the 2D electron gas in the potential shown in Fig. 3b for $V_0 = E_F =$ 50 meV and different mobilities.

over 10^6 initial conditions. A relatively good agreement of the calculated dependences with the experimental data was achieved for the case when the amplitude of the model long-range potential was comparable to the Fermi energy of the 2D electron gas: $V_0 \sim E_F = 50$ meV. This value is much greater than that obtained from B_{max}. We explain this difference by the fact that the model potential only qualitatively describes the actual long-range scattering potential. The results of simulation, which are presented in Fig. 4b, suggest that the relative increase in the negative MR observed for the corrugated MBE structure is primarily caused by an increase in the mobility of 2D electrons rather than by a change in the concentration of the 2D electron gas.

From the experimental results presented above, it follows that the negative MR that occurs in the classical magnetic field range is most pronounced in MBE structures with corrugated growth surfaces, i.e., in the structures in which the scattering of charge carriers is mainly caused by the short-range potential of the doping impurity and the long-range potential associated with the nonplanar configuration of heterointerfaces [7]. The comparison with the theory [3, 11] shows that this experimental observation agrees well with the theoretical prediction [3]: the scattering of a 2D electron gas by both short-range and long-range potentials leads to a semiclassical negative MR.

Thus, on the basis of the comparison of our experimental data with the theory and with the results of the numerical simulation, we showed that the negative MR observed for narrow GaAs quantum wells with corrugated heterointerfaces is semiclassical and is caused by the scattering from the short-range and long-range potentials.

We are grateful to M.V. Éntin and A.G. Pogosov for fruitful discussions. This work was supported by the Russian Foundation for Basic Research, project no. 01-02-16892.

REFERENCES

1. É. M. Baskin, L. I. Magarill, and M. V. Éntin, Zh. Éksp. Teor. Fiz. **75**, 723 (1978) [Sov. Phys. JETP **48**, 365 (1978)].
2. D. V. Khveshchenko, Phys. Rev. Lett. **77**, 1817 (1996).
3. A. D. Mirlin, D. G. Polyakov, F. Evers, and P. Wolfle, Phys. Rev. Lett. **87**, 126805 (2001).
4. A. D. Mirlin, J. Wilke, F. Evers, *et al.*, Phys. Rev. Lett. **83**, 2801 (1999).
5. A. A. Bykov, G. M. Gusev, J. R. Leite, *et al.*, Phys. Rev. B **65**, 035302 (2001).
6. A. A. Bykov, A. K. Bakarov, A. V. Goran, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 182 (2001) [JETP Lett. **74**, 164 (2001)].
7. A. K. Bakarov, A. A. Bykov, N. D. Aksenova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 794 (2003) [JETP Lett. **77**, 662 (2003)].
8. K.-J. Friedland, R. Hey, H. Kostial, *et al.*, Phys. Rev. Lett. **77**, 4616 (1996).
9. A. Dmitriev, M. Dyakonov, and R. Jullien, Phys. Rev. Lett. **89**, 266804 (2002).
10. D. G. Polyakov, F. Evers, A. D. Mirlin, and P. Wolfle, Phys. Rev. B **64**, 205306 (2001).
11. E. M. Baskin and M. V. Entin, Physica B (Amsterdam) **249–251**, 805 (1998).
12. M. M. Fogler, A. Yu. Dobin, V. I. Perel, and B. I. Shklovskii, Phys. Rev. B **56**, 6823 (1997).

Translated by E. Golyamina

JETP Letters, Vol. 78, No. 3, 2003, pp. 138–142. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 170–174.

Theory of Magnetic Field-Induced Charge-Density-Wave Phases¶

A. G. Lebed

Department of Physics, Boston College, Chestnut Hill, MA 02467, USA

Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117334 Russia

Received June 26, 2003

We elaborate an analytical theory of a cascade of magnetic field-induced charge-density-wave (FICDW) phases. It is shown that the following features distinguish it from the well-known spin-density-wave cascade: (1) the FICDW phases exist at temperatures much lower than the characteristic CDW transition temperature at $H = 0$; (2) the cascade of the FICDW phases dramatically changes at certain directions of a magnetic field due to an interplay of Zeeman spin-splitting and electron motion along open Fermi surfaces. Theoretical results are compared with the recent experimental attempts to reveal FICDW phases in the organic conductors α-$(ET)_2MHg(SCN)_4$ (M = K, Tl, Rb, etc.). © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 71.45.Lr; 72.15.Nj; 73.20.Mf

A cascade of magnetic field-induced spin-density-wave (FISDW) phases has been intensively studied since its experimental discovery in the $(TMTSF)_2X$ and $(DMETTSeF)_2X$ organic compounds [1, 2] (for the first theories of this phenomenon, see [1–4]; for the recent theoretical analysis, see [5]; for the recent experiments, see [6]). The possibility that a similar phenomenon, a cascade of magnetic field-induced charge-density-wave (FICDW) phases, may exist in some solids was discussed in [3] and was numerically proved in [7]. Indeed, if a charge-density-wave (CDW) state is destroyed by pressure in a quasi-one-dimensional (Q1D) conductor, it can be restored in a magnetic field, which increases 1D Peierls instability for both SDW and CDW phases due to the "one-dimensionalization" of the electron's motion, as shown in [3–5, 7]. Nevertheless, the existence of the FICDW phases in real materials has not been firmly established, and an analytical theory of FICDW states has not been elaborated so far. The major problem in observing the FICDW phenomenon is sufficiently high CDW transition temperatures, $T_{CDW} \simeq 100$ K, in traditional Q1D conductors. Due to high T_{CDW}, the CDW states normally demonstrate very low responses to the experimental magnetic fields, $H \simeq 10$–30 T, and experimental pressures, $P \leq 10$ kbar.

Very recently, the first experimental indications [8–10] that the FICDW phases perhaps exist in layered organic conductors α-$(ET)_2MHg(SCN)_4$ (M = K, Tl, Rb, etc.) with low enough DW transition temperatures at $H = 0$ and $P = 1$ bar, $T_{CDW} \simeq 8$–10 K, have appeared. To be more specific, the unexpected changes in slope of the low temperature magnetoresistance at $T \ll T_{CDW}$ accompanied by hysteresis [8] (observed under pressure $P \geq 3$ kbar, which destroys the CDW state at $H = 0$) were interpreted [8] in terms of the FICDW phase transitions. Moreover, the recent low-temperature measurements in a tilted magnetic field [9, 10] in α-$(ET)_2MHg(SCN)_4$ conductors were suggested [9] to reflect an interplay of Zeeman spin-splitting and the orbital effects of the electron's motion along open Fermi surfaces (FS). Note that the most popular description of the ground states in α-$(ET)_2MHg(SCN)_4$ compounds at $T \simeq 8$–10 K is based on an idea [11] about a formation of some density-wave (DW) phase. Although the physical nature of this DW phase is still controversial (see, for example, discussions in [7–10, 12–14]), the experimental confirmation [8] of the theory [7] at low enough magnetic fields and the interpretation of the experiment [9] seem to be strong arguments in favor of the CDW scenario.

Our goals are (1) to elaborate an analytical theory of a cascade of FICDW phase transitions in a magnetic field perpendicular to the conducting layers (i.e., [**x**(**a**), **z**(**c**)]-plane) and (2) to suggest a theory of the above mentioned cascade at some "commensurate directions" of a magnetic field. (We note that the possibility that some special directions of a magnetic field may exist due to interplay of the Zeeman spin-splitting and the orbital electron motion along open orbits was discussed in [7, 15]. Nevertheless, no theoretical description of this phenomenon in FICDW phases has yet been proposed.) Below, we reveal some peculiar features of the FICDW phase diagram that distinguish it from the FISDW one [1–5] and present an additional argument in favor of the CDW–FICDW scenario for the ground states in α-$(ET)_2MHg(SCN)_4$ compounds.

¶ This article was submitted by the author in English.

Let us consider the FICDW phases in a tilted magnetic field perpendicular to the conducting chains [i.e., **x**(**a**) axis],

$$\mathbf{H} = (0, \cos\theta, \sin\theta)H, \quad \mathbf{A} = (0, \sin\theta, -\cos\theta)Hx, \quad (1)$$

in a layered conductor with a Q1D FS [1–4, 7, 8, 11, 16]:

$$\epsilon^{\pm}(\mathbf{p}) = \pm v_F(p_x \mp p_F) - t_\perp(p_y, p_z),$$
$$t_\perp(p_y, p_z) = 2t_c\cos(p_z c^*) \quad (2)$$
$$+ 2t_c'\cos(2p_z c^*) + 2t_b\cos(p_y b^*),$$

where θ is the angle between the field direction and the normal to the conducting plane, +(–) stands for the right (left) sheet of the FS; v_F and p_F are the Fermi velocity and Fermi momentum; $t_c \gg t_b \sim t_c'$ are the overlapping integrals of the electron wave functions [1–5, 7, 11].

As in [1–5, 7, 8], we suppose that, under external pressure, the "antinesting term" t_c' is bigger than its critical value, $t_c' > t_c^*$, which corresponds to destruction of the CDW phase [1–4, 7, 8]. Let us find the metal–FICDW phase transition temperature, $T_{\text{FICDW}}(H)$, where the FICDW phases exist at high enough magnetic fields and low enough temperatures. They correspond to the following order parameters:

$$\Delta_{FICDW}(\mathbf{r}) = \Delta\exp(i\mathbf{Q}\mathbf{r}),$$
$$\mathbf{Q} = (2p_F + K, \pi/b^*, \pi/c^*). \quad (3)$$

Using the Green's functions method [1, 3–5, 7, 17], we find that the Green's functions in the mixed representation [3, 4], $G_\sigma^{\pm}(i\omega_m; p_y, p_z; x, x_1)$, where

$$G_\sigma^{\pm}(i\omega_m; p_y, p_z; x, x_1)$$
$$= \exp[\pm ip_F(x - x_1)]g_\sigma^{\pm}(i\omega_m; p_y, p_z; x, x_1), \quad (4)$$

obey the following equations:

$$\left[i\omega_m \pm iv_F\frac{d}{dx}\right.$$
$$\left. - t_\perp\left(p_y - \frac{eH\sin\theta x}{c}, p_z + \frac{eH\cos\theta x}{c}\right) - \mu_B H\sigma\right] \quad (5)$$
$$\times g_\sigma^{\pm}(i\omega_m; p_y, p_z; x, x_1) = \delta(x - x_1).$$

Here, $i\omega_m$ is the so-called Matsubara frequency [17], σ = +(–) stands for electron with spin up (down); μ_B is the Bohr magneton, c is the speed of light; and Planck's constant $h = 1$.

Note that the main difference between Eqs. (4), (5) and the equations determining the Green's functions in the FISDW case [1, 3–5] is the appearance of the Zeeman spin-splitting term, $\mu_B H$, in Eq. (5). Let us solve Eqs. (4), (5) using the procedure described in [3, 4]. Then, let us use the Gor'kov-type equations [1, 3–5, 17] to determine the metal–FICDW phase transition temperature, $T_{\text{FICDW}}(H)$. As a result, we find that, in contrast to the well-known case of the metal–FISDW phase transitions [1, 3–5], the equation

$$\frac{1}{g} = \int_d^\infty \frac{2\pi T_{\text{FICDW}}dZ}{v_F\sinh\left(\frac{2\pi T_{\text{FICDW}}Z}{v_F}\right)}$$
$$\times J_0\left[\frac{4t_c'}{\omega_c\cos\theta}\sin\left(\frac{\omega_c Z\cos\theta}{v_F}\right)\right]\cos\left(\frac{2\mu_B HZ}{v_F}\right)\cos(KZ) \quad (6)$$

determining the metal–FICDW phase transition temperature, $T_{\text{FICDW}}(H)$, contains the Zeeman spin-splitting term, $2\mu_B H$. (Here, g is a bare coupling constant for the CDW instability, $J_0[\ldots]$ is the Bessel function, $\omega_c(H) = eHc^*v_F/c$ is the frequency of electron motion along the open sheets of the FS (2).)

To describe a cascade of the FICDW phases analytically, we suppose that the magnetic field is strong enough to satisfy the condition of the so-called "quantum limit," $\pi T_{\text{FICDW}}(H) \ll \omega_c(H)$ (for discussions, see [5]). In this case, we find with a logarithmic accuracy the following expression for the metal–FICDW phases transition temperature from Eq. (6):

$$T_{\text{FICDW}}(H) \simeq \omega_c(H)\exp\left[-\frac{1}{g_{\text{eff}}(H)}\right]; \quad (7)$$
$$\omega_c(H) = ev_F Hc^*/c,$$
$$g_{\text{eff}}(H) = \frac{1}{2\ln(t_c'/t_c^*)}$$
$$\times MAX\left(\left\langle J_0\left[\frac{4t_c'}{\omega_c(H)\cos\theta}\sin\left(\frac{\omega_c(H)Z\cos\theta}{v_F}\right)\right]\right.\right.$$
$$\times\left(\cos\left[\frac{(Kv_F - 2\mu_B H)Z}{v_F}\right]\right. \quad (8)$$
$$\left.\left.\left. + \cos\left[\frac{(Kv_F + 2\mu_B H)Z}{v_F}\right]\right)\right\rangle_Z\right)_K,$$

where t_c^* is a critical value of the "antinesting term" t_c', which destroys the CDW phase at $H = 0$; $\langle\ldots\rangle_Z$ means averaging over the coordinate Z, $MAX(\ldots)_K$ stands for the maximum value of a function with respect to the variable K.

As follows from Eq. (8), the FICDW effective interaction constant $g_{\text{eff}}(H)$ is nonzero only for the FICDW

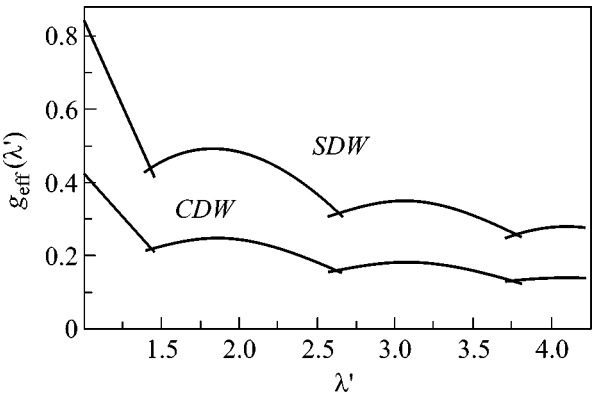


Fig. 1. Upper curve: phase transitions between the FISDW phases [3–5] with $K = 2N\omega_c(H)/v_F$; lower curve: the phase transitions between the FICDW phases with $K = 2\mu_B H/v_F + 2N\omega_c(H)/v_F$ [see Eqs. (3), (12)] calculated in the paper for $\theta = 0$ [see Eqs. (7), (11)], where N is an integer. The following notations and values of the parameters are used: $\lambda' = 2t_c'/\omega_c(H) \sim 1/H$; $\ln(t_c'/t_c^*) = 1.4$. Note that the FICDW phase transition temperatures are much lower than the FISDW phase ones.

phases (3) with the "quantized" longitudinal component of the wave vector,

$$K_1(H, L_1) = +\frac{2\mu_B H}{v_F} + 2L_1\frac{\omega_c(H)\cos\theta}{v_F}, \qquad (9)$$

$$K_2(H, L_2) = -\frac{2\mu_B H}{v_F} + 2L_2\frac{\omega_c(H)\cos\theta}{v_F}, \qquad (10)$$

where L_1 and L_2 are integers. It is possible to rewrite Eq. (8) as follows:

$$g_{\rm eff}(H) = \frac{1}{2\ln(t_c'/t_c^*)} J_N^2\left[\frac{2t_c'}{\omega_c(H)\cos\theta}\right],$$
$$J_N^2\left[\frac{2t_c'}{\omega_c(H)\cos\theta}\right] = MAX\left(J_L^2\left[\frac{2t_c'}{\omega_c(H)\cos\theta}\right]\right)_L, \qquad (11)$$

where the effective FICDW coupling constant (11) defines the metal–FICDW phase transition line for the FICDW order parameters (3) with

$$K_{1,2,3,4}(H) = \pm\frac{2\mu_B H}{v_F} \pm 2N\frac{\omega_c(H)\cos\theta}{v_F}, \qquad (12)$$

where N and L are integers. Note that the order parameters (3) with four possible wave-vectors (12) correspond to the same value of the metal–FISDW transition temperature (see Eqs. (7), (11)). In other words, the layered Q1D metal (2) is unstable in a magnetic field (1) with respect to the formation of four FICDW phases with the "quantized" wave vectors (12). Our theoretical results (7)–(12) are summarized in Fig. 1, where the cascade of the FICDW phase transitions (calculated for $\theta = 0$) is compared with the FISDW cascade [1, 2, 4, 5, 18–20].

In contrast to the FISDW case [1–5, 18–20], the quantization rules (12) for the FICDW wave vector (3) contain the Zeeman spin-splitting term. This leads to significant differences between physical properties of the FICDW and FISDW phases. Indeed, as follows from Eq. (11), the effective interaction constant for the formation of the FICDW phases is two times smaller than the corresponding constant for the formation of the FISDW phases [1–4, 18–20] (see Fig. 1). Therefore, unlike the FISDW phases, the FICDW ones can appear only at temperatures much lower than the CDW characteristic temperature at $H = 0$ and $P = 1$ bar:

$$T_{\rm FICDW}(H) \ll T_{\rm CDW}(H = 0, P = 1\ \text{bar}). \qquad (13)$$

Note that this result is in agreement with the experiment [8] and, thus, supports the hypothesis [7–12] about the CDW ground states in α-$(ET)_2MHg(SCN)_4$ compound at $H = 0$ and $P = 1$ bar.

As follows from Eqs. (7)–(10), the second distinctive feature of the FICDW phases is that the effective FICDW coupling constant changes at "commensurate directions" of a magnetic field,

$$\cos\theta = (1/M)[2\mu_B H/\omega_c(H)], \qquad (14)$$

where $M \neq 0$ is an integer.

Starting from Eq. (8), it is possible to show that the effective coupling constant for the "commensurate directions" of a magnetic field (14) is equal to

$$g_{\rm eff}^M(H) = \frac{1}{2\ln(t_c'/t_c^*)}$$
$$\times MAX\left(J_L^2\left[\frac{2t_c'}{\omega_c(H)\cos\theta}\right] + J_{L+M}^2\left[\frac{2t_c'}{\omega_c(H)\cos\theta}\right]\right)_L, \qquad (15)$$
$$T_{\rm FICDW}(H) = \exp\left[-\frac{1}{g_{\rm eff}^M(H)}\right],$$

where L and N are integers. Thus, the cascades of the FICDW phase transitions at the "commensurate directions" of a magnetic field (14) are qualitatively different from the cascade at $\theta = 0$ (see Fig. 2). As can be seen from Eq. (15) and Fig. 2, the FICDW transition temperatures at "commensurate directions" of a magnetic field can be significantly higher than the FICDW transition temperatures for the standard experimental geometry, where magnetic field is perpendicular to the conducting layers (i.e., at $\theta = 0$). This provides a method to detect the FICDW phases in other Q1D compounds. We speculate that the experimental results [9], where some novel phases were observed only at angles higher than $\theta \geq \pi/4$, may be related to the main "commensurate angles" (14) (see Fig. 2).

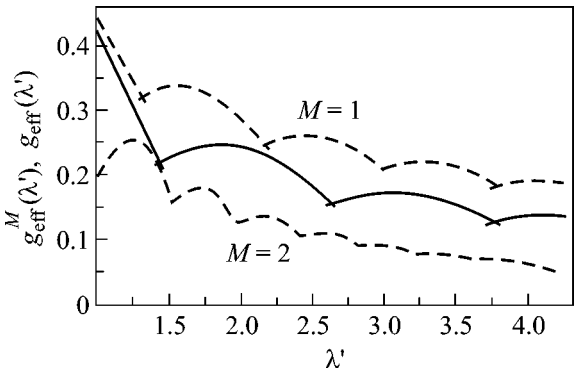


Fig. 2. Phase transitions between the metallic and the FICDW phases are calculated for $\theta = 0$ (solid line) [see Eqs. (3), (7), (11), (12)] and for two "commensurate directions" of a magnetic field [see Eq. (14)]. Upper dashed line corresponds to $\theta = \pi/4$ (i.e., $M = 1$), whereas the lower dashed line corresponds to $M = 2$ [see Eqs. (14), (15)]; $\ln(t_c'/t_c^*) = 1.4$, $2\mu_B H/\omega_c(H) = 1/\sqrt{2}$; $\lambda' = 2t_c'/\omega_c(H) \sim 1/H$. Note that $T_{\rm FICDW}(H, \theta = \pi/4)$ is higher than $T_{\rm FICDW}(H, \theta = 0)$.

Below, we discuss the applicability of the analytical theory of the cascades of FICDW phase transitions suggested in the paper to real experiments. We stress that, in the paper, we have extended the so-called "quantized nesting" (QN) model [1–4, 18–20] to describe the cascade of the FICDW phase transitions. As shown in [5], the analytical QN model in the case of FISDW phase transitions is an approximation that is qualitatively correct in the quantum limit where $\omega_c(H) \geq \pi T_{\rm FICDW}(H)$. Therefore, we expect that the analytical theory suggested by us is qualitatively correct at least if $\omega_c(H) \geq \pi T_{\rm FICDW}(H)$. In contrast to the FISDW case, application of the present theory is restricted by the second condition, $\omega_c(H) \leq 2t_c'$, since Eqs. (7)–(15) are valid only with logarithmic accuracy. Using the typical values of the parameters [8], $\omega_c(H) \simeq 1$ K/T, $T_{\rm FICDW}(H) \simeq 1$ K, and $t_c \simeq 10$ K, we can conclude that the analytical theory suggested in the paper is an appropriate description of the FICDW phases in a broad region of magnetic fields, $3\ {\rm T} \leq H \leq 20$ T.

To summarize, an analytical theory of a cascade of the FICDW phase transitions in a magnetic field perpendicular to the conducting chains in layered Q1D metals (predicted by L.P. Gor'kov and the author in [3] and numerically proved in [7]) is elaborated. As a result, we come to the conclusion that, in contrast to the FISDW phases, the FICDW ones can exist only at sufficiently low temperatures (13). This is in agreement with the experimental data [8] and is an extra argument in favor of the CDW/FICDW nature of the ground states existing at low temperatures in organic conductor α-$(ET)_2MHg(SCN)_4$. We have suggested a theory of the FICDW phases in an inclined magnetic field and calculated the FICDW phase diagram for some "commensurate directions" [7, 15] of a magnetic field (see Eqs. (14), (15) and Fig. 2). It is shown that the FICDW transition temperatures can be higher for some "commensurate directions" of a magnetic field than for the field perpendicular to the conducting plane (see Eqs. (11), (14), (15) and Fig. 2). This provides an experimental method to detect the FICDW phases in other Q1D compounds and may be related to the experiment [9], where some novel phases were observed only in inclined magnetic fields.

I am thankful to M.V. Kartsovnik for sending me the preprint of [8] before its publication. I am thankful to N.N. Bagmet, W. Biberacher, E.V. Brusse, S.B. Brusse, M.V. Kartsovnik, and M.J. Naughton for the numerous discussions. I also wish to thank M.J. Naughton for his kind hospitality.

This work was supported in part by the Department of Energy, grant no. DE-FG02-02ER63404, and by INTAS, grant nos. 2001-2212 and 2001-0791.

REFERENCES

1. T. Ishiguro, K. Yamaji, and G. Saito, *Organic Superconductors*, 2nd ed. (Springer, Heidelberg, 1998).
2. For a review, see P. M. Chaikin, J. Phys. I **6**, 1875 (1996).
3. L. P. Gor'kov and A. G. Lebed, J. Phys. Lett. (Paris) **45**, L-433 (1984); P. M. Chaikin, Phys. Rev. B **31**, 4770 (1985).
4. M. Heritier, G. Montambaux, and P. Lederer, J. Phys. Lett. (Paris) **45**, L-943 (1984).
5. A. G. Lebed, Phys. Rev. Lett. **88**, 177001-1 (2002); JETP Lett. **72**, 141 (2000).
6. A. V. Kornilov, V. M. Pudalov, Y. Kitaoka, *et al.*, Phys. Rev. B **65**, 060404R (2002).
7. D. Zanchi, A. Bjelis, and G. Montambaux, Phys. Rev. B **53**, 1240 (1996); A. Bjelis, D. Zanchi, and G. Montambaux, J. Phys. IV **9**, 10 (1999).
8. D. Andres, M. V. Kartsovnik, W. Biberacher, *et al.*, Phys. Rev. B **64**, 161104-1 (2001).
9. J. S. Qualls, L. Balicas, J. S. Brooks, *et al.*, Phys. Rev. B **62**, 10008 (2000).
10. M. V. Kartsovnik, D. Andres, W. Biberacher, *et al.*, Synth. Met. **120**, 841 (2001).
11. M. V. Kartsovnik, A. E. Kovalev, and N. Kushch, J. Phys. I **3**, 1187 (1993).
12. P. Christ, W. Biberacher, M. V. Katsovnik, *et al.*, JETP Lett. **71**, 303 (2000); R. McKenzie, cond-mat/9706235 (unpublished); N. Biskup, J. A. A. J. Perenboom, J. S. Brooks, and J. S. Qualls, Solid State Commun. **107**, 503 (1998); T. Sasaki, A. G. Lebed, N. Toyota, *et al.*, Phys. Rev. B **54**, 12969 (1996).
13. For a review, see J. Singleton, Rep. Prog. Phys. **63**, 1111 (2000).
14. J. Wosnitza, *Fermi Surfaces of Low-Dimensional Organic Metals and Superconductors* (Springer, Berlin, 1996).
15. A. G. Lebed, J. Phys. I **2**, 2005 (1992).

16. In this paper, we study the generic features of the FICDW phases using, as usually, a model Q1D spectrum (2). Note that the real organic compounds α-$(ET)_2MHg(SCN)_4$, in addition to the Q1D sheets of the FS, also have Q2D parts of the FS. The influence of the Q2D electron orbits on the physical properties of the FICDW phases will be studied elsewere. As shown in [7], from a phenomenological point of view, the FICDW phase is mixed with the component of the FISDW state characterized by zero projection of the total spin on the magnetic field direction. Nevertheless, the CDW-FICDW scenario (see [8, 9, 12]) considered in this paper is based on the suggestion that the SDW instability is a secondary phenomenon comparable with the CDW one. In other words, we assume that a bare coupling constant for the CDW instability is significantly bigger than a bare coupling constant for the SDW instability (see for discussions in [8, 9, 12]).

17. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Dover, New York, 1975).

18. A. G. Lebed, Zh. Éksp. Teor. Fiz. **89**, 1034 (1985) [Sov. Phys. JETP **62**, 595 (1985)].

19. K. Maki, Phys. Rev. B **33**, 4826 (1986); L. Chen, K. Maki, and A. Virosztek, Physica B (Amsterdam) **143**, 444 (1986); A. Virosztek, L. Chen, and K. Maki, Phys. Rev. B **34**, 3371 (1986).

20. K. Yamaji, Synth. Met. **13**, 29 (1986).

JETP Letters, Vol. 78, No. 3, 2003, pp. 143–147. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 175–179.

Electronic Structure of Hexagonal Rare-Earth Manganites $RMnO_3$

A. M. Kalashnikova and R. V. Pisarev*

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

** e-mail: pisarev@mail.ioffe.ru*

Received June 30, 2003

The optical spectra of single crystals of hexagonal rare-earth manganites $RMnO_3$ (R = Sc, Y, Er) are studied in the range from 0.7 to 5.4 eV. It is found that the spectra substantially differ from the spectra of orthorhombic manganites in both the positions of spectral features and their polarization anisotropy. It is shown that the optical absorption edge is determined by an abnormally strong ($k \simeq 1$) and narrow electric dipole transition with the center at approximately 1.6 eV with light polarization in the basal plane of the crystal. This transition can be treated with confidence as charge transfer from oxygen to manganese. The experimental results are in many instances substantially different from the first-principles calculations of the electronic structure of $YMnO_3$ published recently and, hence, may serve as a reliable basis for the further improvement of computational methods.

PACS numbers: 71.20.Eh; 78.20.Ci

The character of electronic states and the nature of the fundamental absorption edge in strongly correlated compounds of 3*d* transition metals have been the subject of intensive studies for several decades. Despite this, many principal questions in this field remain open and experimental studies in some cases lead to ambiguous results and often become the subject of contradictory explanations. This is primarily due to strong electron correlation and an intimate connection between spin, charge, and orbital degrees of freedom in the compounds of 3*d* metals. This complicates the construction of rigorous theoretical models, whereas the same experimental result can be explained within the framework of different models. The most striking examples of strong correlations are the high-temperature conductivity in copper oxide compounds [1], colossal magnetoresistance, and a number of other phenomena in orthorhombic manganites with the perovskite structure [2]. Neither the former nor the latter phenomenon has received a strict and unambiguous explanation so far, and studies in these fields are being actively continued.

In this work, we report investigations into optical properties of hexagonal dielectric manganites $RMnO_3$ (R = Sc, Y, and Er), which can be regarded as an antipode of orthorhombic manganites. In terms of their chemical composition, undoped hexagonal and orthorhombic manganites are identical. However, differences in crystal structures and in the symmetry of the local environment of magnetic manganese and rare-earth ions lead to radical differences between the physical properties of these two groups of materials. Using the ellipsometry technique in the spectral range of 0.7–5.4 eV, we found that the electronic structures of hexagonal and orthorhombic manganites are substantially different from each other. We will show that the results of first-principles calculations of the electronic structure of $YMnO_3$ published recently by two groups [3, 4] are in many instances substantially different from our experimental results.

Hexagonal manganites $RMnO_3$ (R = Sc, Y, In, Er, … Lu) form a group of compounds with an unusual combination of electrical, magnetic, and optical properties. The results of earlier studies of these materials can be found in the reference literature [5], and many new publications have appeared in recent years (see, for example, [6] and references therein). Hexagonal manganites at room temperature are crystallized in the noncentrosymmetric 6*mm* point group and represent ferroelectrics, whereas orthorhombic manganites with the perovskite structure possess the *mmm* symmetry, that is, are centrosymmetric. It can be expected that the electronic structures of these two groups must be substantially different, because Mn^{3+} ions in hexagonal manganites reside in a bipyramidal environment of five O^{2-} oxygen ions, whereas those in orthorhombic manganites are located in a distorted octahedral environment of six O^{2-} ions. In particular, the difference in the local symmetry of Mn^{3+} ions results in the fact that the Jahn–Teller effect is exhibited in orthorhombic manganites, whereas this effect is absent in hexagonal manganites.

Only fragmentary reports of investigations into the optical properties of hexagonal manganites are available in the literature. The optical absorption spectra of thin polycrystalline $YMnO_3$ films of hexagonal and

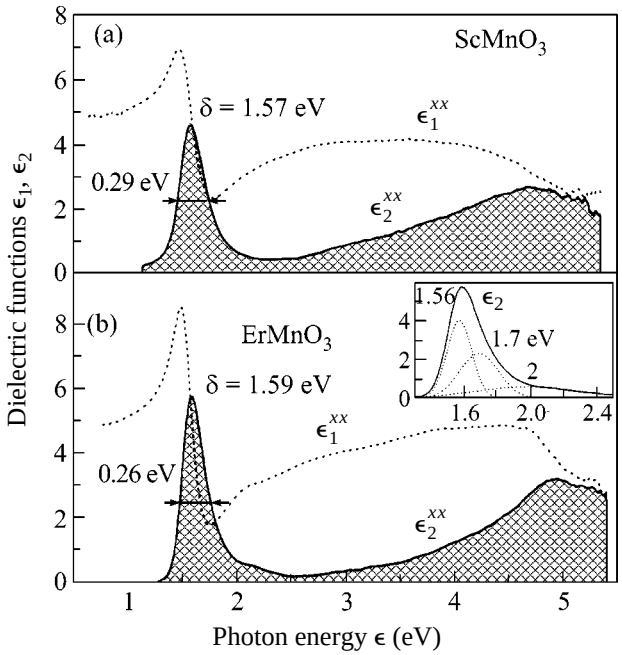


Fig. 1. Spectral dependences of the real and imaginary parts (dotted and solid lines, respectively) of the dielectric function ϵ^{xx} (a) for $ScMnO_3$ and (b) for $ErMnO_3$. The inset shows the result of a decomposition of the absorption peak in the region of 1.6 eV (solid line) with the use of Gaussian functions (dotted lines) for $ErMnO_3$.

orthorhombic modifications were investigated in [7]. It was inferred that the fundamental absorption edge of both materials lies at about 4.2–4.3 eV. The luminescence spectra of a hexagonal $YMnO_3$ single crystal and the unpolarized optical absorption spectra of a solution of $YMnO_3$ and KBr powders were published recently in [8]. The results of studies of second-harmonic generation in hexagonal manganites in the region of local 3*d* transitions were published in several works [9, 10]. The third-order nonlinear optical response in the energy range 1.45–1.62 eV was investigated in [11].

The measurements of the ellipsometric parameters ψ and Δ were carried out at temperature $T = 295$ K in the spectral range 0.7–5.4 eV. We calculated the optical parameters of the materials using the relationship [12]

$$\epsilon^{ps} = \sin^2\theta\left[1 + \tan^2\theta\left(\frac{1 - r_{pp}/r_{ss}}{1 + r_{pp}/r_{ss}}\right)^2\right], \qquad (1)$$

where ϵ^{ps} is the pseudodielectric function and $r_{pp}/r_{ss} = \tan\psi e^{i\Delta}$.

The samples of $ScMnO_3$ and $ErMnO_3$ represented polished plane-parallel plates with the normal along the [0001] hexagonal optical axis. In this case, the pseudodielectric function ϵ^{ps} determines the element $\epsilon^{xx} = \epsilon^{yy}$ of the dielectric tensor along the direction of intersection of the incidence plane and the sample surface [13]. The measurements were performed for incidence angles $\theta = 65°$, 70°, and 75°, and calculations gave coincident results. The $YMnO_3$ sample was a plate with the normal along the *y* axis, that is, with the optical axis in the [0001] plane of the plate. In this case, the investigations were carried out for two orientations of the sample with the optical axis parallel and perpendicular to the incidence plane. These investigations allowed the values of ϵ^{xx} and ϵ^{zz} to be calculated with the use of the ellipsometric relationships and the iteration method proposed for anisotropic crystals in [14]. The accuracy of measurements was mainly limited by the small sizes of the samples, which were of about 2–4 mm^2.

The spectra of the real ϵ_1^{xx} and the imaginary ϵ_2^{xx} components of the dielectric tensor for scandium and erbium manganites are presented in Fig. 1. Below 1.2–1.3 eV, the absorption drops to zero, which is in agreement with the direct transmission measurements of the absorption spectrum. An abnormally intense and narrow absorption peak in the region of 1.57–1.59 eV is the most pronounced feature of the spectra. This peak exhibits asymmetry, and its decomposition into components with the use of Gaussian functions demonstrates the occurrence of at least two components, as is shown in the inset in Fig. 1. An analogous decomposition procedure with the use of Lorentzian functions gives a significantly poorer description of the spectra. The absorption in the region above 2.2–2.4 eV increases with increasing photon energy, and a broad absorption peak is observed in the energy region of 4.6–4.8 eV.

The spectral dependences of $\epsilon_{1,2}^{xx}$ and $\epsilon_{1,2}^{zz}$ for the $YMnO_3$ crystal are shown in Fig. 2. The spectra of ϵ_1^{xx} and ϵ_2^{xx} are similar to the analogous spectra in $ScMnO_3$ and $ErMnO_3$. The agreement between the spectra demonstrates the reliability of the results obtained, because different crystallographic planes were used in the ellipsometric measurements and different analytical relationships were used for dielectric functions in the calculations for $YMnO_3$ on the one hand and for $ScMnO_3$ and $ErMnO_3$ on the other. The spectra of $\epsilon_{1,2}^{xx}$ and $\epsilon_{1,2}^{zz}$ in $YMnO_3$ strongly differ from each other; that is, the crystal is characterized by strong optical anisotropy in the entire studied region. The change in the spectrum in the region of the narrow absorption band at 1.6 eV and also the appearance of a wide absorption band at 3.5 eV should be specially noted. The spectral dependences of linear birefringence $\Delta n = n_x - n_z$ and linear dichroism $\Delta k = k_x - k_z$, which characterize the optical anisotropy of $YMnO_3$, are displayed in Fig. 3.

Let us compare the results obtained for the dielectric functions of hexagonal manganites with the available experimental data. Our results showed that manganites in the visible spectral region (1.8–3.4 eV) are character-

ized by high values of the absorption coefficient and are actually opaque. At the maximum of the band at 1.6 eV, $\alpha = 1.6 \times 10^5$ cm^{-1} in $ScMnO_3$ and 2×10^5 cm^{-1} in $ErMnO_3$. In one of the first works on hexagonal manganites [15], thin single-crystal $LuMnO_3$ plates were studied in a polarization microscope. The plate thickness was not indicated in this work; however, based on our data, we believe that the plates were several micrometers thick. This study showed that $LuMnO_3$ possesses strong linear dichroism in the visible spectral region. In particular, the plates turned dark red in the light polarized along the optical axis and light green in the case of polarization perpendicular to the optical axis. This qualitative observation is in agreement with the spectral dependence of linear dichroism (Fig. 3), which exhibits a change in the dichroism sign between the red and green spectral regions.

Our results demonstrate a significant quantitative and even qualitative disagreement with the absorption spectrum of a thin polycrystalline $YMnO_3$ film [7]. We reason that these discrepancies are associated with the unreliability of the measurement of absorption in thin polycrystalline films over a wide spectral range, when both the quality of the material itself and the difficulties of adequately taking into account the factors of light reflection, scattering, and interference may result in a strong divergence between the true and the measured spectra. As an additional argument, we indicate that the most intense absorption band at 1.6 eV went virtually unnoticed in [7] and significant differences between the $YMnO_3$ films with the hexagonal and orthorhombic structures were also not detected. The absorption curve of a solution of $YMnO_3$ and KBr powders given in the work [8] is also rather far from our results.

The results of first-principles electronic structure calculations for the hexagonal modification of $YMnO_3$ were published recently [3, 4].The LSDA+U approximation was used in [3], and it was shown that $YMnO_3$ must be a semiconductor with a band gap of about 1.5 eV. The top of the valence band is mainly determined by the 2*p* states of the oxygen ions in the basal plane, and the lowest unoccupied band is formed by the $(3z^2–r^2)$ states of the Mn^{3+} ions. These calculations showed that the electronic spectrum of $YMnO_3$ must contain a higher unoccupied band with the lower edge at about 4.3 eV; that is, there is a second forbidden band between the two manganese subbands. This conclusion qualitatively agrees with our results, and the narrow absorption band at about 1.6 eV can be assigned to transitions with charge transfer from the planar oxygen ions to manganese with the formation of a hole in the valence band and with the occupation of the vacant $(3z^2–r^2)$ orbital with an electron. The unusual narrowness of the band can be associated with polaron effects [16], which lead to band narrowing for small-radius polarons. Another first-principles approach based on the LDA+U model was used in the work [4]. The results

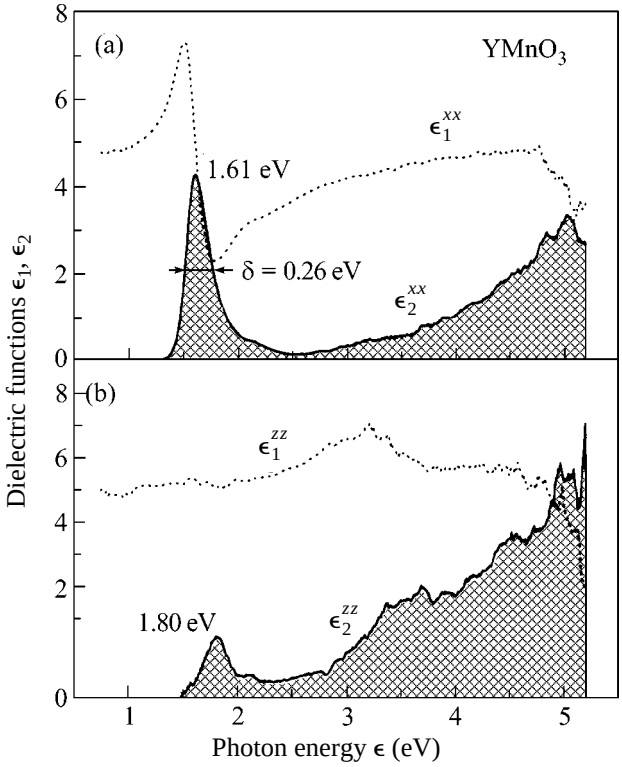


Fig. 2. Spectral dependences of the real and imaginary parts (dotted and solid lines, respectively) of the dielectric functions (a) ϵ^{xx} and (b) ϵ^{zz} for $YMnO_3$.

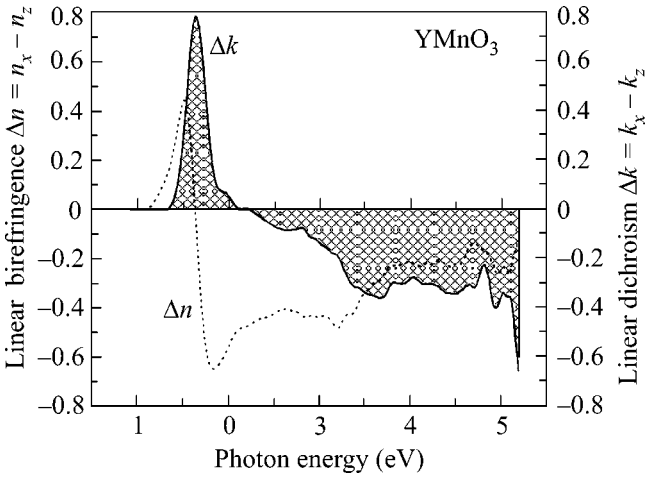


Fig. 3. Spectral dependences of linear dispersion (dotted line) and linear dichroism (solid line) for $YMnO_3$.

of calculations by the two groups [3, 4] markedly differ, though some details of the spectra of electron density of states are in qualitative agreement. In particular, the band-gap width at point Γ of the Brillouin zone was determined in [4] at 0.48 eV, which is significantly

smaller than the value found in the work [3]. We see that the results of the works [3, 4] can only qualitatively be compared with the experimental data, but, nevertheless, these results allow the conclusion that the strong absorption band that we observed at about 1.6 eV can be assigned to a transition with charge transfer from oxygen to manganese.

The spectra of dielectric functions ϵ_1 and ϵ_2 were calculated in the work [4]. In this case, we also may speak of only some qualitative agreement between the calculated and experimental results. The calculations show that the optical absorption edge for ϵ_2^{xx} is located at about 0.5 eV and that for ϵ_2^{zz} is located at 1.6 eV; that is, very strong anisotropy is observed in optical susceptibility. The results presented in Fig. 2 actually indicate that the optical susceptibility of $YMnO_3$ is characterized by strong anisotropy; however, the numerical values of ϵ_1 and ϵ_2 and their spectral dependences strongly differ from the calculated results. The ϵ_2^{xx} susceptibility predominates in the region of the band at 1.6 eV; however, the contribution of ϵ_2^{zz} becomes stronger above 2.4 eV (see Fig. 3) with a wide band at about 3.5 eV.

Let us dwell on another interesting feature of the electronic structure of hexagonal manganites that stems from the comparison of the results for the spectra of ϵ_1 and ϵ_2 and the studies of second-harmonic generation [9, 10]. The spectra of ϵ_1 and ϵ_2 in Figs. 1 and 2 are undoubtedly determined by intense electric dipole transitions between the oxygen valence band and the manganese subbands. Clearly defined features are observed in the second-harmonic spectra in the energy regions of 1.4–1.8 eV and 2.4–2.8 eV. These features with respect to their selection rules can be explained adequately within the model of local transitions between the 3*d* states of Mn^{3+} ions in the crystal field of the five oxygen ions that form the nearest environment of manganese ions [10, 17]. The transitions between the 3*d* states are forbidden within the electric dipole approximation, and their intensity in the oxides of 3*d* transition metals is two or three orders of magnitude weaker than the intensity of allowed electric dipole transitions [18]. No pronounced feature is observed in the spectra of ϵ_2 in the region of 2.4–2.8 eV, where the features of the second-harmonic spectra are observed. This can be readily understood taking into account the differences by two or three orders of magnitude in the intensities of allowed and forbidden transitions.

Thus, as a result of the optical study of three hexagonal manganites $RMnO_3$, we obtained spectral dependences for the dielectric functions ϵ_1 and ϵ_2 in the spectral range from 0.7 to 5.4 eV. The reliability of the results obtained is corroborated by the close similarity of the spectra of ϵ_1 and ϵ_2 in crystals of different compositions, the perfect identity of spectra calculated from measurements at different incident angles, and the similarity of the spectra of ϵ^{xx} in $ScMnO_3$ and $ErMnO_3$ to the spectra of ϵ^{xx} in $YMnO_3$. These latter were obtained with samples differing in their orientation and with the use of another procedure for the calculation of spectra based on the ellipsometric values Δ and ψ. All this testifies that the published data on absorption in thin $YMnO_3$ films and in solutions of $YMnO_3$ [7] and KBr powders [8] give no way of deducing correct conclusions on the electronic structure of hexagonal manganites. The detection of an intense electric dipole absorption band with the center at about 1.6 eV, which can be interpreted as a band due to charge transfer from the oxygen band to the lower manganese subband, is an important result of our work. The upper manganese subband is located at a higher energy with the center at about 4.6–4.8 eV. Thus, the electronic structure of hexagonal manganites substantially differs in its character from the structure of orthorhombic manganites, in which charge-transfer transitions start in the region at about 4 eV [19]. It should also be noted that our results in many important instances substantially differ from the calculations published by two theoretical groups [3, 4].

The authors are grateful to K. Kohn for presenting some samples for investigations and to Th. Rasing for the possibility of performing ellipsometric measurements.

This work was partially supported by the Russian Foundation for Basic Research and the European program “Dynamics.”

REFERENCES

1. M. A. Kastner, R. J. Birgeneau, G. Shirane, and Y. Endoh, Rev. Mod. Phys. **70**, 897 (1998).
2. Y. Tokura and Y. Tomioka, J. Magn. Magn. Mater. **200**, 1 (1999).
3. J. E. Medvedeva, V. I. Anisimov, M. A. Korotin, *et al.*, J. Phys.: Condens. Matter **12**, 4947 (2000).
4. M. Qian, J. Dong, and D. Y. Xing, Phys. Rev. B **63**, 155101 (2001).
5. *Landolt-Börnstein. Numerical Data and Functional Relationships* (Springer, Berlin, 1981), Group III, Vol. 16a.
6. A. V. Goltsev, R. V. Pisarev, Th. Lottermoser, and M. Fiebig, Phys. Rev. Lett. **90**, 177204 (2003); T. Katsufuji, M. Masaki, A. Machida, *et al.*, Phys. Rev. B **66**, 134434 (2002); M. Fiebig, Th. Lottermoser, D. Fröhlich, *et al.*, Nature **419**, 818 (2002); N. A. Hill and A. Filippetto, J. Magn. Magn. Mater. **242**, 976 (2002); H. Sugie, N. Twata, and K. Kohn, J. Phys. Soc. Jpn. **71**, 1558 (2002); B. B. van Aken, J.-W. G. Bos, R. A. de Groot, and T. T. M. Palstra, Phys. Rev. B **63**, 125127 (2001); D. Tomuta, S. Ramakrishnan, G. J. Nieuwenhuys, and J. A. Mydosh, J. Phys.: Condens. Matter **13**, 4543 (2001).
7. Woo-Chul Yi, Sook-Il Kwun, and Jong-Gul Yoon, J. Phys. Soc. Jpn. **69**, 2706 (2000).

8. J. Takahashi, K. Kohn, and E. Hanamura, J. Lumin. **100**, 141 (2002).

9. M. Fiebig, D. Fröhlich, K. Kohn, *et al.*, Phys. Rev. Lett. **84**, 5620 (2000).

10. C. Degenhardt, M. Fiebig, D. Fröhlich, *et al.*, Appl. Phys. B **73**, 139 (2001).

11. A. V. Kimel, R. V. Pisarev, F. Bentivegna, and Th. Rasing, Phys. Rev. B **64**, 201103 (2001).

12. R. V. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977; Mir, Moscow, 1981).

13. D. E. Aspnes, J. Opt. Soc. Am. **70**, 1275 (1980).

14. A. G. Pakhomov, V. A. Yakovlev, and A. F. Konstantinova, Zh. Tekh. Fiz. **51**, 1013 (1981) [Sov. Phys. Tech. Phys. **26**, 607 (1981)].

15. H. L. Yakel, W. C. Koehler, E. F. Bertaut, and E. F. Forrat, Acta Crystallogr. **16**, 957 (1963).

16. *Polarons*, Ed. by Yu. A. Firsov (Nauka, Moscow, 1975).

17. T. Iizuka-Sakano, E. Hanamura, and Y. Tanabe, J. Phys.: Condens. Matter **13**, 3031 (2001).

18. A. B. P. Lever, *Inorganic Electronic Spectroscopy*, 2nd ed. (Elsevier, Amsterdam, 1984).

19. A. E. Bocquet, T. Mizokawa, K. Mirokawa, *et al.*, Phys. Rev. B **53**, 1161 (1996).

Translated by A. Bagatur'yants

JETP Letters, Vol. 78, No. 3, 2003, pp. 148–151. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 180–183.

Point Impurities Remove Degeneracy of the Landau Levels in a Two-Dimensional Electron Gas

A. M. Dyugaev[1, 2], P. D. Grigoriev[1, 3, *], and Yu. N. Ovchinnikov[1, 2]

[1]*Landau Institute of Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia*

[2]*Max Planck Institute for the Physics of Complex Systems, Dresden D-01187, Germany*

[3]*High Magnetic Field Laboratory, MPI–FRF and CNRS, BP166, Grenoble F-38041, France*

**e-mail: pashag@itp.ac.ru*

Received July 3, 2003

The density of states of a two-dimensional electron gas in a magnetic field has been studied taking into account the scattering on point impurities. It is demonstrated that allowance for the electron–impurity interaction completely removes degeneracy of the Landau levels even for a small volume density of these point defects. The density of states is calculated in a self-consistent approximation taking into account all diagrams without intersections of the impurity lines. The electron density of states ρ is determined by the contribution from a cut of the one-particle Green's function rather than from a pole. In a wide range of the electron energies ω (measured from each Landau level), the value of ρ(ω) is inversely proportional to the energy |ω| and proportional to the impurity concentration. The obtained results are applicable to various two-dimensional electron systems such as inversion layers, heterostructures, and electrons on the surface of liquid helium. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 73.20.At; 71.70.Di; 73.20.Hb

1. Two-dimensional (2D) electron systems exist at the interfaces between two condensed media, typical examples being electrons on the surface of liquid helium [1] and in inversion layers and heterostructures [2]. In a magnetic field, the spectrum of electron energies in such a 2D system is discrete and infinitely degenerate. It is of interest to consider the removal of degeneracy as a result of the interaction between electrons and point defects.

It was established [3] that the infinite degeneracy of the Landau levels is only partly removed in the presence of a small surface density n_s of point impurities. For $N < S/2\pi l_H^2$ (where $N = n_s S$ is the total number of impurities, S is the area accessible for electrons, and l_H is the magnetic length), the degree of degeneracy of the Landau level is $\frac{S}{2\pi l_H^2} - N$. In this case, N electron states are split from each level to form an impurity band. An elegant proof of this statement [3] was based on the possibility to determine the electron wave function on the Landau level in such a way that this function is zero at the impurity location sites. Note that the point impurities do not modify the spectrum of $S/2\pi l_H^2 - N$ electron states.

However, this approach requires a proper definition of the surface density of impurities n_s. Indeed, the 2D electron systems are usually open. The motion of electrons along the z axis perpendicular to the conducting plane is characterized by the wave function $\varphi(z)$ limited by a very large macroscopic scale z_∞ depending on the geometry of a measuring instrument. The typical form of φ as a function of z is as follows [1, 2]:

$$\begin{aligned} \varphi^2(z) &= \frac{1}{2z_0} z_*^2 e^{-z_*}, \quad 0 < z < z_\infty; \\ \varphi(z) &= 0, \quad z < 0, \quad z_* \equiv z/z_0. \end{aligned} \tag{1}$$

Since the entire region of $z > 0$ (rather than only $z \le z_0$) is accessible for an electron, the surface density of impurities $n_s = n_{imp} z_\infty$ is very large even for a small volume density n_{imp} of these impurities. Therefore, we can assume that $n_s = \infty$ and, hence, the infinite degeneracy of the Landau levels is completely, rather than partly, removed as a result of the interaction of electrons with point impurities.

Below, the electron density of states ρ is determined for a small concentration of impurities $C_0 \equiv n_{imp} 2\pi l_H^2 z_0 \ll 1$. In other words, it is assumed that the number of impurities is small in a surface layer ($z \le z_0$) but large in the entire region of $z < z_\infty$, where $C_\infty = C_0(z_\infty/z_0) \gg 1$. This assumption is usually valid in experiment [1, 2]. For electrons on the surface of liquid helium, the roles of point impurities are played by the heavy atoms of helium vapor.

2. The density of states $\rho(E)$is related to the imaginary part of the Green's function of electrons $G(r, r', E)$ by the equation [4]

$$\rho(E) = \frac{-1}{\pi}\mathrm{Im}G(E), \quad G(E) = \int d^3 r G(r, r, E).$$

Since the electron system is homogeneous in the conducting plane (x, y), the Green's function $G(r, r, E)$ is independent of the coordinates in this plane: $G(r, r, E) = G(z, E)$. The gap between electron levels in the z axis is usually large and all electrons occur on the lowest energy level and have the same wave function $\varphi(z)$, so that $G(z, E) = G(E)\varphi^2(z)$. The magnetic field H is assumed to be strong, so that the interaction $V(\mathbf{r})$ between electrons and impurities does not mix their wave functions on various Landau levels. Therefore, the electron density of states $\rho(E)$ depends only on the variable $\omega \equiv E - \varepsilon_0 - (n + 1/2)\omega_c$, where ε_0 is the energy of the lowest level in the z axis and ω_c is the cyclotron energy.

The consideration below is restricted to point impurities. The potential of interaction between an electron and such impurities is given by the expression

$$V(\mathbf{r}) = u_0 \sum_i \delta(x - x_i)\delta(y - y_i)\delta(z - z_i), \qquad (2)$$

where x, y, z are the electron coordinates and x_i, y_i, z_i are the coordinates of point impurities. The electron scattering length is

$$a = -mu_0/2\pi\hbar^2. \qquad (3)$$

The Green's function $G(\omega)$ is related to the seeding function $G_0(\omega)$ of the pure electron system by the well-known relation [4]

$$G(\omega) = \frac{1}{G_0^{-1}(\omega) - \Sigma(\omega)}, \quad G_0(\omega) = \frac{1}{\omega}, \qquad (4)$$

where the function $\Sigma(\omega)$ can be determined using the well developed cross technique [4–6]:

$$\Sigma(\omega) = u_0 n_{imp} \int_0^{z_\infty} \frac{\varphi_0^2(z)dz}{1 - \dfrac{u_0}{2\pi l_H^2}\varphi_0^2(z)G(\omega)}. \qquad (5)$$

This expression is obtained by summing over all the cross diagrams without intersections of the impurity lines [6], followed by averaging over the positions of impurities. Relations (4) and (5) are valid only for the point impurities interacting with electrons as described by the potential (2).

It is convenient to define a reduced wave function φ_* for the transverse electron motion by the following relation

$$\varphi_0^2(z) = \frac{1}{z_0}\varphi_*^2(z_*), \quad z_* \equiv \frac{z}{z_0}, \quad \int \varphi_*^2(z_*)dz_* = 1, \qquad (6)$$

where z_0 is a characteristic scale of the function $\varphi_0(z)$. Then, using Eqs. (3), (5), and (6), we obtain

$$\Sigma(\omega) = -G^{-1}(\omega)C_0 J(\omega);$$
$$J(\omega) \equiv \int_0^{z_\infty/z_0} \frac{\varphi_*^2(z_*)dz_*}{\varphi_*^2(z_*) - 1/[\omega_0 G(\omega)]}, \qquad (7)$$

where $C_0 \equiv 2\pi l_H^2 z_0 n_{imp}$ and $\omega_0 \equiv -(a/z_0)\omega_c$.

3. Let us consider the problem to within the first-order approximation in the impurity concentration. To this end, the exact Green's function G in formulas (5) and (7) has to be replaced by the seeding function $G_0 = 1/\omega$. This approximation corresponds to the summation of all cross diagrams corresponding to the electron scattering on a single impurity. Expressions (4) and (7) indicate that the Green's function G formally has a pole at $\omega = 0$:

$$G = \frac{1}{\omega[1 + C_0 J(\omega)]}. \qquad (8)$$

However, it follows from (7) that the integral with respect to z_* diverges as $\omega \longrightarrow 0$. This makes the pole residue small, which is related to a large value of the truncation parameter z_∞:

$$\frac{1}{\pi}\mathrm{Im}G(\omega) = \delta(\omega)\frac{1}{1 + C_\infty}; \quad C_\infty = C_0\frac{z_\infty}{z_0}. \qquad (9)$$

Taking $z_\infty = \infty$, we arrive at a conclusion that the infinite degeneracy of the Landau levels is completely removed already in the first-order approximation with respect to the impurity concentration. The density of states $\rho(\omega)$ is determined only by features of the $J(\omega)$ function given by formula (7):

$$\rho(\omega) = \frac{-C_0\mathrm{Im}J(\omega)}{\pi\omega[(1 + C_0\mathrm{Re}J(\omega))^2 + C_0^2(\mathrm{Im}J(\omega))^2]}. \qquad (10)$$

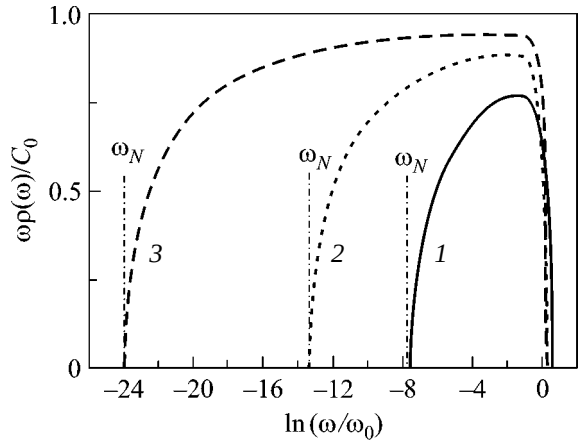
Consider a model system with the reduced wave function φ_* defined by the relations

$$\varphi_*^2(z_*) = e^{-z_*} \text{ at } z_* > 0,$$
$$\varphi_*^2(z_*) = 0 \text{ at } z_* < 0. \qquad (11)$$

This function corresponds to a boundary condition at $z = 0$ for a narrow and deep potential well [17]. Here, the function $J(\omega)$ (7) acquires the following form:

$$J(\omega) = \ln\frac{\omega - \omega_0}{\omega}. \qquad (12)$$

Assuming that there is only one type of point impurities with a negative scattering length a (which corresponds



Reduced density of states $\omega\rho(\omega)/C_0$ plotted in a logarithmic scale for three values of the impurity concentration $C_0 = 0.2$ (*1*), 0.1 (*2*), and 0.05 (*3*). The values of ω_N correspond to energies near which $(\omega/C_0)\rho(\omega)$ is a universal function of ω/ω_N.

to repulsion of electrons from the impurity) and using formulas (10) and (12), we obtain

$$\rho(\omega) = \frac{C_0}{\omega\left[\left(1 + C_0\ln\frac{\omega_0 - \omega}{\omega}\right)^2 + C_0^2\pi^2\right]}. \quad (13)$$

In this case, the density of states $\rho(\omega)$ is determined by a cut of the one-particle Green's function $G(\omega)$, rather than by the pole. Taking into account the definition of ω_0 in formula (7), we obtain a criterion of weak electron–impurity interaction $V(\mathbf{r})$ in terms of the parameter a given by formula (3): $a \ll z_0$. Thus, if the length of electron scattering from impurities is smaller than the characteristic scale of the transverse quantization in the z axis, the approximation of a single Landau level is valid and the densities of states for various levels exhibit no intersections.

According to formula (13), the density of states $\rho(\omega)$ in a wide range of the electron energies is inversely proportional to the ω value:

$$\rho(\omega) = \frac{C_0}{\omega} \text{ at } C_0\ln\frac{\omega_0}{\omega} \ll 1. \quad (14)$$

The density of states $\rho(\omega)$ described by formula (13) is integrable and, to within the terms proportional to the square impurity concentration C_0, obeys the relation

$$\int\rho(\omega)d\omega \cong \int_0^{\omega_0}\frac{C_0 d\omega}{\omega\left[1 + C_0\ln\frac{\omega_0 - \omega}{\omega}\right]^2} \cong 1. \quad (15)$$

Therefore, the number of electron states remains unchanged in the presence of the electron–impurity interaction, although the δ-like peak of the seeding density of states $\rho_0(\omega)$ is absent in the $\rho(\omega)$ function. This implies that, for an arbitrarily small value of the parameter ω, we can find a point impurity with the coordinate $z_i = z_0\ln(\omega_0/\omega) \gg z_0$ for which the degeneracy of the Landau level will be removed. This conclusion is independent of the form of the wave function $\varphi(z)$. For a physical function $\varphi(z)$ determined by relations (1), the density of states $\rho(\omega)$ is given by the same expression (13) for $\omega \ll \omega_0$ to within the terms on the order of $\ln(\omega_0/\omega)$. The only important condition is that $\varphi(z) \sim e^{-z/z_0}$ for $z \gg z_0$.

Moreover, we can always find a broad range of ω where $\rho(\omega)$ is inversely proportional to the energy ω. For example, in the case of an oscillator wave function $\varphi_*(z_*) \sim e^{-z_*^2}$, expression (10) yields for $\omega \ll \omega_0$

$$\rho(\omega) = \frac{C_0}{2\omega(\ln(\omega_0/\omega))^{1/2}[1 + C_0(\ln(\omega_0/\omega))^{1/2}]^2}. \quad (16)$$

In the presence of a strong electric field forcing electrons to move at the interface of two condensed media, $\varphi_*^2(z_*) \sim e^{-z_*^{3/2}}$, expression (10) yields for $\omega \ll \omega_0$

$$\rho(\omega) = \frac{2C_0}{3\omega(\ln(\omega_0/\omega))^{1/3}[1 + C_0(\ln(\omega_0/\omega))^{2/3}]^2}. \quad (17)$$

Thus, allowance for the electron–impurity interaction in open 2D electron systems completely removes degeneracy of the Landau levels. The density of states $\rho(\omega)$ is described by expressions (13)–(17). Generalization to the case of impurities of several types is trivial. For example, in the case of repulsing and attracting impurities, an analogue of relation (14) is

$$\begin{aligned}\rho(\omega) &= C_-/\omega \text{ at } \omega > 0,\\ \rho(\omega) &= C_+/\omega \text{ at } \omega < 0,\end{aligned} \quad (18)$$

where C_- and C_+ are the total concentrations of the repulsive and attractive impurities, respectively.

4. The density of states $\rho(\omega)$ described by formula (13) exists for

$$\omega_0 e^{-1/C_0} \ll \omega, \quad \omega_0 - \omega \gg e^{-1/C_0}. \quad (19)$$

Outside this region, it is necessary to take into account terms of the higher orders in the impurity concentration C_0, since these terms are multiplied by a large quantity $\ln[(\omega_0 - \omega/\omega)]$. In order to study the Green's function in a wider interval of energies ω, let us consider the self-consistent approximation determined by formulas (4), (5), and (7). Substituting wave function (11) into Eqs. (4) and (7), we obtain an equation for the function G:

$$G = \frac{1}{\omega}[1 - C_0\ln(1 - \omega_0 G)]. \quad (20)$$

For $\omega \ll \omega_0$ and $\omega_0 G \gg 1$, Eq. (20) can be simplified to

$$G = \frac{-C_0}{\omega}[\ln(-\omega_0 G) - 1/C_0] = \frac{-C_0}{\omega}\ln\left(-\frac{e\omega_N G}{C_0}\right); \quad (21)$$

where $\omega_N \equiv \omega_0 \exp(-1/C_0)C_0/e$ is the new energy scale. Introducing the new function $y \equiv -\omega G/C_0$, we can write Eq. (21) as

$$y = \ln(y\omega_N/\omega) + 1, \quad (22)$$

from which it follows that y is the function of a single real variable $y = y(\omega/\omega_N)$. Separating the real and imaginary pats of this function, $y = A + iB$, and using relation (22), we obtain a system of two equations for the real quantities A and B. This system can be presented in the following form:

$$A = B/\tan B,$$
$$e^{B/\tan B}\frac{\sin B}{B} = \frac{\omega_N e}{\omega}. \quad (23)$$

This equation determines the density of states $\rho(\omega) = C_0 B/\pi\omega$. We are interested in a solution such that $B(\omega/\omega_N) \geq 0$. The function $B(\omega/\omega_N)$ monotonically increases from zero at $\omega = \omega_N$ to $B = \pi$ for $\omega/\omega_N \longrightarrow \infty$. In the vicinity of $\omega = \omega_N$, this function can be approximated as $B \approx \sqrt{2(\omega - \omega_N)/\omega_N}$. For $\omega_N \ll \omega \ll \omega_0$, we have $B = \pi[1 - 1/\ln(\omega/\omega_N)]$ to within the double logarithmic terms in ω/ω_N. Thus, for $\omega \ll \omega_0$, the density of states is given by a universal function of ω/ω_N, the asymptotic form of which for $C_0 \ll 1$ is as follows:

$$\omega\rho(\omega)/C_0 = \sqrt{2(\omega - \omega_N)/\omega_N}/\pi,$$
$$\omega - \omega_N \ll \omega_N; \quad (24)$$
$$\omega\rho(\omega)/C_0 = 1 - 1/\ln(\omega/\omega_N), \quad \omega \gg \omega_N.$$

Now consider the region of $|\omega - \omega_0| \ll \omega_0$ in which $\mathrm{Im}\,G_0(\omega)$ exhibits a singularity in the first-order approximation with respect to the concentration C_0. The solution $G(\omega)$ of the self-consistent Eqs. (4) and (5) is not singular in this region and monotonically decays to zero at $\omega = \omega_x \equiv \omega_0(1 + x_0)$, where x_0 is a solution to the algebraic equation $x_0 = C_0[\ln(1/C_0) + 1 + \ln(1 + x_0)]$. For $C_0 \ll 1$, we obtain $x_0 \approx C_0[\ln(1/C_0) + 1 + C_0\ln(1/C_0)]$. In the vicinity of $\omega = \omega_x$, Eq. (20) yields (after transformations analogous to those used to derive (24))

$$B(\omega) = \sqrt{2(\omega_x - \omega)[1/(\omega_0 C_0) - 1/\omega_x]}.$$

The reduced density of states $\omega\rho(\omega)/C_0 = B/\pi$ for three concentrations of the point impurity is plotted in the figure. For $\omega \ll \omega_0$, all profiles fit a universal curve that shifts along the abscissa axis together with ω_N. This solution correctly describes the behavior of the density of states $\rho(\omega)$ until this function goes to zero (i.e., in the interval of $\omega_N < \omega < \omega_x$). Outside this interval, the density of states exhibits tails, which are exponentially small and can be determined only using diagrams with intersections of the impurity lines.

In conclusion, it should be emphasized that the proposed physical pattern is at variance with the notions formulated in the literature. Some rigorous results of the theory of electron–impurity interactions in 2D electron systems are reported in [8–10]. However, these works consider the 2D systems with point impurities whose potential $V(r)$ in Eq. (2) has the form of

$$V(r) = u_0\delta(x - x_i)\delta(y - y_i). \quad (25)$$

Since the impurity coordinate in the z axis is undetermined, the potential $V(r)$ described by expression (25) corresponds to a δ-like filament, rather than a point impurity. In contrast, the results presented above were obtained for three-dimensional impurities. Should the averaging over the coordinate z_i be "forgot," we pass to the nonphysical limit of δ-like filaments.

This study was supported by the Russian Foundation for Basic Research and by INTAS (grant no. N01-0791).

REFERENCES

1. V. B. Shikin and Yu. P. Monarkha, *Two-Dimensional Charged Systems in Helium* (Nauka, Moscow, 1989).
2. *The Quantum Hall Effect*, Ed. by R. Prange and S. M. Girvin (Springer, New York, 1987; Mir, Moscow, 1989).
3. E. M. Baskin, L. N. Magarill, and M. V. Éntin, Zh. Éksp. Teor. Fiz. **75**, 723 (1978) [Sov. Phys. JETP **48**, 365 (1978)].
4. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Prentice Hall, Englewood Cliffs, N.J., 1963).
5. Yu. A. Bychkov, Zh. Éksp. Teor. Fiz. **39**, 1401 (1960) [Sov. Phys. JETP **12**, 971 (1960)].
6. Tsunea Ando, J. Phys. Soc. Jpn. **36**, 1521 (1974); J. Phys. Soc. Jpn. **37**, 622 (1974).
7. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1974; Pergamon, New York, 1977).
8. F. Wegner, Z. Phys. B **51**, 279 (1983).
9. E. Brezin, D. I. Gross, and C. Itzykson, Nucl. Phys. B **235**, 24 (1984).
10. W. Apel, J. Phys. C **20**, L577 (1987).

Translated by P. Pozdeev

JETP Letters, Vol. 78, No. 3, 2003, pp. 152–155. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 184–187.

Electric and Photoelectric Properties of GaAs/ZnSe–Ge/ZnSe/Al Structures with Ge Quantum Dots

I. Yu. Borodin, I. A. Litvinova, I. G. Neizvestny, A. V. Prozorov, S. P. Suprun, A. B. Talochkin, V. N. Sherstyakova, and V. N. Shumsky

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

Received July 3, 2003

Current–voltage characteristics and spectral dependences of photovoltage are investigated at $T = 4.2$ and 300 K in stress-free structures with germanium quantum dots (QDs) in the GaAs/ZnSe/QD–Ge/ZnSe/Al system. The "Coulomb staircase" type features in the current–voltage characteristic observed at room temperature without illumination are due to the Coulomb interaction of electrons in resonant tunneling through intrinsic levels in QDs. The features in the photovoltage spectra are related to the absorption of radiation in the system of discrete levels of QDs. An energy band diagram of the structure is constructed based on the experimental data. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 73.63.Kv

Interest in studying the spectra of electronic states of semiconducting quantum dots (QDs) is associated with new properties that are not observed in the systems of higher dimensionality and also with the prospects of using these properties for the development of new devices on their basis [1, 2]. Semiconducting QDs are obtained using molecular-beam epitaxy (MBE) in various systems such as Ge/Si, InAs/GaAs, and others [3–5] in which QDs are strongly stressed because of the difference in the lattice constants of the QD and matrix materials. Elastic stresses and their nonuniform distribution complicate the analysis of the spectrum of electronic states [6]. This paper reports the results of studying current–voltage characteristics and spectral dependences of photovoltage in new stress-free structures with germanium quantum dots in the GaAs/ZnSe/QD–Ge/ZnSe/Al system. With the difference in the band gap of ZnSe and Ge equal to approximately 2 eV, the misfit of lattice parameters for the ZnSe–Ge heterosystem comprises about 0.2%. Thus, this system is free of elastic stresses; nevertheless, it is possible to obtain an array of germanium QDs on the epitaxial layer of zinc selenide in certain MBE modes [7, 8].

As was shown in [7], the sizes of islands in the growth plane can be varied in the range 5–20 nm with their height equal to 1–5 nm. According to the images obtained by a scanning tunneling microscope (STM), the distinctive feature of the QD array is that the spatial ordering of islands is slightly pronounced and their density is high, namely, $(2–8) \times 10^{11}$ cm^{-2}. Larger values were observed for the (001) and (110) orientations of the substrate. The high density of islands and the low degree of ordering are apparently determined by the absence of elastic fields in the heterostructure under consideration.

The structures were obtained by the MBE technique. GaAs plates with the (001) orientation and *n*-type conduction doped up to a level of 10^{18} cm^{-3} were used as the substrate. First, an epitaxial ZnSe layer 6–10 nm thick was grown on the substrate, then the deposition of germanium QDs was carried out, and the deposited QDs were covered by a ZnSe overlayer of the same thickness. With the aim of forming an active region of small sizes, an electron-beam resist was deposited on the structure obtained, in which windows 100–150 nm in diameter were opened. The size of the window was controlled by an atomic force microscope. In the closing stage, aluminum was deposited and contact regions were formed. The scheme of the structure is shown in the inset in Fig. 1.

With the structures obtained in this way, current–voltage characteristics and spectral dependences of the photovoltage were measured at temperatures of 300 and 4.2 K. Because of small currents and the occurrence of strong spurious signals, it was not possible to measure current–voltage characteristics for the structures with a contact diameter of 150 nm at a low temperature in the cryostat. Measurements of current–voltage characteristics of test structures of a large area showed that, as the temperature varies from 300 to 4.2 K, the current decreases only by a factor of 30–40, which points to the predominance of tunneling processes.

The current–voltage characteristics measured for the structure with a contact diameter of 150 nm at room temperature exhibited features that were more pronounced after processing the curves by subtracting the ohmic component of the shunt current. The transverse current through the structure is determined by not only tunneling through QDs but also the current through the

ZnSe film. The ZnSe film is not a dielectric and can conduct a current. In addition, in injection from contacts, a current limited by the space charge can be observed. In this case, a linear voltage dependence of current must be observed in weak fields and a quadratic dependence must be observed in sufficiently strong electric fields. This averaged voltage dependence of current (given features of the "staircase" type) is actually observed; however, because an analysis of the passage of current in ZnSe is beyond the scope of this paper, Fig. 1 demonstrates a current–voltage characteristic in relatively weak fields.

It is evident in the figure that variations of current of the Coulomb staircase type are observed, but the fact that these variations start at a voltage of less than 50 mV and are not strictly periodic in voltage is the main feature of the current–voltage characteristic.

Spectral characteristics were measured by an alternating signal with a frequency of ~10^2 Hz with illumination on the substrate side. The photoresponse spectra of the structure at the temperatures T = 300 K and T = 4.2 K without an external bias are given in Fig. 2. The occurrence of photovoltage indicates that a built-in filed exists in the structure, which is due to the contact potential difference. In saturation, the open-circuit photovoltage reached 0.12 V.

It is of interest that the long-wavelength edge of the spectrum starts approximately with 1.3 μm at room temperature and with 1.15 μm at liquid-helium temperature. This position of the edge does not correspond to the absorption edge in bulk germanium; at the same time, its shift correlates well with the change in the band gap in Ge upon the change in temperature from 300 to 4.2 K. The characteristic measured at a temperature of 4.2 K exhibits features that can be associated with the discrete character of the spectrum. In the high-energy region, the spectrum edge sharply terminates, being cut off by the substrate absorption.

The features of the "Coulomb staircase" type observed in the current–voltage characteristic are not strictly periodic. Irregularity in the period of Coulomb oscillations (and in the repeatability of the stairs of the Coulomb staircase) is a typical phenomenon for vertical QDs with a small number of electrons in them because of strong effects of electron–electron interaction [9]. In our case, however, because tunneling can proceed through QDs with somewhat differing sizes, the irregularity can also be associated with this circumstance. Moreover, the shape of the current–voltage characteristic will be determined by not only Coulomb interaction but also the density of the intrinsic energy levels of QDs and their width.

It can be suggested that the appearance of current and the occurrence of features in the current–voltage characteristic at low voltages, that is, in the case when the difference of the Fermi levels in aluminum and gallium arsenide is small indicate that at least the first

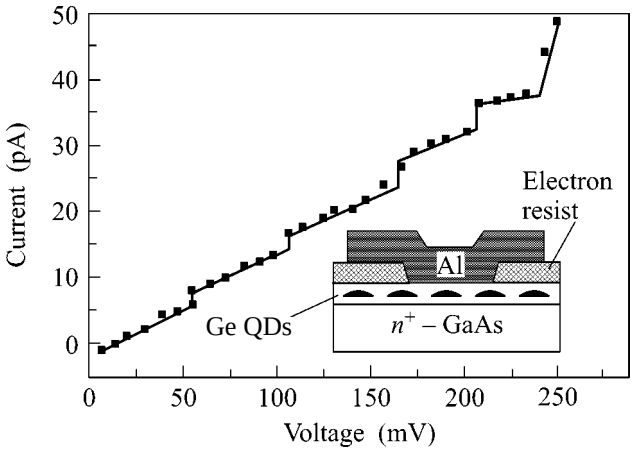


Fig. 1. Current–voltage characteristic of a structure with a contact 150 nm in diameter at T = 300 K. Points correspond to experimental data, and solid line corresponds to an approximation of the current–voltage characteristic.

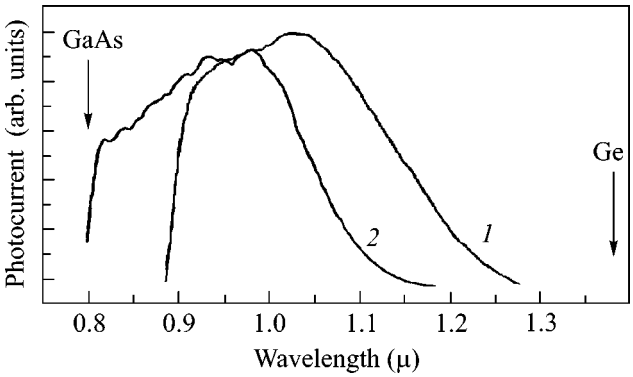


Fig. 2. Spectral dependence of a photosignal without a shift. Illumination on the GaAs side. Curve *1* was measured at T = 300 K, and curve *2* was measured at T = 4.2 K. Arrows indicate the edge of the GaAs band gap and the onset of direct transitions in bulk Ge at 4.2 K.

quantum level of an electron in QDs lies below the Fermi level of the metal in the equilibrium state (at V = 0).

Consider the region of the structure in which photoactive absorption arises, leading to the appearance of photovoltage. Photovoltage can form only through excitation of electrons and holes (electron–hole pairs) followed by their spatial separation caused by the built-in electric field. It is evident that neither GaAs nor ZnSe can be regions of electron–hole pair generation, because the band-gap width in these compounds at T = 4.2 K is larger than 1.5 eV and impurity absorption cannot give rise to photovoltage. Absorption in the QD array can lead to the formation of localized electrons and holes, but a mechanism of their spatial separation must exist.

In the structure under study, a built-in electric field exists, which is generated by the contact potential dif-

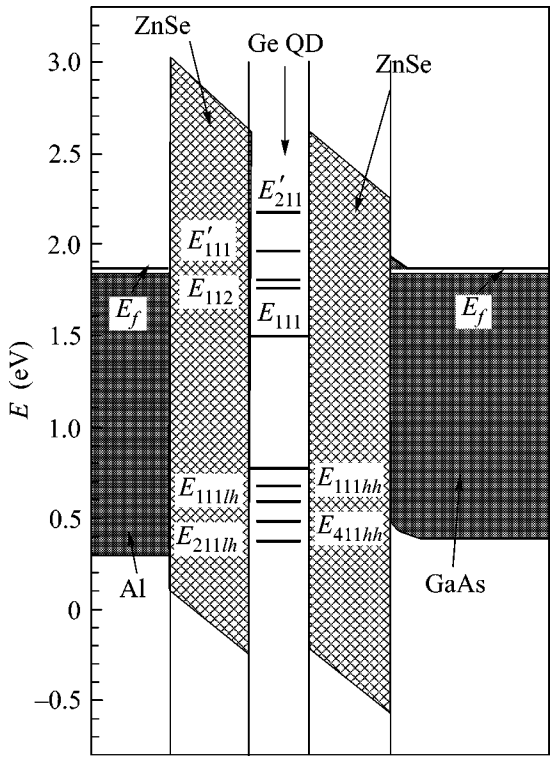


Fig. 3. Low-energy levels (three digits, the first one corresponds to the QD height) of the subband of heavy (subscript *hh*) and light (*hl)* holes and electrons for the *L* valley and the Γ valley (primed).

ference between Al and GaAs, and its value is approximately equal to $(3–5) \times 10^5$ V/cm. As the measured current–voltage characteristics show, the ZnSe layers are tunnel-transparent for electrons. Hence, upon illumination, the photoexcited electron of a generated electron–hole pair can tunnel in this field from a QD to GaAs. Further, two possibilities can be in principle accomplished: a nonequilibrium hole from a QD can tunnel to Al, or an electron from Al can pass to a QD. In any case, a photocurrent will flow in the external circuit and a photovoltage will arise in the system in the open-circuit mode.

Experimental data (the known sizes of QDs, current–voltage characteristics, and the spectral dependence of the photovoltage) along with the data on the magnitude of band discontinuities for the heterojunctions entering into the structure allow a band diagram to be constructed for the GaAs/ZnSe/QD–Ge/ZnSe/Al structure.

The band diagram was calculated from the values of the valence band discontinuity between GaAs and ZnSe equal to 1.3 eV and that between ZnSe and Ge equal to 1.0 eV. The values of the band discontinuity between the materials entering into the structure that are available in the literature vary in a wide range. This is connected with differences in technological conditions (substrate preparation and growth temperature), substrate stoichiometry, the presence of foreign atoms, the plane of substrate orientation, and a number of other factors. In the work by G. Kremer [10], an analysis of experimental data was carried out, and the conclusion was made that, in the case of growth on the nonpolar (110) plane in the system of GaAs, ZnSe, and Ge materials with regard to the layer deposition sequence, the values of the band discontinuity equal

$$\Delta E_{v\,\mathrm{ZnSe/GaAs}} = 0.96\ \mathrm{eV}, \quad \Delta E_{c\,\mathrm{ZnSe/GaAs}} = 0.35\ \mathrm{eV},$$
$$\Delta E_{v\,\mathrm{Ge/ZnSe}} = 1.52\ \mathrm{eV}, \quad \Delta E_{c\,\mathrm{Ge/ZnSe}} = 0.5\ \mathrm{eV},$$
$$\Delta E_{v\,\mathrm{ZnSe/Ge}} = 1.29\ \mathrm{eV}, \quad \Delta E_{c\,\mathrm{ZnSe/Ge}} = 0.73\ \mathrm{eV}.$$

At the same time, the valence band discontinuity $\Delta E_{v\,\mathrm{Ge/ZnSe}}$ found in [11] comprises 0.44 eV for the (001) plane, which substantially differs from the value of 0.93 eV for the (110) plane given in the same work and disagrees with the data published in [10]. With this scatter in the data, the values of the band discontinuity were adjusted from the following conditions:

(1) the position of the first quantum electron level in the *L* valley must be lower than the position of the Fermi level of the system in equilibrium, which was determined from the difference in the work function of Al and GaAs equal to 0.8 eV;

(2) the optical transition with the lowest energy measured in the experiment must correspond to the transition of an electron from a quantum level in the QD valence band to the unoccupied electronic level closest to the Fermi level. Note that the observed optical transition energy corresponds to a transition in the vicinity of the $\Gamma_{25'}$ maximum of the valence band to the $\Gamma_{2'}$ minimum of the conduction band.

The lowest radiation energy at the absorption of which photovoltage is observed can be written as

$$E_{\min} : E_{111hh} + E_{g\,\mathrm{Ge}} + (E_f - E_{c\,\mathrm{Ge}}),$$

where E_{111hh} is the first quantum level of heavy holes, $E_{g\,\mathrm{Ge}}$ is the germanium band-gap width at $T = 4.2$ K and $E_f - E_{c\,\mathrm{Ge}}$ is the difference between the Fermi level and the bottom of the germanium conduction band, which (rather than quantum levels occupied by electrons) determines the optical transition with the lowest photon energy. For $T = 4.2$ K, the lowest experimental quantum energy at which a photosignal is observed comprises approximately 1.1 eV.

STM studies of QDs provide their shape and distributions by height and diameter as functions of the average thickness of the deposited germanium layer [7, 8]. Based on these data, the pattern of the energy spectrum of electron and hole levels in QDs $3 \times 15 \times 15$ nm in size was qualitatively evaluated. Here, the shape of QDs was approximated by a parallelepiped with regard to the effective mass of the corresponding subband of holes but without regard to Coulomb interactions and with the use of Bastard's boundary conditions. The values of the effective mass of carriers in ZnSe and various valleys of Ge and also the values of the band-gap width

in Ge and ZnSe at a temperature close to 0 K were used in the calculations.

The energy band diagram of the structure constructed with the use of these calculations, published data, and the experimental conditions indicated above is presented in Fig. 3.

From the band diagram presented here, it can be seen that QDs in an equilibrium state contain levels positioned below the Fermi level of the system. Therefore, the tunneling of electrons from the aluminum contact must proceed upon applying a potential corresponding to the self-capacitance, that is, at $V \approx 40$–50 mV. Here, the occurrence of a small potential in GaAs must not substantially limit tunneling, because the concentration of electrons in GaAs is high ($n = 10^{18}$ cm^{-3}), and the thickness of the space-charge region near the GaAs–ZnSe heteroboundary does not exceed several tens of angstroms. As the voltage increases, the current is limited by only the charging of QDs, because the difference in energy between the neighboring electron levels

$$\Delta E = E_{1i+1j+1} - E_{1i+1j} = E_{1i+1j+1} - E_{1ij+1} < 2 \text{ meV}.$$

Here, the first subscript of the energy corresponds to the height, and the second and the third subscripts correspond to the base of the QD.

Because absorption in QDs is limited on the short-wavelength side by absorption in GaAs, it is evident that transitions to the E'_{211} level cannot be observed. The occurrence of features in the spectral characteristic of the photosignal can be associated with the difference in the absorption coefficient for different transitions. The absence of a pronounced discrete spectrum is explained by the fact that the experimental spectrum of the photovoltage contains optical transitions in QDs of different sizes. The occurrence of dots with a scatter in sizes follows from the histogram of the distribution by height obtained in the work [6].

Thus, the current–voltage characteristic and photoelectric properties of a stress-free n+GaAs/ZnSe/Ge–QD/ZnSe/Al structure with QDs were studied. Features associated with Coulomb effects were detected in the current–voltage characteristic at room temperature. Photovoltage arising upon light absorption in QDs was detected. A band diagram of the system was constructed based on experimental data. This diagram does not contradict the data on the values of band discontinuities available in the literature for the corresponding heterojunctions. It adequately describes the observed current–voltage characteristic and the spectral dependence of the photovoltage.

The authors are grateful to A.E. Plotnikov for help in manufacturing the structures.

This work was supported by the Russian Foundation for Basic Research, project no. 02-02-17800, and by the projects “Surface atomic structures” and “Low-dimensionality quantum structures.”

REFERENCES

1. F. A. Reboredo and Alex Zunger, Phys. Rev. B **62**, R2275 (2000).
2. M. Palummo, G. Onida, R. Del Sole, *et al.*, Phys. Status Solidi B **224**, 247 (2001).
3. N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 385 (1998) [Semiconductors **32**, 343 (1998)].
4. I. L. Krestnikov, M. V. Maximov, A. V. Sakharov, *et al.*, J. Cryst. Growth **184/185**, 545 (1998).
5. Y. W. Zhang, Phys. Rev. B **61**, 10388 (2000).
6. M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11969 (1995).
7. I. G. Neizvestny, S. P. Suprun, A. B. Talochkin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 1135 (2001) [Semiconductors **35**, 1088 (2001)].
8. I. G. Neizvestny, S. P. Suprun, V. N. Shumsky, *et al.*, Nanotechnology **12**, 437 (2001).
9. L. P. Kouwenhoven, Ch. M. Marcus, P. L. Mceuen, *et al.*, in *Proceedings of Advanced Study Institute on Mesoscopic Electron Transport*, Ed. by L. Sohn, L. P. Kouwenhoven, and G. Schön (Kluwer Academic, Dordrecht, 1997).
10. H. Kroemer, Surf. Sci. **132**, 543 (1983).
11. G. Bratina, R. Nicolini, L. Sorba, *et al.*, J. Cryst. Growth **127**, 387 (1993).

Translated by A. Bagatur’yants

JETP Letters, Vol. 78, No. 3, 2003, pp. 156–161. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 188–193.

On the Effect of Far Impurities on the Density of States of Two-Dimensional Electron Gas in a Strong Magnetic Field¶

I. S. Burmistrov[1, 2] and M. A. Skvortsov[1]

[1] *Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117940 Russia*

[2] *Institute for Theoretical Physics, University of Amsterdam, 1018 XE Amsterdam, The Netherlands*

e-mail: burmi@itp.ac.ru, skvor@itp.ac.ru

Received July 3, 2003

The effect of impurities situated at different distances from a two-dimensional electron gas on the density of states in a strong magnetic field is analyzed. Based on the entire result of Brezin, Gross, and Itzykson, we calculate the density of states in the entire energy range, assuming the Poisson distribution of impurities in the bulk. It is shown that, in the case of small impurity concentration, the density of states is qualitatively different from the model case when all impurities are located in the plane of the two-dimensional electron gas. © *2003 MAIK "Nauka/Interperiodica"*.

PACS numbers: 73.20.At; 73.20.Hb

1. INTRODUCTION

Two-dimensional electrons in a quantizing magnetic field H have been attracting much attention [1], especially since the discovery of the quantum Hall effect [2]. The properties of two-dimensional electrons in the magnetic field are affected by the presence of electron–electron interactions, as well as by impurities. Investigation of the density of states as a function of the magnetic field and filling fraction allows us to estimate the inhomogeneities caused by impurities in experimental samples [3]. Although the electron–electron interaction should usually be taken into account, the question of the density of states in the simplest model of noninteracting electrons is also rather interesting.

In the absence of interaction, impurities near a two-dimensional electron gas (2DEG) provide the only mechanism for broadening of Landau levels. In a weak magnetic field, a large number of Landau levels, $N \gg 1$, are filled. One can therefore use the self-consistent Born approximation that is justified by the small parameter $\ln N/N \ll 1$. This results in the well-known semicircle shape for the density of states [4]. Beyond the self-consistent Born approximation, one can find exponentially small tails in the density of states [5].

In the opposite limit of a strong magnetic field, only the lowest Landau level is partially occupied. In this case, one can neglect the influence of the other empty Landau levels assuming $\omega_H \gg T, \tau^{-1}$. Here, $\omega_H = eH/m$ denotes the cyclotron frequency, where e and m are the electron charge and mass, respectively, T stands for the temperature, and τ is the elastic collision time. The density of states on the lowest Landau level strongly depends on the statistical properties of the random potential created by impurities and on the value of the dimensionless parameter n_S/n_L, where $n_L = 1/(2\pi l_H^2)$ with the magnetic field length $l_H = 1/\sqrt{m\omega_H}$ and n_S stands for the two-dimensional impurity density. For the white-noise distribution of the random potential, the density of states was found exactly by Wegner [6]. For arbitrary statistics of the random potential, the density of states was obtained exactly in a beautiful paper by Brezin, Gross, and Itzykson [7]. If the number of impurities is less than the number of states on the Landau level, $n_S < n_L$, the Landau level remains partially degenerate. In the opposite case, $n_S \geq n_L$, the presence of impurities leads to complete lifting of the degeneracy of the Landau level [7, 8].

In experimental samples, impurities can be found rather far from the 2DEG [1, 2]. In such a situation, the two-dimensional electron system is subject to the three-dimensional random potential. This means that an electron localized at the heterojunction feels impurities situated at distances much larger than the width z_0 of the 2DEG. This situation was considered recently by Dyugaev, Grigor'ev, and Ovchinnikov [9]. In the lowest order of the perturbation theory in the concentration n_{imp} of three-dimensional scatterers, they calculated the density of states $D(E)$ in the limit when the multiple scattering on the same impurity provides the main contribution. Assuming exponential decay of the wave function in the transverse direction, $\varphi^2(z) \propto \exp(-z/z_0)$, they obtained a universal regime where $D(E) = n_{\text{imp}} z_0/E$, and energy E is measured from the unperturbed Landau level. Being bounded from the sides of both small and large energies by many-impurity effects, this interval contains most of the states of the unper-

¶ This article was submitted by the authors in English.

turbed Landau level. Though the analysis of [9] holds for an arbitrary Landau level, it cannot be generalized to the limits of small and large energies, where a nonperturbative treatment of impurity scattering is required.

The main objective of the present letter is to present the full analysis of the effect of far impurities on the density of states of a two-dimensional electron gas in a strong magnetic field. Employing the remarkable result of Brezin, Gross, and Itzykson [7], we calculate the broadening of the lowest Landau level by the three-dimensional short-range impurities with the Poisson distribution in the bulk.

2. RESULTS

Usually, impurities occupy a rather large volume near a two-dimensional electron gas and, consequently, their number exceeds the number of states at the Landau level, $N_{\rm imp} \gg n_L S$, with S being the area of the two-dimensional electron system. Therefore, the degeneracy of the Landau level is removed completely by impurities [7, 9]. The behavior of the density of states is determined by the new dimensionless parameter

$$f = \frac{n_{\rm imp} z_0}{n_L}, \tag{1}$$

which will be referred to as impurity concentration. Here, $n_{\rm imp}$ is the three-dimensional impurity density and z_0 stands for the spatial extent of the electron wave function in the direction perpendicular to the 2DEG, explicitly defined in Eq. (4).

In experiments, there is usually a small amount of impurities in a layer of width z_0 near the two-dimensional electron gas [1, 2]; i.e., impurity concentration is small, $f \ll 1$. In this case, we obtain the following density of states at the lowest Landau level as a function of the deviation E from the unperturbed level $\omega_H/2$:

$$\frac{D(E)}{n_L} = \begin{cases} 0, & E < 0, \\ \dfrac{1}{E_0} S_f\left(f \ln\dfrac{E_0}{E}\right), & 0 \le \dfrac{E}{E_0} \ll e^{-1/f}, \\ \dfrac{f}{E}, & e^{-1/f} \ll \dfrac{E}{E_0} \ll 1, \\ \dfrac{2\pi^{-1/2} E^2}{(f_1 E_1^2)^{3/2}} \exp\left(-\dfrac{E^2}{f_1 E_1^2}\right), & E_0 \ll E, \end{cases} \tag{2}$$

where

$$S_f(\xi) = \sqrt{\frac{f}{2\pi}} \frac{\xi - 1}{\Gamma(\xi)} \exp\left(\frac{\xi}{f} - \frac{(\xi-1)^2}{2f} - \frac{f}{2}\ln^2\xi\right). \tag{3}$$

In deriving Eq. (2), we assumed that the wave function $\varphi(z)$ decays in the transverse direction as

$$\varphi^2(z) \sim \frac{A}{z_0} \exp\left(-\frac{z}{z_0}\right), \quad z \gg z_0, \tag{4}$$

where z_0 defines the width of the 2DEG and A is a constant of order 1. The form (4) corresponds to a rectangular well confining potential [1]. The energy scale

$$E_0 = e^{\gamma} A \frac{n_L u_0}{z_0} \tag{5}$$

is introduced by impurities, where $u_0 > 0$ is the strength of the repulsive disorder potential [cf. Eq. (14) below] and $\gamma \approx 0.577$ denotes Euler's constant. The result of $E \gg E_0$ is governed by the parameters

$$f_1 = \frac{n_{\rm imp} z_1}{n_L} \sim f, \quad E_1 = \frac{n_L u_0}{z_1} \sim E_0, \tag{6}$$

where $z_1 \sim z_0$ is the width of the wave function as determined via its fourth moment:

$$\frac{1}{z_0} = \int_0^\infty \varphi^4(z)\, dz. \tag{7}$$

The fact that the density of states vanishes for $E < 0$ is expected, since the random potential is purely repulsive. Since $\lim_{\xi \to \infty} S_f(\xi) = 0$, the density of states also vanishes at the position of the unperturbed Landau level, $D(0) = 0$.

In the interval $0 \le E \ll E_0 e^{-1/f}$, the density of states exhibits the maximum

$$D_{\max}(E_0) = \sqrt{\frac{f}{2\pi}} \frac{n_L}{E_0} \exp\left[\frac{3}{2f}\right] \tag{8}$$

at the exponentially small energy

$$E_* \sim E_0 \exp\left(-\frac{2}{f}\right). \tag{9}$$

In the region $E_0 e^{-1/f} \ll E \ll E_0$, the density of states $D(E) = f n_L/E$ is linear in impurity concentration coinciding with the perturbative result obtained by Dyugaev, Grigor'ev, and Ovchinnikov [9]. This indicates that the multiple scattering on the same impurity provides the main contribution to the density of states for energies $E_0 e^{-1/f} \ll E \ll E_0$. This energy interval contains the major part of the states formed from the lowest Landau level.

The result found in Eq. (2) in the limit $0 \le E \ll E_0 e^{-1/f}$ is applicable for $f \ln(E_0/E) - 1 \gg \sqrt{f}$ (cf. Eq. (31)). At the border of applicability, Eq. (3) gives $S_f[f \ln(E_0/E)] \sim fE_0/E$, and, thus, $D(E)$ merges with the universal result at $E \gg E_0 e^{-1/f}$.

In the region of rather large energies $E \gg E_0$, the tail of the density of states is described by the same expression, as if all impurities were situated in the plane of the 2DEG, with the effective two-dimensional parameters

$$u_0^{(2D)} = \frac{u_0}{z_1}, \quad n_{\text{imp}}^{(2D)} = n_{\text{imp}} z_1. \tag{10}$$

We mention that the tail of the density of states corresponds to some optimal fluctuation of the random potential, as happens for the purely two-dimensional problem [10, 11].

For large impurity concentration, $f \gg 1$, the Poisson distribution can be replaced by the white-noise distribution of impurities on the plane with the effective parameters (10). The density of states is therefore given by the well-known formula [6, 7]

$$D(E) = \frac{n_L}{\pi E_1 \sqrt{f_1}} W\left(\frac{E - f_1 E_1}{E_1 \sqrt{f_1}}\right), \tag{11}$$

where we introduce the function

$$W(z) = \frac{2}{\sqrt{\pi}} e^{z^2}\left[1 + \frac{4}{\pi}\left(\int_0^z e^{x^2} dx\right)^2\right]^{-1}. \tag{12}$$

The shift of the maximum of $D(E)$ to positive energies is related to the repulsive character of the impurities' potential. Equation (11) describes the density of states for the Poisson distribution only approximately, since the exact density of states should vanish $E \le 0$. However, the deviation of Eq. (11) from the exact answer is exponentially small ($e^{-f} \ll 1$) for positive E.

3. MODEL

The spin-polarized two-dimensional electron gas in the presence of the random potential $V(\mathbf{r}, z)$ and the strong perpendicular magnetic field $\mathbf{H}$ is described by the following one-particle Hamiltonian:

$$\mathcal{H} = -\frac{1}{2m}(\nabla - ie\mathbf{A})^2 + V(\mathbf{r}, z) + U_{\text{conf}}(\mathbf{r}, z). \tag{13}$$

Here, $\mathbf{A}$ stands for the vector potential, $\mathbf{H} = \text{rot}\,\mathbf{A}$, and $U_{\text{conf}}(\mathbf{r}, z)$ denotes the confining potential that creates the two-dimensional electron gas. We use units such that $\hbar = 1$ and $c = 1$.

We assume that impurities situated near the two-dimensional electron gas are zero-range repulsive ($u_0 > 0$) scatterers producing the random potential

$$V(\mathbf{r}) = u_0 \sum_{j=1}^{N_{\text{imp}}} \delta^{(2)}(\mathbf{r} - \mathbf{r}_j)\delta(z - z_j)! \tag{14}$$

Assuming that the confining potential U_{conf} depends only on the z coordinate, we can represent the electron wave function as follows:

$$\Psi(\mathbf{r}, z) = \psi(\mathbf{r})\varphi(z), \tag{15}$$

where $\varphi(z)$ is the ground-state wave function for the electron motion in the direction perpendicular to the 2DEG in the absence of disorder and $\psi(\mathbf{r})$ describes the electron motion in the plane of 2DEG. The decomposition (15) is equivalent to the projection onto the lowest level of dimensional quantization and is analogous to the projection onto the lowest Landau level states $\psi(\mathbf{r})$. Since, in experiment, the energy separation between the lowest and the first excited level of dimensional quantization is usually larger than the cyclotron gap, the accuracy of projection onto $\varphi(z)$ is higher than the accuracy of projection onto the lowest Landau level. With the help of the ansatz (15), the original three-dimensional problem (13) reduces to the two-dimensional one with the effective two-dimensional random potential

$$V_{\text{eff}}(\mathbf{r}) = u_0 \sum_{j=1}^{N_{\text{imp}}} \varphi^2(z_j)\delta^{(2)}(\mathbf{r} - \mathbf{r}_j). \tag{16}$$

Thus, the distribution of impurities along the z direction leads to an additional random distribution of the potential strengths $u_0\varphi^2(z_j)$ effectively felt by two-dimensional electrons.

By using the general result of Brezin, Gross, and Itzykson [7] for the random potential (16), we obtain for the density of states at the lowest Landau level

$$D(E) = \frac{n_L}{\pi}\text{Im}\frac{\partial}{\partial E}\ln F(E), \tag{17}$$

where

$$F(E) = \int_0^\infty dt \exp\left(\frac{iEt}{n_L} + \int_0^t \frac{d\beta}{\beta} g(\beta)\right). \tag{18}$$

The properties of the random potential are encoded in the function $g(\beta)$, which is defined as

$$\exp\{n_L \int d^2\mathbf{r}\, g[\beta(\mathbf{r})]\} = \left\langle \exp\left[-i\int d^2\mathbf{r}\,\beta(\mathbf{r})V_{\text{eff}}(\mathbf{r})\right]\right\rangle, \tag{19}$$

where the average $\langle \ldots \rangle$ is with respect to the distribution of the random potential $V_{\text{eff}}(\mathbf{r})$.

We assume that the three-dimensional scatterers (14) with equal strengths u_0 obey the Poisson statistics, being uniformly distributed along the z direction. Then, averaging over $V_{\text{eff}}(\mathbf{r})$ in Eq. (16) reduces to integration over the z coordinate:

$$g(\beta) = \frac{n_{\text{imp}}}{n_L}\int_0^\infty dz\,(e^{-i\beta u_0 \varphi^2(z)} - 1). \tag{20}$$

On writing Eq. (20), we employed the fact that the wave function $\varphi(z)$ vanishes for $z < 0$.

4. EVALUATION OF THE DENSITY OF STATES

The density of states is generally given by the integral representation (17), (18), and (20). However, Eq. (18) cannot be calculated in a closed form valid for arbitrary values of impurity concentration and energies. Below, we analyze the most interesting asymptotic cases.

First of all, we note that $D(E)$ vanishes for energies $E < 0$ regardless of the form of $\varphi(z)$. This follows from the fact that, for $E < 0$, the function $F(E)$ is purely imaginary, which can be obtained by performing the Wick rotation $t \longrightarrow -i\tau$ of the integration contour in Eq. (18).

The density of states can also be easily calculated in the limit of either large impurity concentration ($f \gg 1$) and arbitrary energies or small impurity concentration ($f \ll 1$) but large energies $E \gg E_0$. In both cases, integral (18) is determined by small values of t, which allows one to expand the function $g(\beta)$ given by Eq. (20):

$$\int_0^t \frac{d\beta}{\beta} g(\beta) \simeq -i\frac{n_{\text{imp}}u_0 t}{n_L} - \frac{n_{\text{imp}}u_0^2 t^2}{4n_L z_1}, \qquad (21)$$

where z_1 is defined in Eq. (7). The quadratic term in Eq. (21) describes the Gaussian (white-noise) distribution of impurities [7], whereas the linear term accounts for the energy shift due to the nonzero average potential of impurities. Employing the result of [7], we arrive at Eq. (11). Using the asymptotic expression $W(x) \simeq 2\sqrt{\pi}x^2 e^{-x^2}$ valid at $x \gg 1$, we obtain the result (2) for $f \ll 1$ in the regime $E \ll E_0$.

The most interesting is the behavior of $D(E)$ in the limit of *small impurity concentrations*, $f \ll 1$, and *sufficiently small energies*, $E \ll E_0$. In this limit, assumed hereafter, the function $F(E)$ given by Eq. (18) is determined by large values of t, which makes it possible to use asymptotic formula (4) for calculation of $g(\beta)$ in Eq. (20). Introducing the dimensionless energy $\varepsilon = E/E_0$, where the energy scale E_0 is defined in Eq. (5), and rescaling t accordingly, we rewrite the expression for the density of states as

$$D(\varepsilon) = \frac{n_L}{\pi E_0}\text{Im}\frac{\partial}{\partial\varepsilon}\ln F(\varepsilon), \qquad (22)$$

where

$$F(\varepsilon) = \int_0^\infty dt e^{ie^\gamma\varepsilon t}e^{-fh(t)}, \qquad (23)$$

$$h(t) = \int_0^t \frac{d\beta}{\beta}\int_0^\beta \frac{ds}{s}(1-e^{-is}). \qquad (24)$$

The function $h(t)$ is positive at the negative part of the imaginary axis, $t = -i\tau$, having the following asymptotic behavior at $\tau \gg 1$:

$$h(-i\tau) = \frac{1}{2}\ln^2(e^\gamma\tau) + c_0 + \hat{h}(\tau), \qquad (25)$$

where c_0 is a constant of the order 1 and $\hat{h}(\tau)$ decays exponentially at large τ:

$$\hat{h}(\tau) = -\int_1^\infty \frac{dx}{x}e^{-\tau x}\ln x \approx -\frac{1}{\tau^2}e^{-\tau}. \qquad (26)$$

The $\ln^2 t$ asymptotics of $h(t)$ is specific to the problem with distributed strengths $u_0\varphi^2(z_j)$ of impurities and asymptotic behavior (4) of the wave function $\varphi(z)$ far from the 2DEG and should be contrasted with the $\ln t$ dependence for the case of the Poisson distribution with constant impurity strengths. For another decay law of the wave function, $\varphi^2(z) \sim \exp[-(z/z_0)^\alpha]$, the leading asymptotics would be $h(t) \sim \ln^{1+1/\alpha}t$.

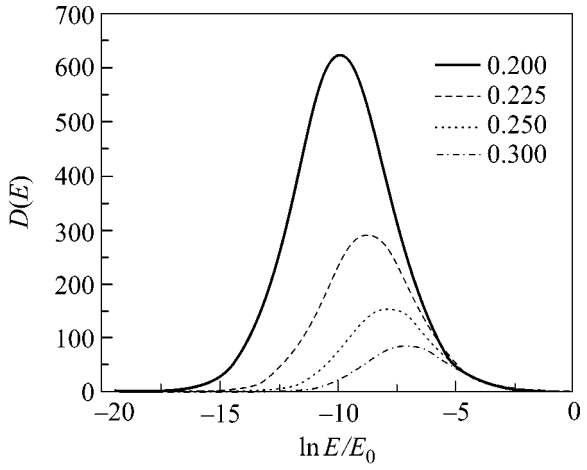
The function $F(\varepsilon)$ in Eq. (23) is given by an oscillating integral. Therefore, it is desirable to deform the integration contour to get rid of oscillations. However, for $\varepsilon > 0$, such a deformation in Eq. (23) is impossible: the first factor prohibits deformation into the lower half-plane, whereas the second factor leads to a divergent integral if deformed into the upper half-plane. This complication can be overcome by splitting the integrand into two parts, singling out the leading log-square asymptotics:

$$F(\varepsilon) = \int_0^\infty dt e^{ie^\gamma\varepsilon t - (f/2)\ln^2(e^\gamma it)}[1 + (e^{-f\hat{h}(it)} - 1)], \qquad (27)$$

where we have omitted the irrelevant factor e^{-fc_0}. In the limit $\varepsilon < 1$, the integral with the second term $(e^{-f\hat{h}(it)} - 1)$ in the square brackets allows deformation of the contour to the negative part of the imaginary axis, where the integrand is purely real. It can be shown that the resulting contribution can be neglected compared to the integral with the first term in the square brackets. The latter can be calculated by deforming the integration contour to the upper part of the imaginary axis. After a proper rescaling of variables, one finds

$$F(\varepsilon) = \int_0^\infty d\tau \exp\left(-\tau - \frac{f}{2}\left[\ln\left(\frac{\tau}{\varepsilon} + i\pi\right)\right]^2\right), \qquad (28)$$

where we have again omitted the irrelevant factor $ie^{-\gamma}/\varepsilon$.



The density of states $D(E)$ in units of $n_L/\pi E_0$ as a function of $\ln E/E_0$ for different values of impurity concentration f.

Equations (22) and (28) give the integral representation for the density of states at $E \ll E_0$. Its behavior depends on the value of the parameter

$$\xi = f \ln\frac{1}{\varepsilon}. \tag{29}$$

For *small* $\xi < 1$, i.e., not too close to the unperturbed Landau level ($e^{-1/f} \ll e \ll 1$), one can calculate $F(\varepsilon)$ perturbatively. Expanding Eq. (28) in f, one can easily recover the perturbative result of [9], as well as the leading correction to it:

$$D(\varepsilon) = \frac{n_L f}{E_0 \varepsilon}\left[1 - 2\zeta(3) f^2 \ln\frac{1}{\varepsilon} + \ldots\right]. \tag{30}$$

Retaining only the leading term, we obtain result (2) in the regime $E_0 e^{-1/f} \ll E \ll E_0$.

For *large* $\xi > 1$, corresponding to energies close to the unperturbed Landau level, evaluation of Eq. (28) is subtler. In this case, the ratio $\mathrm{Im}F(\varepsilon)/\mathrm{Re}F(\varepsilon)$ is exponentially small, and special care must be taken in order to extract $\mathrm{Im}F(\varepsilon)$. On the other hand, $\mathrm{Re}F(\varepsilon)$ can easily be calculated for $\xi - 1 \gg \sqrt{f}$. Making the substitution $\tau = \varepsilon e^p$ and calculating the resulting Gaussian integral over p, one finds

$$\mathrm{Re}F(\varepsilon) = -\varepsilon\sqrt{\frac{2\pi}{f}}\exp\left[\frac{1}{2f}\right], \quad \xi - 1 \gg \sqrt{f}. \tag{31}$$

To extract $\mathrm{Im}F(\varepsilon)$, we find it convenient to pass to another representation for the function $F(\varepsilon)$. To this end, we decouple the square term in the exponential of Eq. (28) by the Hubbard–Stratonovich transformation, and, integrating over τ, we obtain

$$F(\varepsilon) = \int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi f}}\Gamma(1 + iz)\exp\left[-\frac{z^2}{2f} - \pi z + iz\ln\frac{1}{\varepsilon}\right]. \tag{32}$$

This representation in terms of the Γ function is suitable for numerical simulation due to rather fast convergence of the integral, contrary to the initial representation (23).

To proceed, we shift the integration contour to the upper part of the complex plane: $z = i\xi + x$, with x being the new real integration variable. As soon as $\xi \geq 1$, in doing the contour transformation, we have to cross the poles of the Γ function at $z = ik$ with integer $k > 0$. As a result, we obtain

$$F(\varepsilon) = \frac{1}{\sqrt{2\pi f}}\left\{2\pi\sum_{k=1}^{[\xi]}\frac{(-1)^k}{(k-1)!}\exp\left[\frac{k^2 - 2\xi k}{2f}\right] + \exp\left(-\frac{\xi^2}{2f}\right)\Phi(\xi, f)\right\}, \tag{33}$$

where $[\xi]$ is an integer part of ξ, and the function $\Phi(\xi, f)$ is defined as

$$\Phi(\xi, f) = e^{-i\pi\xi}\int_{-\infty}^{\infty} dx\Gamma(1 - \xi + ix)e^{-\pi x}\exp\left[-\frac{x^2}{2f}\right]. \tag{34}$$

An advantage of this representation is that the pole contribution in Eq. (33) is purely real and, hence, $\mathrm{Im}F(\varepsilon)$ is determined solely by $\mathrm{Im}\Phi(\xi, f)$. Employing the identity $\Gamma(1 - \eta)\Gamma(\eta) = \pi/\sin(\pi\eta)$ with $\eta = \xi - ix$, we obtain for the imaginary part of $\Phi(\xi, f)$

$$\mathrm{Im}\,\Phi(\xi, f) = -2\pi\,\mathrm{Re}\int_0^{\infty} dx\frac{\exp\left[-\frac{x^2}{2f}\right]}{\Gamma(\xi - ix)}. \tag{35}$$

In the limit $\xi \gg 1$, the term ix in the argument of the Γ function can be taken into account as $\Gamma(\xi - ix) \simeq \Gamma(\xi)e^{-ix\ln\xi}$. Thereby, we find the following estimate:

$$\mathrm{Im}\Phi(\xi, f) = -\frac{\pi\sqrt{2\pi f}}{\Gamma(\xi)}\exp\left[-\frac{f}{2}\ln^2\xi\right]. \tag{36}$$

Though Eq. (36) is formally derived for $\xi \gg 1$, it can also be applied at $\xi \geq 1$ as well, with the error being small by virtue of the inequality $f \ll 1$.

Now, with the help of Eqs. (31), (33), and (36), we obtain, for $\xi - 1 \gg \sqrt{f}$,

$$\frac{F(\varepsilon)}{\mathrm{Re}F(\varepsilon)} = 1 + i\sqrt{\frac{\pi f}{2}}\frac{\exp\left(-\frac{(\xi-1)^2}{2f} - \frac{f}{2}\ln^2\xi\right)}{\Gamma(\xi)}. \tag{37}$$

Finally, using Eq. (22), we find

$$D(\varepsilon) = \frac{n_L}{E_0}S_f(\xi), \tag{38}$$

where $S_f(\xi)$ is defined in Eq. (3). Equation (38) gives result (2) in the region $0 \leq E \ll E_0 e^{-1/f}$.

The whole profile of $D(E)$ for $\varepsilon \ll 1$ can be obtained by numerical evaluation of Eqs. (22) and (32). The density of states numerically calculated for several values of the impurity concentration f is presented in the figure.

5. CONCLUSION

In conclusion, we evaluated the density of states of a two-dimensional electron gas in the presence of a strong magnetic field and impurities. The fact that impurities are situated at different distances from the two-dimensional electron gas leads to a dramatic change of the density of states in the case of small impurity concentration compared to the case when all impurities are situated at the same distance from the 2DEG.

Using the exact result of [7], we obtained the density of states in the entire energy range for the case of the wave function with the asymptotic behavior (4). The density of states vanishes at the position of the unperturbed Landau level and has a maximum at an exponentially small energy (9). The major part of the states is localized by single impurities in accordance with the findings of [9].

The functional form of the result will be different for asymptotic behavior of $\varphi(z)$ differing from the simple exponential decay (4). However, the qualitative structure of the density of states is supposed to be preserved.

We acknowledge useful discussions with M.V. Feigelman, S.V. Iordansky, A.S. Iosselevich, D. Lyubshin, and P.M. Ostrovsky. We are grateful to A.M. Dyugaev, P.D. Grigoriev, and Yu.N. Ovchinnikov for bringing their paper to us prior to publication. We thank Forschungszentrum Jülich (Landau Scholarship) (I.S.B.), Dynasty Foundation, ICFPM, the Russian Foundation for Basic Research (grant no. 01-02-17759), and the Russian Ministry of Science (M.A.S.) for financial support.

REFERENCES

1. T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
2. *The Quantum Hall Effect*, Eds. by R. E. Prange and S. M. Girvin (Springer, Berlin, 1987).
3. I. V. Kukushkin, S. V. Meshkov, and V. B. Timofeev, Usp. Fiz. Nauk **155**, 219 (1988) [Sov. Phys. Usp. **31**, 511 (1988)].
4. T. Ando and Y. Uemura, J. Phys. Soc. Jpn. **36**, 959 (1974); J. Phys. Soc. Jpn. **36**, 1521 (1974); T. Ando, J. Phys. Soc. Jpn. **37**, 1233 (1974).
5. K. B. Efetov and V. G. Marikhin, Phys. Rev. B **40**, 12 126 (1989).
6. F. Wegner, Z. Phys. B **51**, 279 (1983).
7. E. Brezin, D. J. Gross, and C. Itzykson, Nucl. Phys. B **235**, 24 (1984).
8. É. M. Baskin, L. I. Magarill, and M. V. Éntin, Zh. Éksp. Teor. Fiz. **75**, 723 (1978) [Sov. Phys. JETP **48**, 365 (1978)].
9. A. M. Dyugaev, P. D. Grigoriev, and Yu. N. Ovchinnikov, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 179 (2003).
10. L. B. Ioffe and A. I. Larkin, Zh. Éksp. Teor. Fiz. **81**, 1048 (1981) [Sov. Phys. JETP **54**, 556 (1981)].
11. I. Affleck, J. Phys. C **17**, 279 (1984).

JETP Letters, Vol. 78, No. 3, 2003, pp. 162–167. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 194–200.

Simple Delay Scheme for Quantum Cryptography Based on a Mach–Zehnder Optic Fiber Interferometer

S. N. Molotkov

*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia**
Faculty of Computational Mathematics and Cybernetics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received June 10, 2003; in final form, July 1, 2003

A simple experimental scheme is proposed for a relativistic quantum cryptosystem based on a Mach–Zehnder optic fiber interferometer. In this scheme, quantum mechanical laws, along with the restrictions imposed by the Special Relativity, ensure the detection of any eavesdropping attempt. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 03.67.Dd; 42.50.St; 07.60.Ly

Many schemes for quantum cryptography have already been proposed and experimentally implemented [1]. Almost all schemes of secret key distribution use two quantum mechanical exclusions: exclusion of the copying of an unknown quantum state (no-cloning theorem [2]) and exclusion of perturbation-free distinguishing between nonorthogonal quantum states [3]. These two closely connected exclusions make it possible to detect any eavesdropping attempt and ensure unconditional security. Nonrelativistic quantum mechanics does not forbid the perturbation-free cloning and distinguishing of orthogonal states [2, 3]. As was previously shown in [4–6], the security of the key can be guaranteed up to a theoretical limit of (e.g., for the BB84 protocol) 11% for the admissible errors at the receiver end [4, 6]. This limit arises in fact because it is impossible to distinguish between errors due to noise in the channel and eavesdropper-induced errors. Therefore, to guarantee key security, one must consider that all errors are induced by the eavesdropper.

The above schemes in no way explicitly use the fact that information can be transmitted at long distances only by photons (states of a massless quantized electromagnetic field). Since the quantized massless photon field propagates in a vacuum with the highest possible velocity, Special Relativity exclusions can be used to ensure secrete key distribution in quantum cryptography. In this case, the security of the key is achieved even with the use of orthogonal states. Relativistic schemes of quantum cryptography on orthogonal states of the photon field were previously proposed in [7, 8], where orthogonal states with an extent longer than the length of the communication channel were used in communication protocols. This property can be in principle achieved, but its experimental implementation is difficult. It has been shown recently that single-photon states of any length, including those shorter than the length of the communication channel, can be used for secret key distribution under exclusions imposed by both the quantum nature of the states and Special Relativity [9]. Moreover, the security of the key can be guaranteed up to a theoretical limit of 43.75% (7/16) for the errors at the receiver end. This limit is noticeably higher than the admissible error limit for nonrelativistic schemes, because eavesdropper-induced errors can be partially distinguished from the errors due to noise in the channel.

In this paper, a simple realization of a quantum cryptosystem based on a Mach–Zehnder optic fiber interferometer is proposed. Contrary to the previous schemes, it does not require polarization control and ideal balance of the arms of the interferometer at the transmitter and receiver ends. Moreover, the admissible error probability at the receiver end is equal to about 25%. Although this value is lower than the theoretical limit for relativistic schemes [9], it is twice as high as that for the nonrelativistic schemes based on similar interferometers. This circumstance is due to the fact that exclusions imposed by quantum mechanics and Special Relativity are more stringent than purely quantum-mechanical exclusions.

The impossibility of copying arbitrary quantum states $|\varphi_0\rangle$ and $|\varphi_1\rangle$ means the impossibility of the process [2]

$$|\varphi_0\rangle \mapsto |\varphi_0\rangle|\varphi_0\rangle, \quad |\varphi_1\rangle \mapsto |\varphi_1\rangle|\varphi_1\rangle. \tag{1}$$

This exclusion for orthogonal states is absent in quantum mechanics [2]. The impossibility of acquiring information about one of the quantum states $|\varphi_0\rangle$ and

* Basic affiliation.

0021-3640/03/7803-0162$24.00 © 2003 MAIK "Nauka/Interperiodica"

$|\varphi_1\rangle$ without their perturbation means the impossibility of the process [3]

$$U(|\varphi_0\rangle A\rangle) \mapsto |\varphi_0\rangle|A_0\rangle,$$
$$U(|\varphi_1\rangle A\rangle) \mapsto |\varphi_1\rangle|A_1\rangle, \quad |A_0\rangle \neq |A_1\rangle, \tag{2}$$

if the states are nonorthogonal; i.e., $\langle\varphi_0|\varphi_1\rangle \neq 0$. The reliable perturbation-free distinguishing between orthogonal states is not forbidden [3]. More precisely, the theorem proved in [3] does not apply to this case. This theorem is often interpreted as follows: an orthogonal state "passes" through the auxiliary system $|A\rangle$, interacts with it, and changes its state. However, the theorem does not imply this interpretation. The theorem is purely geometric and states that the state vector of the auxiliary system $|A\rangle$ can be unitarily transformed to the new state $|A_0\rangle$ or $|A_1\rangle$ for the input vector $|\varphi_{0,1}\rangle$, respectively, without a change in the input vector. In this case, it is implicitly assumed that the input vector $|\varphi_{0,1}\rangle$ is available as a holistic object; i.e., the unitary transformation U requires access to the entire state space $\mathcal{H}_{\varphi_{0,1}}$ in which the state support is nonzero. Otherwise, the transformation is not unitary. The fact that the state vector is treated in the proof only as a holistic object $|\varphi_{0,1}\rangle$ without internal coordinates just implies that the unitary transformation involves the state vector as a whole.

The Hilbert space $\mathcal{H}_{\varphi_{0,1}}$ for any real physical system is inevitably referred to the Minkowski spacetime, where a state has the amplitude (smoothing wave function). Access to the Hilbert space inevitably implies the access to the spacetime region where the amplitude (wave function) of the state is nonzero. If only the spacetime region where the amplitude of the states is nonzero is accessible, even orthogonal states cannot be reliably copied or distinguished. The last statement is more or less evident, because the outcome probability of any process including copying or distinguishing cannot be higher than the state-normalization fraction in the accessible spacetime region, i.e., in the accessible region of the Hilbert space. Roughly speaking, to reliably copy or distinguish orthogonal states, they must be simultaneously and entirely accessible.

Thus, if the state amplitude is nonzero in a certain finite spacetime region, the accessibility of the entire state means that this region is accessible. In nonrelativistic quantum mechanics, where the maximum velocity is absent, access to any finite region is instantaneous. In quantum field theory, where the maximum possible velocity exists, access to the entire state is possible only if the extended state is previously unitarily transformed to a state whose amplitude is nonzero only in an infinitely small space region. After that, one can use the theorem proved in [2, 3]. This unitary transformation of the state specified in a finite spacetime region to the state localized in an infinitely small space region can be made only in a finite time by virtue of the relativistic causality principle [10]. The minimum necessary time is determined by the condition that the initial space region where the state amplitude is nonzero is covered by the past light cone (see Fig. 1). The vertex of this cone is in the infinitely strongly localized region (point) to which the initial state amplitude is transformed. Each of the pair of orthogonal states unitarily transformed (contracted) to the localized region can be then reliably copied or distinguished. Since we consider massless states of a quantized field (photons), which propagate with the maximum possible velocity, this unitary transformation and further copying will lead to a shift (delay) of the states in spacetime as compared to the original free evolution (propagation) of the states. This circumstance makes it possible to detect any eavesdropping attempt. We note that exclusions imposed on measurements in the relativistic region were analyzed in [11, 12].

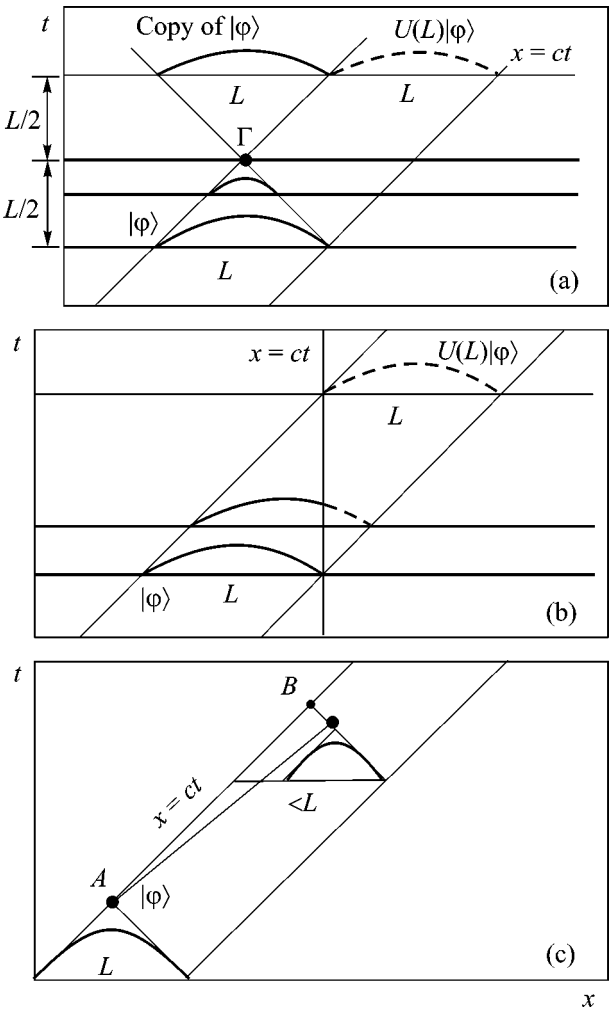


Fig. 1.

In other words, the exclusion theorem for orthogonal states of a massless quantized field is formulated as follows. Orthogonal states can be copied with a proba-

bility arbitrarily close to unity. However, copying provides states that have amplitudes of the same shape but shifted (translated in spacetime); i.e., a weaker process is allowed as compared to the process given by Eq. (1) for the nonrelativistic case:

$$|\varphi_0\rangle \mapsto (U_L|\varphi_0\rangle)(U_L|\varphi_0\rangle),$$
$$|\varphi_1\rangle \mapsto (U_L|\varphi_1\rangle)(U_L|\varphi_1\rangle). \quad (3)$$

Here, U_L is the operator of spacetime translation along the light cone branch for $L = \Delta(x - t)$, i.e., the size of the region where the state amplitude is nonzero (for brevity, we consider that both states are nonzero in the same spacetime region, but they are distinguished by the amplitude shape $\varphi_{0,1}(x - t)$).

The theorem proved in [3] about distinguishing between orthogonal states is modified similarly; i.e., only a weaker process is allowed as compared to the nonrelativistic case (2)

$$|\varphi_0\rangle|A\rangle \mapsto (U_L|\varphi_0\rangle)|A_0\rangle,$$
$$|\varphi_1\rangle|A\rangle \mapsto (U_L|\varphi_1\rangle)|A_1\rangle, \quad |A_0\rangle \neq |A_1\rangle. \quad (4)$$

It is convenient to illustrate the above conclusions by Figs. 1a and 1b. Since the amplitude of states of a massless quantized field that propagate in one direction of the x axis depends only on the difference $x - t$, the time can be fixed and the coordinate can be considered as variable or vice versa. We will do this for both cases. These two cases represent all possible situations. Let one of the orthogonal states that have the amplitude $\varphi(x - t)$ and propagate with the speed of light ($c = 1$, and state subscript 0 or 1 is omitted for brevity) be specified. Let the state be localized in the region L; i.e., $\int_L |\varphi(x - t_0)|^2 dx \approx 1$, where $\varphi_{0,1}(x - t_0)$ is the amplitude at time t_0.

To determine the state amplitude at t_0 for all x values at time t_0 in the region where it is nonzero, it is necessary to carry out the unitary transformation of the entire state. After the unitary transformation of the state amplitude $U\varphi_{0,1}(x - t_0) = \tilde{\varphi}_{0,1}(x' - t)$, the amplitude of the new state $\tilde{\varphi}(x' - t)$ can be nonzero in a smaller spatial region. The minimum size of the region in x' at time t is determined by the relativistic causality principle [10]. The matrix elements of the unitary operator are nonzero only if the points (x, t_0) and (x', t) lie within the past light cone that has the vertex at the point Γ and covers the region where the state amplitude at time t_0 is nonzero. To the time no earlier than L, the initial state can be unitarily transformed to the state with the amplitude localized in an arbitrarily small vicinity of the point Γ. It is fundamentally important that this state differs from the initial state $\varphi(x - t_0)$. To the time Γ, the values of the state amplitude for all x values are immediately (instantaneously) accessible. Then, the measurement outcome can be acquired instantaneously, which provides complete (with unit probability) information about the state. If a pair of the initial states is orthogonal, a pair of orthogonal states can be obtained to the time Γ by the unitary transformation and, therefore, they can be distinguished (in this case, the theorem about reliable distinguishability of orthogonal states [2] can be used). We emphasize again that these states differ from the initial states. A state can also be "reconstructed" or copied by the inverse unitary transformation "directed" to the future. The state with the initial amplitude shape can be acquired no earlier than the time determined by relativistic causality. The state amplitude with the initial shape lies in the future light cone with the vertex at the point Γ. This state also differs from the initial state: it is delayed in time with respect to the initial state, which would propagate in x to time L by the L value if any attempt of copying or acquiring information about it were absent (Fig. 1a). We discuss acquiring information about states in the channel with unit probability. The same consideration is applicable to acquiring information with a probability of less than unity. In this case, the delay is less than L (Figs. 1a, 1b).

The consideration is also applicable to the nonrelativistic case. In this case, argumentation concerning the light cone must be omitted in the above consideration and unitary transformations can be formally instantaneous. Moreover, the explicit presence of coordinates can be excluded from the consideration taking into account only that states are entirely accessible under the unitary transformation (the entire spatial region is instantaneously accessible).

A similar consideration is applicable when the state is unitarily transformed to the state of an auxiliary localized system. Such a unitary transformation is realized when "trapping" light [13]. This unitary transformation transforms the photon-field state to the vacuum state because it is massless and the propagation velocity cannot be equal to zero. This unitary transformation simultaneously transforms the state of the atomic system to a certain new state. Being unitary, the transformation requires access to all the values of the photon-packet amplitude at the point of the localization of the atomic system. This access is naturally achieved as the packet propagates with the speed of light and reaches the localized atomic system (the packet entirely "enters" the atomic system). To acquire the result with unit probability in this process, it also takes time L (single-photon packet must entirely enter into the atomic system). In this case, the photon field is in another (vacuum) state, and the auxiliary system is in a new state depending on the initial photon state. To the time L, it is possible to determine the initial state with unit probability and to prepare the same state but with inevitable delay by L with respect to the free propagation of the initial packet (Fig. 1b).

For further analysis, it is also important that any evolution of a massless quantized field interacting with the environment (other quantum or classic degrees of freedom in the channel) cannot lead to the contraction

of the state in the sense that the normalization of the state is gained in a smaller spatial region falling outside the light cone as compared to free propagation (see Fig. 1c). As a rule, this interaction leads to a mixed state, but the support of the density matrix in spacetime cannot be contracted and removed outside the light cone (Fig. 1c). Otherwise, quantum states would transmit information faster than the speed of light. Indeed, let one of a pair of orthogonal quantum states be known (Fig. 1c). Alice can acquire classical information from the quantum state not earlier than the time determined by the condition that the state amplitude is covered by the past light cone. After that, she can transmit classical information to Bob. Such a transmission cannot be accomplished faster than the speed of light (users are connected by the branch of the light cone, Fig. 1c). Let the quantum state in the channel can be compressed so that the vertex of the past light cone covering the state is in the spacelike region of the light cone whose vertex is located at Alice and one of whose branches passes through Bob. In this case, Bob could extract classical information from the quantum state earlier than it could be transmitted at the speed of light by Alice, because the vertex of the light cone covering the contracted quantum state falls within the spacelike region.

The above analysis for cryptography means that noise in the channel prevents both copying and acquiring information by the eavesdropper earlier than it is determined by the diagrams shown in Figs. 1a and 1b [eavesdropper error under the passage of the time delay test is no less than Eq. (12), see below].

We now describe a cryptosystem based on an optic fiber interferometer (Fig. 2). A pair of orthogonal single-photon states of the form

$$|\varphi_{0,1}\rangle = \int\limits_{\Delta k_{0,1}} d\hat{k}\tilde{\varphi}_{0,1}\delta(\hat{k}^2)\theta(k_0)a^+(\hat{k})|0\rangle = \int\limits_{\Delta k_{0,1}} \frac{dk}{\sqrt{k}}\frac{\tilde{\varphi}(k, k_0 = |k|)}{\sqrt{k}}|k\rangle, \quad (5)$$

where the polarization index immaterial for further discussion is omitted and $\hat{k} = (k, k_0)$, serves as the input states. States are specified in nonoverlapping frequency bands $\Delta k_{0,1}$, which will be considered as coinciding for brevity. We consider the states propagating in the same direction. These states carry information between remote users. We denote $\varphi_{0,1}(k) \equiv \tilde{\varphi}(k, k_0 = |k|)/\sqrt{k}$. It is convenient to represent the states in the coordinate–time representation as

$$|\varphi_{0,1}\rangle = \int\limits_{-\infty}^{\infty} d\tau\varphi_{0,1}(\tau)|\tau\rangle, \quad (6)$$

$$\varphi_{0,1}(\tau) = \int\limits_{\Delta k} dke^{-ik\tau}\varphi(k), \quad |\tau\rangle = \int\limits_{\Delta k}\frac{dk}{\sqrt{k}}e^{ik\tau}|k\rangle, \quad (7)$$

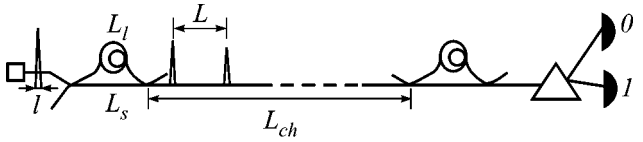


Fig. 2.

where $\tau = x - t$. The amplitude of these states depends only on $\tau = x - t$; i.e., if a measurement outcome takes place at time t in the range $(x, x + dx)$, the same outcome can be acquired at time t' in the range $(x', x' - x + t + dx)$. For brevity, we will say that the amplitudes $\varphi_{0,1}(\tau)$ are specified on a branch of the light cone.

We take states sufficiently localized so that almost complete normalization (arbitrarily close to unity) is gained in the region of size l:

$$\int\limits_{l} d\tau|\varphi_{0,1}(\tau)|^2 \approx 1. \quad (8)$$

The degree of spacetime localization is determined by Δk and $l \approx 1/\Delta k$ (more precise relations see in [9]).

Now, we describe the protocol. Alice randomly chooses one of the states $|\varphi_0\rangle$ and $|\varphi_1\rangle$ in each pulse at a priori known times. It is assumed that the length of the communication channel is known and clocks at both ends are synchronized. The accuracy of synchronization δt must be such that $L \gg \delta t$ ($L = L_l - L_s$ is the difference between long L_l and short L_s arms of the interferometer, Fig. 2). The accuracy of instants of the transmission of states to the communication channel is approximately equal to $\approx l \sim 1/\Delta k$.

The interferometer arm at the transmitter end is in essence necessary for the extension of the short input state with the size $\sim l$ to a longer state consisting of two "halves" at the distance $L \gg l$. This is technically much easier than the preparation of an initially extended state with the "length" L and the correspondingly narrower frequency spectrum. The communication channel after the interferometer arm at the transmitter end has two fiber-optic beam splitters each with one working and blank (vacuum) entry and exit, as well as delay lines in one of the arms (Fig. 2). The state at the working exit in the communication channel has the form (up to the normalization constant and common translation by the arm length)

$$|\varphi_{0,1}\rangle + |\varphi_{0,1}(L)\rangle, \quad (9)$$

where the half of the state is delayed by L:

$$|\varphi_{0,1}(L)\rangle = \int d\tau\varphi_{0,1}(\tau - L)|\tau\rangle. \quad (10)$$

At the receiver end, two halves of the extended state are united together by the inverse unitary transformation realized similarly to the input unitary transformation.

Since the onset time for each pulse, as well as the state length L and communication channel length L_{ch}, is

known, the time of arrival at detectors after uniting two halves at the receiver end is also known. At the receiver end, Bob carries out measurements by a detector with the time constant $\tau_d \ll L$ to detect possible delays with an accuracy better than L. Thus, the detector may not distinguish the arrival times with accuracy l. In other words, times $\approx l$ are considered as indistinguishable (zero).

A prism (or diffraction pattern) at the receiver end is necessary for the separation of the states with different frequency bands for 0 and 1. Moreover, since $l \ll L$, exact balance of the arms between transmitter and receiver ends of the interferometer is not required. In other words, halves of the state at the receiver end may not be exactly united to the state localized in the time window l. It is only necessary that the delay due to the difference between arm lengths at the receiver and transmitter ends is much shorter than $\ll L$.

Further, only measurements in pulses that pass the time delay test retain in the key. In other words, these measurements provide outcomes in the time window $\tau \in (L_s + L_l + L_{ch}) \pm \delta l$ (where the time window δl covers several state lengths l). In fact, only undelayed measurement outcomes such that the first and second halves pass through the long and short arms at the transmitter end, respectively, are retained. At the receiver end, the first and second halves pass through the short and long arms, respectively. For these outcomes, the probability that the eavesdropper knows the bit transmitted by Alice and passes the time delay test is equal to

$$\begin{aligned} &\mathrm{Pt}_E(\mathrm{bit}_E \\ &= \mathrm{bit}_A \wedge \mathrm{test}(\tau \in \tau \in L + L_{ch} + \delta l) = OK) \\ &= \Pr_E(\tau \in \delta l) \cdot 1 \cdot 1 + \Pr_E(\tau \bar{\in} \delta l) \cdot \frac{1}{2} \cdot 1 \\ &\leq \frac{1}{2} \cdot 1 \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} \cdot 1 = \frac{3}{4} < 1. \end{aligned} \tag{11}$$

The first factor in the first term is the probability of the detection of the front half of one of the states. Although this detection leads to the delay $\sim l$, this delay is not detected. This probability is no more than 1/2 due to the localization of the halves of the state at the scale l. If the detection occurs, the states are uniquely identified (second factor), because their frequency bands do not overlap. The third factor is the probability of passing the time delay test at the receiver end and is equal to unity, because the detector time constant is $\tau_d \approx> \delta l$. The first factor in the second term is the probability of the absence of the detection in the time window δl. In this case, the probability of the state identification is equal to 1/2. The third factor is the probability of passing the time delay test and is equal to unity, because the state passes "through" the eavesdropper.

The eavesdropper can increase the probability of identification only by waiting for the second half of the state, i.e., by unitary transformations for the assembly of the state (see the above discussion), which results in a delay of about L detected with probability ~1. Since Bob retains only outcomes passing the time delay test, outcomes with delay $\sim L$ are removed.

Thus, the error probability for the eavesdropper is equal to

$$\begin{aligned} \delta_E \geq 1 - \mathrm{Pt}_E(\mathrm{bit}_E = \mathrm{bit}_A \wedge \mathrm{test}(\tau \in \tau \in L_s \\ + L_l + L_{ch} \pm \delta l) = OK) = \frac{1}{4} \end{aligned} \tag{12}$$

for measurements passed the time delay test. We note that, since states are orthogonal and all events in the protocol occur in real time, it is not necessary to consider collective measurements, because they do not increase the probability given by Eq. (12). It is impossible to determine what is transmitted when the test is passed at the receiver end, because states are quantum and exclusions of Special Relativity exist. Moreover, noise in the channel cannot increase probability (12) due to relativistic exclusions (see the above discussion).

Let the number of pulses passing the test be equal to $2n \gg 1$. Legal users randomly select n positions, open them, compare the bit values (0 or 1) in each position, and estimate the error probabilities. These errors can be induced by noise in the communication channel rather than by the eavesdropper. It is important that the probability given by Eq. (12) cannot be exceeded due to noise. Let δ_{AB} be the error probability estimated from the disclosed part of the sequence. For a sufficiently long sequence, the error probability in the closed part coincides with δ_{AB}.

If $\delta_{AB} < \delta_E$, a random binary code $[n, k]$ with the rate [14, 15]

$$k/n \leq R < C(\delta_{AB}) - \varepsilon, \quad \forall \varepsilon > 0, \tag{13}$$

can be taken so that its error probability is arbitrarily small. However, this code does not correct the error whose probability $\delta_E > \delta_{AB}$. For a sufficiently long sequence ($n \gg 1$), the number of the identical bits enclosed by Alice and Bob is approximately equal to $\approx nC(\delta_{AB})$ after correction, where the transmission capability of the classical symmetric binary communication channel [14, 15] is determined as

$$\begin{aligned} C(\delta_{AB}) &= 1 - H(\delta_{AB}), \\ H(x) &= -x\log x - (1-x)\log(1-x). \end{aligned} \tag{14}$$

In essence, when $\delta_E > \delta_{AB}$, the transmission rate exceeds the transmission capability $C(\delta_E)$ of the channel between Alice and the eavesdropper. The eavesdropper recognizes the correction of errors by codes that are bad for him (i.e., when $\delta_{AB} < \delta_E$) as the transmission of pulses with the rate exceeding the transmission capability of the communication channel between Alice and him.

In this case, when the transmission rate is higher than the transmission capability, we can use the estimate [16]

$$p_E > 1 - 4\frac{\text{const}}{n(C(\delta_{AB}) - C(\delta_E))^2} - \exp\left\{-\frac{n(C(\delta_{AB}) - C(\delta_E))}{2}\right\} \quad (15)$$

of the symbol-error probability for the eavesdropper. This means that the maximum allowable probability of errors in the communication channel is estimated as $\delta_{AB} < 1/4 = 25\%$.

After the correction of errors, the remaining bit sequence of length $\approx nC(\delta_{AB})$ is identical for legal users. The eavesdropper can know no more than $\approx nC(\delta_E)$ bits with an arbitrarily low error probability. The number of secret bits that can be extracted by the legal users from their sequence of the length $\approx nC(\delta_{AB})$ does not exceed $\approx nC(\delta_{AB}) - C(\delta_E))$. Further, using contraction (caching) through the open channel, Alice and Bob can enhance the key security by decreasing the sequence length. When contracting the key, the users can use estimate (15) of the error probability for the eavesdropper in the initial sequence. As a result, the key identical for the legal users arises with the unit probability. The probability that the eavesdropper knows this key is arbitrarily small (after contraction).

The characteristic proper time constant τ_d of the detector must satisfy the inequalities $l < c \geq \tau_d \ll L$. This requirement arises because records for the delay test must be accumulated in the time interval L. Detectors with the time constant $\tau_d \approx 10^{-8}$–10^{-9} s are standard instruments. Therefore, the inequality $l/c < \tau_d$ can be satisfied with a reserve of two orders of magnitude for the input state with the duration (length) $l/c \approx$ 10–100 ps (l = 0.3–3 cm). The value $L \approx 10c\tau_d \approx$ 3–30 m is sufficient for the separation of the halves of the state. In this case, it is sufficient to balance the arms at the receiver and transmitter ends with an accuracy of 3 cm. A real fiber-optic communication channel is not a straight line connecting Alice and Bob. This circumstance imposes a certain extra limit on the separation of the halves of the state. This separation cannot be shorter than $L_{\text{curve}} - L_{ch}$, where L_{curve} is the real length of the optic fiber and L_{ch} is the length of the straight line connecting Alice and Bob. Moreover, since the speed of light c' in the optic fiber is somewhat lower than the speed of light in vacuum ($c' < c$), the effective length of the state cannot be shorter than $c(L_{\text{curve}} - L_{ch})/c'$. We also note that the cryptosystem based on the frequency states must be more stable than systems based on polarization states.

I am grateful to S.S. Nazin for stimulating discussions and critical remarks. This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16289 and project nos. 40.020.1.1.1170 and 37.029.1.1.0031).

REFERENCES

1. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, quant-ph/0101098; Rev. Mod. Phys. **74**, 145 (2002).
2. W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
3. C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
4. D. Mayers and A. Yao, quant-ph/9802025.
5. E. Biham, M. Boyer, P. O. Boykin, *et al.*, quant-ph/9912053.
6. P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).
7. L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239 (1995).
8. S. N. Molotkov and S. S. Nazin, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 767 (2001) [JETP Lett. **73**, 682 (2001)]; S. N. Molotkov, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 79 (2002) [JETP Lett. **76**, 71 (2002)].
9. S. N. Molotkov, Pis'ma Zh. Éksp. Teor. Fiz. (in press).
10. N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, 3rd ed. (Nauka, Moscow, 1973; Wiley, New York, 1980).
11. L. D. Landau and R. Peierls, Z. Phys. **69**, 56 (1931); L. D. Landau, in *Collected Works* (Nauka, Moscow, 1969), Vol. 1, p. 56.
12. N. Bohr and L. Rosenfeld, Math. Fys. Medd. **12**, 3 (1933); in *Collected Scientific Works* (Nauka, Moscow, 1969), Vol. 1, p. 39.
13. M. Fleischauer and M. D. Lukin, Phys. Rev. Lett. **84**, 5094 (2000).
14. C. E. Shannon, Bell Syst. Tech. J. **28**, 658 (1949).
15. E. J. Mac Williams and N. J. A. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1977).
16. J. Wolfowitz, Illinois J. Math. **1**, 591 (1957).

Translated by R. Tyapaev

JETP Letters, Vol. 78, No. 3, 2003, pp. 168–178. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 3, 2003, pp. 201–212.

Kinetics of $L1_2$-Type and $L1_0$-Type Orderings in Alloys

V. G. Vaks[1]

Russian Research Centre Kurchatov Institute, Moscow, 123182 Russia

[1] *e-mail: vaks@mbslab.kiae.ru*

Received June 24, 2003

A review is given of the current concepts of microstructure evolution in $L1_2$-type and $L1_0$-type ordering phase transitions in fcc alloys. Theoretical methods developed for the description of this evolution and the main results obtained with the use of these methods are presented. Theoretical results are compared with the available experimental observations. © *2003 MAIK "Nauka/Interperiodica"*.

PACS numbers: 64.70.Kb; 61.50.Ks; 05.70Fh; 81.30.Hd

1. Introduction. Studies of microstructural evolution under phase transformations (PT) associated with alloy ordering or decomposition (called "diffusive" transformations) are of both fundamental and applied interest. From the fundamental side, the creation and evolution of ordered antiphase domains (APDs) or of precipitates of a new phase is a classical example of the dynamics of nonequilibrium systems and self-organization phenomena, which are intensively studied in many areas of physics and chemistry. From the applied side, the macroscopic properties of alloys, such as strength, ductility, and magnetic properties, generally substantially depend on their microstructure, in particular, on the structure and distribution of antiphase or interphase boundaries (APB or IPB) separating different APDs or different phases, and this microstructure in turn strongly depends on the thermal and mechanical history of the sample, in particular, on the kinetic path of PTs. In this connection, studying the kinetics of phase transformations in alloys is one of the most relevant problems of physical materials science.

The simplest ordered structure corresponds to the B2-type (CuZn-type) ordering. In this case, two cubic sublattices are formed in the A_cB_{1-c} bcc alloy. These two sublattices are enriched respectively in atoms A and B and are displaced with respect to each other by the vector [111]$a/2$, where a is the bcc lattice constant. This ordering is described by one order parameter η proportional to the difference of the occupation probabilities of a particular sublattice with atoms A and atoms B. In this case, there are only two types of APDs differing in the sign of η and only one type of APBs between such domains. Just this simplest case has been considered in most studies of ordering kinetics; see, for example, [1–3]. At the same time, orderings in real alloys are usually much more complicated and involve many types of APDs and APBs. In particular, in the case of $L1_2$- or $L1_0$-type orderings characteristic of fcc alloys, which will be discussed below, there are, respectively, four or six different APDs and several types of APBs. This "multivariant" character of ordering leads to a number of essential differences of their kinetics from simple B2 ordering, including a wide diversity of various evolution scenarios and types of intermediate structures. Many of these structures exhibit peculiar properties that are important for applications and strongly depend on both the alloy composition and the evolution conditions.

Previously, the kinetics of $L1_2$ and $L1_0$ orderings was described theoretically using both direct Monte Carlo simulations (see, for example, [4]) and phenomenological kinetic equations [5–7]. However, direct simulations here are complicated, and until now they have provided little information on the details of evolution. Phenomenological equations are simpler for applications, and Khachaturyan *et al.* [5–7] used them for the description of a number of effects of elastic deformations on PTs. However, phenomenological approaches involve many arbitrary assumptions, which may result in significant distortions of real evolution [8], and the relation of such approaches to a microscopic description is usually not clear. Recently, consistent statistical methods have been developed for the description of nonequilibrium alloys [9–12] and have been applied to studying the kinetics of $L1_2$ and $L1_0$ orderings in the papers [12–15]. The main results of these papers are described below.

2. Basic equations. In this section, we discuss the basic equations of the statistical theory of nonequilibrium alloys [9–12] that are employed for the description of diffusive PTs. For definiteness, we consider a binary alloy A_cB_{1-c} with $c \le 0.5$. Various distributions of atoms over lattice sites i are described by the sets of occupation numbers $\{n_i\}$, where $n_i = n_{Ai}$ equals 1 or 0 if site i is occupied, respectively, by atom A or atom B. The Hamiltonian H has the form

$$H = \sum_{i>j} V_{ij} n_i n_j + \sum_{i>j>k} V_{ijk} n_i n_j n_k + \ldots, \qquad (1)$$

0021-3640/03/7803-0168$24.00

where $v_{i...j}$ are effective interactions. It is convenient to write the general expression for the probability P of finding distribution $\{n_i\}$ as the "generalized Gibbs distribution"

$$P\{n_i\} = \exp\left[\beta\left(\Omega + \sum_i \lambda_i n_i - Q\right)\right];$$
$$Q = \sum_{i>j} a_{ij} n_i n_j + \sum_{i>j>k} a_{ijk} n_i n_j n_k + \ldots \quad (2)$$

Here, $\beta = 1/T$ is the inverse temperature, λ_i and $a_{i...j}$ are parameters of the distribution, and the "generalized thermodynamic potential" Ω is determined from the normalizing condition. As was discussed in [17], "quasi-interactions" $a_{i...j}$ in Eq. (2) for conventional conditions of phase transformations can be considered equal to interactions $v_{i...j}$ in Eq. (1), whereas "chemical potentials of lattice sites" λ_i in the absence of equilibrium are generally not equal to each other. Then, using the master equation for the evolution of probability P in Eq. (2) and the conventional thermal activation model for the probability of atom exchange between lattice sites, we can obtain the following master equation describing the evolution of mean site occupations $\langle n_i\rangle = c_i$ averaged over the distribution (2) [9–12]:

$$dc_i/dt = \sum_j M_{ij} 2\sinh[\beta(\lambda_j - \lambda_i)/2]. \quad (3)$$

Here, the chemical potentials of lattice sites $\lambda_i(c_j)$ are found from the self-consistency conditions

$$c_i = \langle n_i\rangle = \mathrm{Tr}(n_i P\{\lambda_j\}), \quad (4)$$

where Tr(…) means the summation over all sets $\{n_i\}$. In writing the equation for the "generalized mobility" M_{ij} in Eq. (3), we use for simplicity the pair interaction model, when Hamiltonian H in Eq. (1) includes only the first sum, and the model of direct exchange by atoms A and B between neighboring sites i and j (a generalization to a more realistic model of vacancy-mediated exchange is given in [10]). Then M_{ij} is given by the equation [12]

$$M_{ij} = \gamma_{ij}\langle n_i' n_j' \exp[\beta A_{ij}(n_k)/2]\rangle;$$
$$A_{ij} = \lambda_i + \lambda_j - \sum_k (v_{ik} + u_{ik} + v_{jk} + u_{jk}) n_k, \quad (5)$$

where $n_i' = n_{Bi} = (1 - n_i)$, $u_{ij} = V_{ij}^{AA} - V_{ij}^{BB}$ is the "asymmetric potential" introduced in [16], and the factor γ_{ij} is proportional to the probability of exchanging atoms A and B between sites i and j per unit time.

Explicit expressions for $\lambda_i(c_j)$ and $M_{ij}(c_k)$ can be found from Eqs. (4) and (5) with the use of one or another approximate method of statistical physics. The simplest method is the kinetic mean-field approximation (MFA), in which each operator n_i on the right-hand side of Eqs. (4) and (5) is replaced by its mean value $c_i = c_i(t)$. Then, λ_i and M_{ij} take the form

$$\lambda_i^{\mathrm{MFA}} = T\ln(c_i/c_i') + \sum_j v_{ij} c_j; \quad (6)$$

$$M_{ij}^{\mathrm{MFA}} = \gamma_{ij}\left\{c_i c_j c_i' c_j' \exp\left[\beta\sum_k (u_{ik} + u_{jk}) c_k\right]\right\}^{1/2}, \quad (7)$$

where $c_i' = 1 - c_i$.

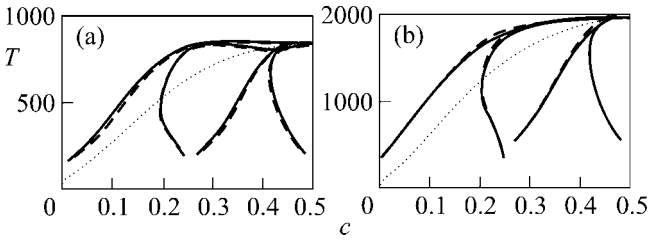


Fig. 1. Equilibrium concentration–temperature (c–T) phase diagrams for models 2 and 4. Solid and dashed lines correspond to calculations within the TCFM and the CVM; dotted line corresponds to the stability limit for the disordered phase in the TCFM [12].

In addition to the simple MFA, more accurate methods can also be used to solve Eqs. (4) and (5). In particular, an essential refinement of the description of B2- and $D0_3$-type orderings in the bcc lattice is reached if the pair cluster approximation (PCA) is used, in which the dependences $\lambda_i(c_j)$ and $M_{ij}(c_k)$ can also be written analytically [17]. However, both the MFA and PCA are inadequate to describe the fcc alloys with the $L1_2$- and $L1_0$-type orderings, because strong many-particle correlations are typical of these alloys and these correlations tend to impede these orderings. These correlations can be adequately described by the cluster-variation method (CVM) with the use of 4-particle (tetrahedral) or larger clusters (see [12] and references therein). However, the CVM is rather cumbersome, and it can hardly be used for the essentially nonuniform systems of our interest. In this connection, a simplified version of the CVM was proposed in [12], namely, the tetrahedron cluster-field method (TCFM), which describes the thermodynamics of $L1_2$ and $L1_0$ orderings for realistic interaction models as accurately as the CVM. This is illustrated by Fig. 1, which presents the concentration–temperature phase diagrams calculated within the TCFM and the CVM for models 2 and 4 discussed below. At the same time, the calculations in the TCFM are much simpler than in the CVM, which makes it possible to apply this method to nonequilibrium systems as well [11–14]. Just as in the MFA and the PCA, the dependence $\lambda_i(c_j)$ in the TCFM can be written explicitly, though, instead of the analytical formulas like Eq. (6), this dependence for each lattice site i is found

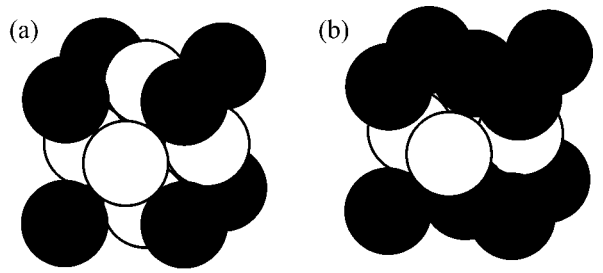


Fig. 2. Phase structures: (a) $L1_2$ and (b) $L1_0$.

from a system of four algebraic equations, which is simply solved by numerical methods. At the same time, the above correlations are insignificant for mobility M_{ij} (Eq. (5)) [12], and the simple expression (7) can be used for this quantity.

Equations (3) can be solved by the methods described above without significant difficulties for systems that contain up to 10^5–10^6 atoms and for times $t \gtrsim (10^3$–$10^4)\gamma_{ij}^{-1}$, which is generally sufficient for studying the main features of microstructural evolution [11–14].

3. Methods and models used. Phase transformation (PT) kinetics after a quench of an alloy from the disordered fcc phase A1 to the phase $L1_2$, to the two-phase region A1+$L1_2$, or to the phase $L1_0$ was studied by computer simulation methods based on Eq. (3) and the TCFM described above [11–14]. A number of models were employed with different interaction radii r_v.

(1) Models 1, 2, and 3 with short-range interaction in which the "reduced interaction" $v_n' = v_n/v_1$ (where v_n corresponds to the interaction of the nth neighbors) is nonzero only for the second neighbors, $v_1 = 1000$ K, and the values of v_2' are, respectively, 0.125, 0.25, and 0.5.

(2) Model 4 with the values of v_n estimated in [18] from the experimental data for Ni–Al alloys: $v_1 = 1680$, $v_2 = -210$, $v_3 = 35$, and $v_4 = -207$ K. These values correspond to an "intermediate" value of r_v.

(3) Model (5) with an "extended" interaction: $v_1 = 1000$ K, $v_2' = -0.5$, $v_3' = 0.25$, and $v_4' = -0.125$.

In studying A1 $\longrightarrow$ A1+$L1_2$ and A1 $\longrightarrow$ A1+$L1_0$ PTs, models 1'–5' and 1"–5" were also considered. In these models, the "deformational" or "elastic" interactions related to the local lattice distortions around various atoms are added to the "chemical" interactions v_n in models 1–5 discussed above. These interactions were estimated from the data for Ni–Al and Co–Pt alloys, respectively, as was described in the papers [12] and [14].

The structures of the $L1_2$ and $L1_0$ phases are shown in Fig. 2. The occupations c_i of fcc lattice sites $\mathbf{R}_i$ in these phases are described by three order parameters η_α corresponding to three superlattice vectors $\mathbf{k}_\alpha$:

$$c_i = c + \eta_1 e^{i\mathbf{k}_1\mathbf{R}_i} + \eta_2 e^{i\mathbf{k}_2\mathbf{R}_i} + \eta_3 e^{i\mathbf{k}_3\mathbf{R}_i};$$
$$\mathbf{k}_1 = (100)\frac{2\pi}{a}, \quad \mathbf{k}_2 = (010)\frac{2\pi}{a}, \quad \mathbf{k}_3 = (001)\frac{2\pi}{a}, \tag{8}$$

where a is the fcc lattice constant. In the homogeneous $L1_2$ phase, the parameters η_α obey the equations $|\eta_1| = |\eta_2| = |\eta_3| = \eta_0$, where η_0 is the equilibrium value of $|\eta_\alpha|$ and $\eta_1\eta_2\eta_3 > 0$, and so four types of ordered domains are possible. One of them is displayed in Fig. 2a, and three others are obtained from it by displacing the sublattice of minority atoms (dark) by vector (011)a/2, (101)a/2, or (110)a/2. In the $L1_0$ phase with the tetragonal axis α, only one parameter $\eta_\alpha = \pm\eta_0$ is nonzero; thus, six types of ordered domains are possible, two domains for each of the three directions α.

The partially ordered states under consideration can be conveniently described with the use of the "local" order parameters $\eta_{\alpha i}$ and the concentrations $\bar{c}_i$ that correspond to spatial averaging over a certain region. Below, we use the parameters $\eta_{\alpha i}^2$ and $\bar{c}_i$ averaged over the nearest neighbors $j_{nn}(i)$ of each lattice site i and the quantities η_i^2 that characterize the total degree of the local order:

$$\eta_{\alpha i}^2 = \frac{1}{16}\left(c_i + \frac{1}{4}\sum_{j_{nn}(i)} c_j e^{i\mathbf{k}_\alpha\mathbf{R}_{ji}}\right)^2;$$
$$\bar{c}_i = \frac{1}{4}\left(c_i + \frac{1}{4}\sum_{j_{nn}(i)} c_j\right); \quad \eta_i^2 = (\eta_{1i}^2 + \eta_{2i}^2 + \eta_{3i}^2), \tag{9}$$

where $\mathbf{R}_{ji}$ is $\mathbf{R}_j - \mathbf{R}_i$. It was shown in [12] that the distribution of η_i^2 values is close to the intensity distribution observed in transmission electron microscopy (TEM) experiments [19, 20]. Therefore, the results of simulations in the figures below are usually presented in the "η^2 representation" in which the gray level varies linearly with η^2 between its minimum and maximum values from completely dark to completely bright.

Our simulations were performed in fcc simulation boxes with the volume $V_b = L^2 \times H$ (where L and H below are given in units of the lattice constants a) with periodic boundary conditions. Both three-dimensional (3D) simulation with $H = L$ and quasi-2D simulation with $H = 1$ were used. Below, we mainly present the results of quasi-2D simulations, in which more sizable structures can be studied. However, these results were verified and supplemented with 3D simulations in the original papers [12–14].

4. Kinetics of $L1_2$-type ordering. In the consideration of the $L1_2$-type orderings, the following problems will be discussed.

(1) The dependence of microstructural evolution under the A1 ⟶ $L1_2$ PT on the type of interactions V_{ij} as well as on the concentration c and temperature T of an alloy.

(2) The effect of elastic interactions on the kinetics of decomposition with ordering of the A1 ⟶ A1+$L1_2$ type.

In studying these problems in the papers [11–13], it was found that the character of the microstructural evolution strongly depends on the type of interaction V_{ij} (especially, on its effective radius r_V) as well as on the degree of nonstoichiometry $\delta c = (c - 0.25)$ and temperature T. An increase in the radius r_V, nonstoichiometry δc, and temperature T makes the microstructures more isotropic and makes the APBs more diffuse and mobile. At the same time, for the short-range-interaction alloys at small δc and not high T, the microstructures are highly anisotropic and the APBs are thin and low-mobile.

These structural features are illustrated by Figs. 3–8. Figure 3 demonstrates the evolution of the domain structure under the A1 ⟶ $L1_2$ PT for model 4 with an "intermediate" interaction range and a nonstoichiometric composition $c = 0.22$. It is evident that the distribution of APBs is entirely isotropic. The main evolution mechanism is the growth of larger domains at the expense of smaller ones by the motion of APBs, which is also typical of the simple B2 ordering, and only this mechanism was discussed previously [1–3]. At the same time, Fig. 3 shows that another mechanism, the fusion of in-phase domains, which is absent for the simple B2 orderings with only two types of domains, is important for the multivariant orderings under consideration. It is seen that there are two types of such processes: (a) the splitting of APBs between two APDs that separate the in-phase domains to be fused and (b) the disappearance of an intermediate domain. Examples of processes of type (a) are seen in the lower half of frames 3b–3e, and two processes of type (b) are seen on the left-hand upper part of frames 3b–3d.

Another characteristic feature of $L1_2$ orderings with isotropic APBs is that approximately equiangular triple junctions with angles ≃120° between the adjacent APBs occur at later stages of evolution. A comparison with the TEM observations shown in Fig. 4 (and explained below in the discussion of Fig. 10) shows that this conclusion agrees with the experiments.

Figure 5 shows the evolution of the same model as in Fig. 3 but with $c = 0.25$. It is evident that the microstructures notably change when one passes to the stoichiometric composition even though the change in concentration is quite small. The distribution of APBs reveals some anisotropy and a tendency to the formation of (100)-oriented APBs (which are called "conser-

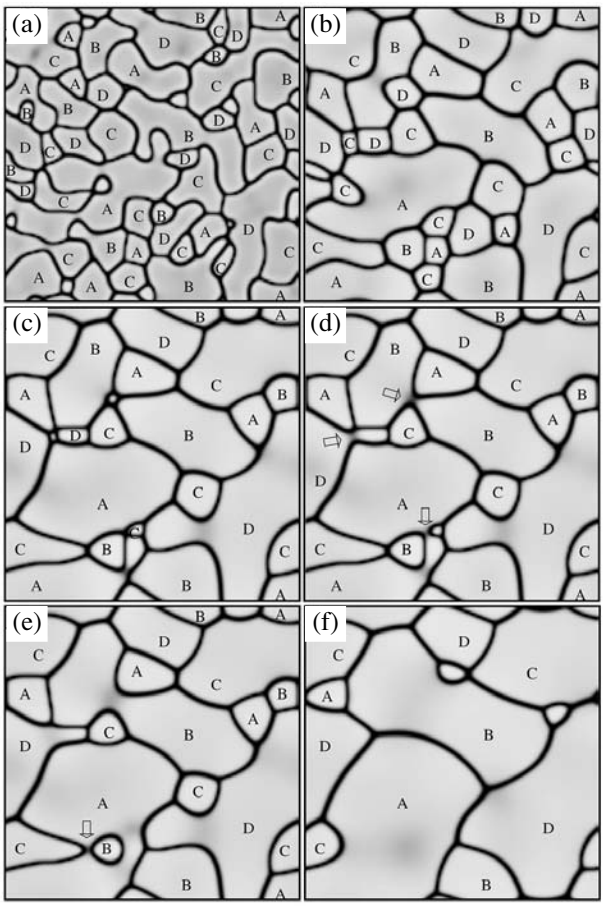


Fig. 3. Temporal evolution of model 4 under the A1⟶ $L1_2$ PT shown in the η^2-representation for simulations in the volume $V_b = 128^2 \times 1$ at $c = 0.22$, $T = 1150$ (from here on in K) and the following values of the reduced time $t' = \gamma_{ij}t$: (a) 5, (b) 50, (c) 120, (d) 125, (e) 140, and (f) 250. Symbols A, B, C, and D indicate the type of ordered domains, and thick arrows indicate the processes of their fusion.

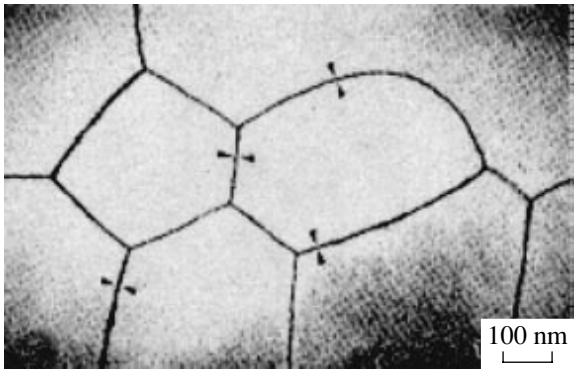


Fig. 4. TEM image of APBs observed in the $Cu_{0.83}Pd_{0.17}$ alloy under the A1 ⟶ A1+$L1_2$ PT at the stage of congruent ordering and then wetted by the disordered phase [19].

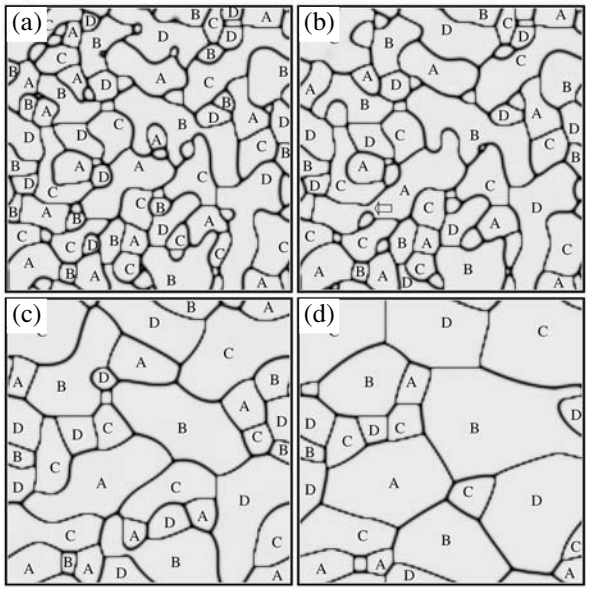


Fig. 5. As Fig. 3, but at $c = 0.25$ and $t' =$ (a) 5, (b) 10, (c) 50, and (d) 250.

vative" APBs and are discussed below). The nonconservative APBs have a tendency to "faceting" and to the formation of steplike structures, in which the energy gain for the (100)-oriented segments exceeds the energy loss due to an increase in the total length of such an APB.

The kinetics of $L1_2$ ordering in alloys with short-range interaction is illustrated in Fig. 6. It is evident that here the microstructures are strongly anisotropic and the conservative APBs mentioned above predominate. In the case of nonzero interaction between only the nearest neighbors ($V_{n>1} = 0$), the energy of such APBs equals zero [21]. Therefore, at small $|V_2/V_1| \simeq 0.1$ considered here, the energy of such APBs is small too and their predominance in the structure is natural. The width of these APBs (which at $V_{n>1}$ is just one atomic layer [21]) is notably smaller than the width of the nonconservative APBs.

Figure 6 also illustrates other structural features of the $L1_2$ ordered alloys with short-range interaction: the "steplike" APBs with the conservative segments mentioned above; the triple junctions of APDs with one nonconservative APB and two conservative APBs normal to each other; the "quadruple" junctions of APDs (for example, on the left-hand side and in the lower part of frames 6b–6f), in which the nonconservative part of two triple junctions has the atomic-order length; and so on. All these features have been observed in the TEM

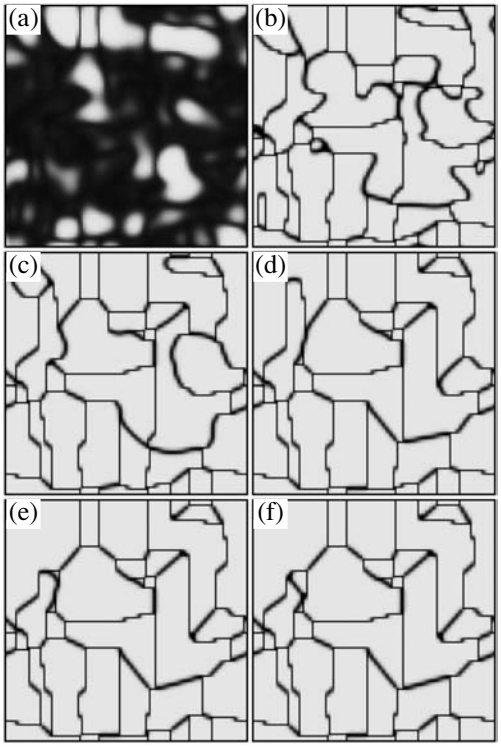


Fig. 6. As Fig. 5, but for model 1 with $V_2' = -0.125$ at $V_b = 64^2 \times 1$, $T = 350$, and $t' =$ (a) 2, (b) 3, (c) 20, (d) 100, (e) 177, and (f) 350.

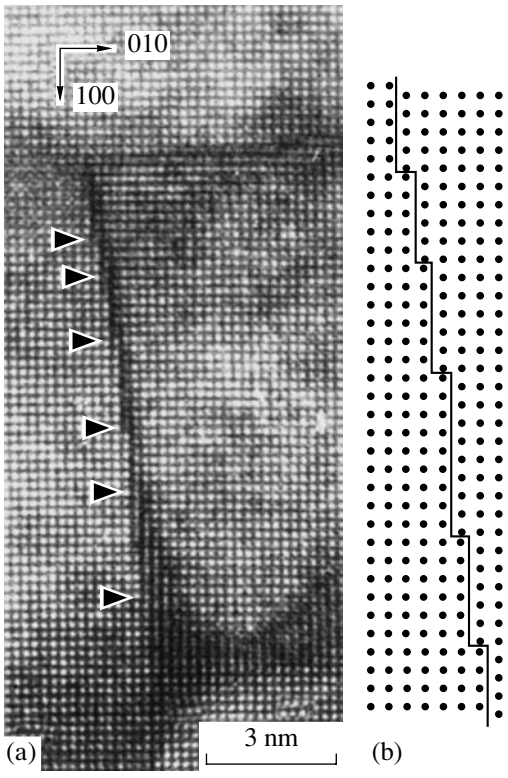


Fig. 7. (a) High-resolution electron microscopy (HREM) image of an APB in the Cu_3Au alloy; (b) schematic view of the atomic structure of this APB.

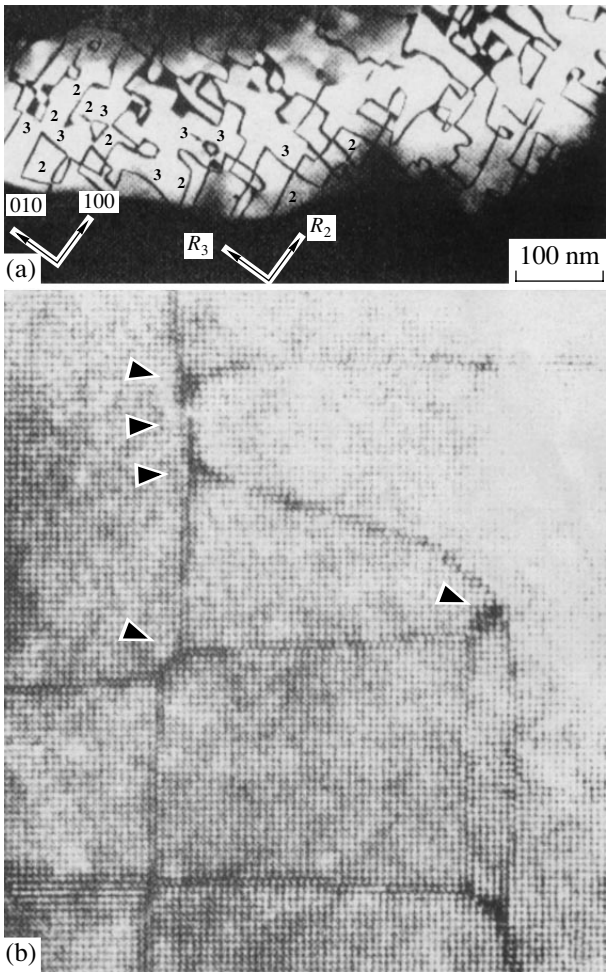


Fig. 8. (a) TEM image of the Cu_3Au alloy showing an APB with the displacement vectors $\mathbf{R}_2 = [101]a/2$ and $\mathbf{R}_3 = [011]a/2$; (b) HREM image of some APBs in the Cu_3Au alloy [19].

Fig. 9. Metallographic image of a two-phase A1+$L1_2$ structure formed by cuboids of the $L1_2$ phase with a size of about 0.2 μm in the disordered A1 phase of a Ni–Al type superalloy [22].

and HREM studies of a Cu_3Au alloy [19, 20] (see Figs. 7 and 8).

The smallness of the energy of conservative APBs makes them extremely low-mobile. Therefore, the evolution and growth of domains is realized by the motion of the nonconservative APBs. Figure 6 illustrates the peculiarity of the processes that take place in this case. Thus, the processes of "sweeping" of a pair of vertical conservative APBs by a moving nonconservative APB are seen in the left-hand lower part of frame 6b and in the left-hand upper part of frame 6d. The process of "wetting" of a conservative APB with the adjacent nonconservative APBs is seen in the left-hand upper corner of frames 6b and 6c. After that, the domain bounded by the nonconservative APBs rapidly "collapses." An unusual process of "splitting" of a nonconservative APB into two conservative and one nonconservative APBs with the formation of a new APD is seen to the left of and higher than the center of frames 6d–6f, and so on. All these peculiar kinetic effects, being predicted by the theory, are due to the smallness of the energy of the conservative APBs in the alloys under consideration.

Consider now the ordering PTs under alloy decomposition of the A1 $\longrightarrow$ A1+$L1_2$ type. In the two-phase region forming in this case, each $L1_2$-ordered domain is surrounded with the disordered A1 phase and the possible anisotropy of the structure is determined by the anisotropy of the IPBs. It was shown in the papers [21, 12] that, in the absence of long-range elastic forces, that is, with only "chemical" interactions, the IPB energies are virtually isotropic for all the models, including the models with short-range interaction (as distinct from the APB energies discussed above). However, when the misfit $\varepsilon = |a_{A1} - a_{L1_2}|/a_{A1}$ between the lattice constants of the A1 and $L1_2$ phases is notable, elastic forces arise in the vicinity of the IPB and tend to orient the IPB along the "elastically soft" (100) directions. The magnitude of these elastic forces grows with increasing size l of the precipitates. Therefore, the elastic forces grow with increasing l in the course of the evolution, and the anisotropy of the IPBs becomes high even at relatively small values of $\varepsilon \lesssim 0.01$. This leads to the formation of peculiar structures with specific properties, in particu-

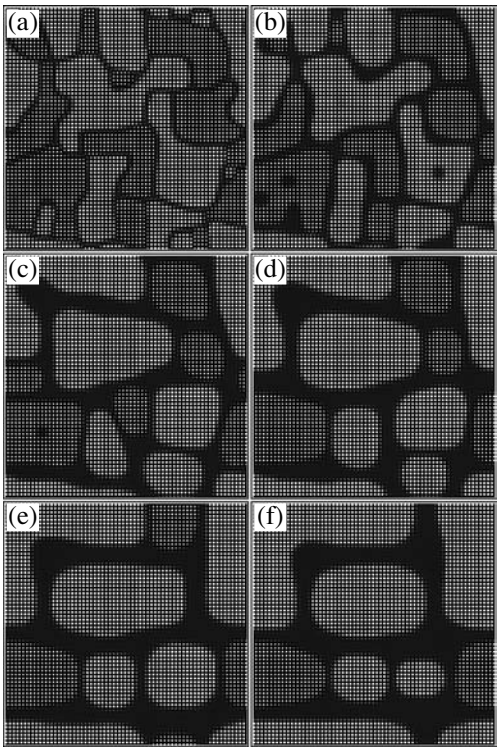


Fig. 10. Temporal evolution of model 2' under the A1 ⟶ A1+$L1_2$ PT for $V_b = 64^2 \times 1$, $c = 0.17$, $T = 400$, and t' = (a) 10, (b) 100, (c) 500, (d) 1000, (e) 2000, and (f) 4200. The gray level varies linearly with c_i between its minimum and maximum values from completely dark to completely bright.

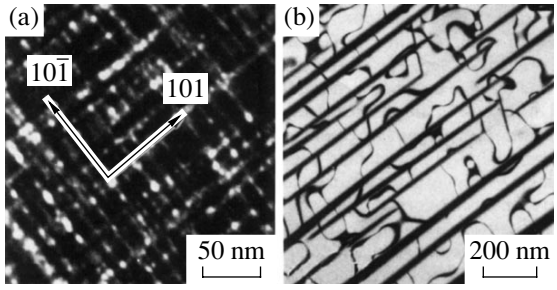


Fig. 11. TEM image of the FePd alloy under the A1 ⟶ $L1_0$ PT and an anneal at $T < T_c$ for (a) 3 h and (b) 61 h [5]. Light areas correspond to domains with one of the directions of the *c* axis (type A and $\bar{A}$ or B and $\bar{B}$); dark areas, to all other areas.

lar, "superalloys," whose structure is illustrated in Fig. 9. These alloys exhibit outstanding strength and refractory properties and are widely used in the aerospace industry [22].

A phenomenological theory of the formation of structures of this type was developed by Khachaturyan *et al.* [6, 7]. The microscopic approach proposed in [11, 12] enables one to specify this theory and to explain a number of observed effects. As an illustration, Fig. 10 presents the results of simulation of the A1 ⟶ A1+$L1_2$ PT for model 2'. Frame 10a shows the initial stage of the "congruent" (that is, proceeding without changing the initial concentration) ordering as well as the beginning of wetting of the as-formed APBs, which gradually transforms these APBs into IPBs. Frame 10b shows the next stage, in the course of which the starting APDs transform into the ordered precipitates, and the excess majority atoms (dark) diffuse toward the surrounding IPBs. Later on, the large precipitates start to grow at the expense of the "evaporation" of smaller ones (frames 10c–10f), and the above effects of the alignment of the IPBs along the (100) directions become pronounced. At the same time, in spite of an extreme simplicity and the small size of the model, the simulation reproduces and explains a number of structural features typical of real superalloys. Thus, "funnels" like those shown in the left-hand upper corner of frames 10c and 10d are seen in Fig. 9 at its left-hand edge. Γ-shaped ledges like those shown at the left-hand upper corner of frame 10f are seen in Fig. 9 at its upper edge and to the left of its center. "Chains" of precipitates like those shown in the lower third of frames 10e and 10f are seen at many places in Fig. 9, and so on.

5. Kinetics of $L1_0$-type ordering. The $L1_0$ structure, as distinct from the $L1_2$ structure, has a distinct axis normal to the alternating planes of different atoms in Fig. 2b, which is called below the axis *c*. As the sizes and interactions of different atoms are different, a tetragonal distortion ε exists along axis *c*. In addition, the tetragonal structure of APDs in the $L1_0$ phase leads to the occurrence of two different types of APBs: shift-APBs separating APDs with the same *c* axis and flip-APBs separating APDs with the perpendicular *c* axes.

Depending on the importance of the distortion ε for the kinetics of ordering, the evolution in the course of the A1 ⟶ $L1_0$ PT can be divided into three stages.

I. The initial stage of the formation of the finest $L1_0$-ordered domains, when their tetragonal distortion only slightly affects the evolution and all six types of domains are approximately equally present in each microstructure.

II. The next, intermediate, stage, which corresponds to the TEM images of the tweed type illustrated in Fig. 11a. The tetragonal distortion of APDs here leads to the predominant orientation of flip-APBs along the (110)-type directions. The number of APDs with the unfavorable (001) orientation of the *c* axis in each of these locally oriented regions is decreased but is still comparable with the number of APDs with the "favorable" (100) and (010) orientations.

III. The final, "twin," stage, when the tetragonal distortion of APDs becomes the main factor of the evolution and leads to the formation of "twin" bands along the (110)-type directions (see Fig. 11b). Each band includes only two types of APBs with the same *c* axis,

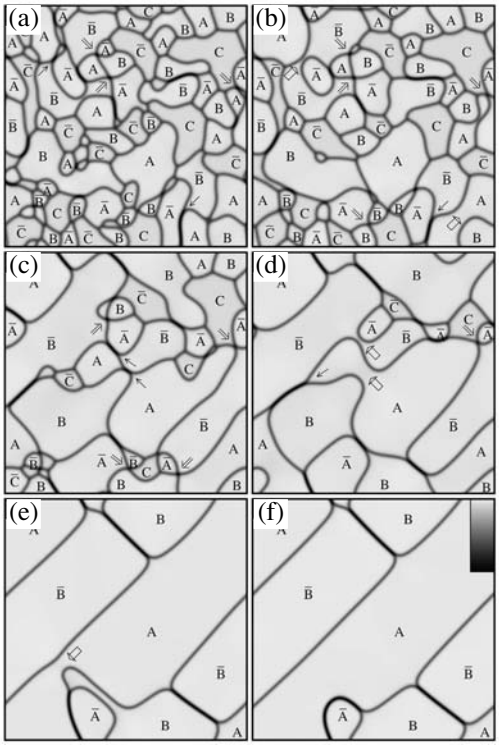


Fig. 12. Evolution of model 5" under the A1 ⟶ $L1_0$ PT for $V_b = 128^2 \times 1$, $c = 0.5$, reduced temperature $T' = T/T_c = 0.7$, maximum distortion $|\varepsilon_m| = 0.1$, and t' = (a) 10, (b) 20, (c) 50, (d) 100, (e) 250, and (f) 280. The inset shows the gray level upon varying η_i^2 from zero up to $\eta_{\max}^2 \simeq 0.20$. The symbol A, $\bar{A}$, B, $\bar{B}$, C, or $\bar{C}$ designates the type of ordered domain, and a thick, thin, or ordinary arrow indicates, respectively, the process of domain fusion, a junction of four APDs, or the process of APB splitting.

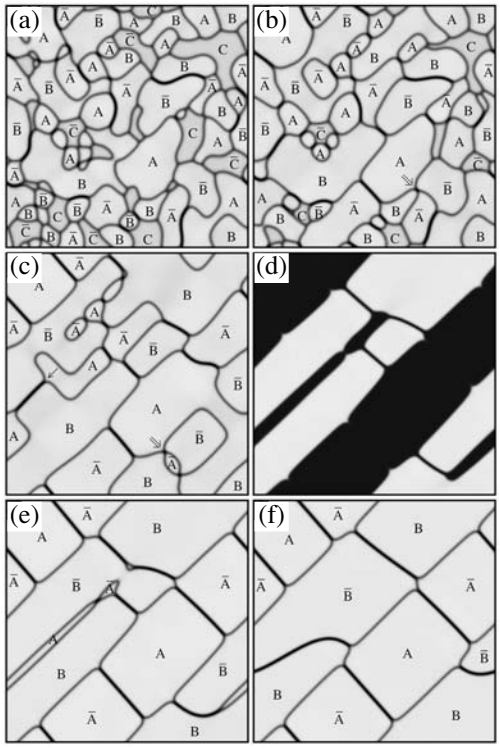


Fig. 13. As Fig. 12, but at $|\varepsilon_m| = 0.15$, $\eta_{\max}^2 \simeq 0.21$, and t' = (a) 10, (b) 20, (c) 50, (d) 150, (e) 172, and (f) 350. In frame (d), the gray level varies linearly with η_{2i}^2 when it is varied from zero to $\eta_{2\max}^2 \simeq 0.21$ from completely dark to completely bright.

and the (100) and (010) orientations of this axis alternate in the adjacent bands.

The physical reason for the (110) orientation of flip-APBs is the gain in elastic energy for the adjacent domains: at other orientations of APBs, this energy increases with the growth of an APD proportionally to its volume [5, 15]. For an APD with the characteristic size l, the surface S_d, and the shear modulus c_s, the elastic energy $E_{el} \sim c_s\varepsilon^2 l S_d$ begins to affect the evolution only when it becomes comparable with the surface energy $E_s \sim \sigma S_d$, where σ is the APB surface tension. The tweed stage II corresponds to the relation $E_{el} \sim E_s$ or to the characteristic size of the APD

$$l_0 \sim \sigma/c_s\varepsilon^2, \tag{10}$$

which sharply increases with increasing distortion ε.

Some results of simulations of the A1 ⟶ $L1_0$ PT are displayed in Figs. 12–17 [14]. The symbols A or $\bar{A}$, B or $\bar{B}$, and C or $\bar{C}$ in these figures indicate APDs with the c axis directed along the principal (100), (010), and (001) crystal axis and with a positive or negative value, respectively, of the relevant order parameter η_α. Temperature T is given in terms of the reduced values $T' = T/T_c$, where T_c is the critical (maximum) temperature of the $L1_0$ ordering. For models 1", 2", 3", 4", and 5", this T_c equals, respectively, 614, 840, 1290, 1950, and 2280 K. The local distortion ε is proportional to the local order parameter squared [14], and the scale of ε is characterized by its maximum equilibrium value $|\varepsilon_m|$, which corresponds to $T = 0$ and $c = 0.5$.

Figures 12–17 illustrate quasi-2D simulations for which microstructures include only edge-on APBs normal to the (001) plane in the figures. In this geometry, the bulk elastic energy can vanish only for the (100)- or (010)-oriented domains A and $\bar{A}$ or B and $\bar{B}$, separated by a flip-APB with the (110) orientation (or the (1, –1, 0) one, which is not mentioned below for brevity). On the other hand, in the (001)-oriented domains C and $\bar{C}$, this elastic energy is always present. Therefore, the tweed and twin stages in our simulations correspond to

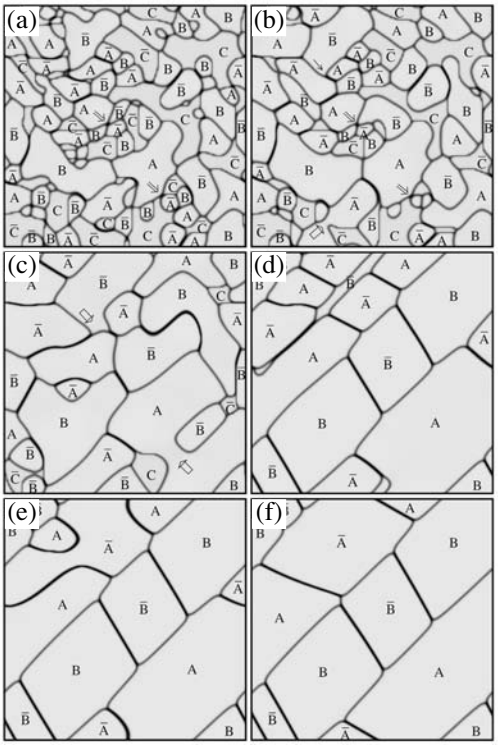


Fig. 14. As Fig. 12 but for model 4" at $|\varepsilon_m| = 0.15$, $T' = 0.67$, $\eta^2_{max} \simeq 0.24$, and t' = (a) 10, (b) 20, (c) 50, (d) 170, (e) 200, and (f) 700.

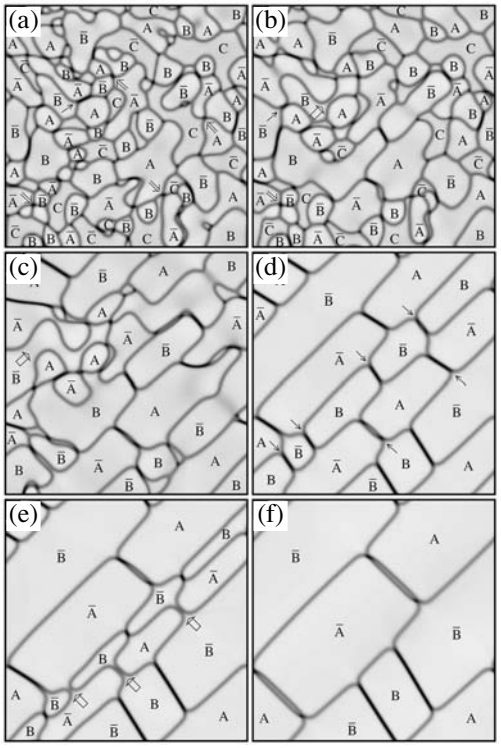


Fig. 15. As Fig. 14, but at $c = 0.44$, $\eta^2_{max} \simeq 0.20$, and t' = (a) 10, (b) 20, (c) 50, (d) 400, (e) 750, and (f) 1100.

both the predominance of (110)-oriented APBs and the decrease in the portion of C and $\bar{C}$ domains.

First we discuss the evolution for the systems in which the interaction range is not small (see Figs. 12–15). Here, frames 12a–12b, 13a, 14a, and 15a correspond to the initial stage; frames 12c–12d, 13b–13c, 14b–14c, and 15b correspond to the tweed stage; and the others correspond to the twin stage. It is evident that the following features are characteristic of both the initial and the tweed stages.

(a) The presence of abundant processes of fusion of in-phase domains, which are among the main mechanisms of evolution at this stage.

(b) The presence of peculiar long-living configurations, the junctions of four different APDs (4-junctions) of the type $A_1A_2\bar{A}_1A_3$, where A_2 and A_3 can correspond to any two of four types of APDs differing from A_1 and $\bar{A}_1$.

(c) The presence of many processes of "splitting" of a shift-APB which lead to either the fusion of in-phase domains ($s \longrightarrow f$ process) or the formation of a 4-junction ($s \longrightarrow 4j$ process).

Thus, for example, the $s \longrightarrow f$ processes can be followed in frames 12a–12b, 12c–12d, 12d–12e, 13c–13d, etc. A fusion with the disappearance of an intermediate domain is seen in the right-hand lower part of frames 12a–12b. Several long-lived 4-junctions are seen in frames 12a–12d and 13c–13d, whereas the $s \longrightarrow 4j$ process can be followed in the right-hand lower part of frames 12a–12c. Note that the microstructural features (b) and (c) are naturally explained by the fact that in this case the surface tension σ for the shift-APBs substantially exceeds σ for the flip-APBs [14].

Frames 12c–12d, 13b–13c, 14b–14c, and 15b (as well as 16a–16b) illustrate the processes of the (110)-type alignment of flip-APBs and the "dying out" of APDs of the C and $\bar{C}$ types, which are characteristic of the tweed stage. Frames 12c–12d also indicate that the characteristic size of APDs in the tweed stage is about $l_0 \sim (20–40)a$ in the simulations with a realistic value of the distortion $|\varepsilon_m| = 0.1$ (estimated from the data on the deformation of the CoPt alloy). This agrees with the order of magnitude of l_0 observed in the FePd and FePt alloys, which are structural analogues of CoPt (see Fig. 11a).

In the discussion of the final, twin, stage of evolution, note that, as is evident in Figs. 11–16, this stage can also be divided into two stages: the initial stage III(a), during which the twin bands still include many curved shift-APBs and small twin bands ("microtwins"), and the final, "quasi-equilibrium," stage III(b), when the APBs are mainly linear and microtwins are no longer present. In comparison with the experiment, it must be taken into account that, because of the relative smallness of the simulation volume, the width of the

twin bands d in the simulations has the same order of magnitude as the size of domains at the tweed stage l_0. However, in the experiments, usually $d \gg l_0$ (see, for example, Fig. 11). Therefore, the distribution of APBs in the simulations of stage III(a) is considerably closer to an equilibrium one than in standard TEM experiments. In spite of that, the simulation reproduces and explains some features of this stage observed in Fig. 11b. Thus, characteristic semiloop-like and S-shaped configurations in Fig. 11b correspond to the shift-APBs that originated from approximately equiaxed domains, which are typical of the beginning of the twin stage, because of the disappearance of the adjacent APDs "wrongly" oriented with respect to the particular twin band. This is seen, for example, in frames 12d–12f, 13e–13f, 14c–14e, and others. The formation of narrow and short microtwins, whose ends are usually adjoined with shift-APBs, is illustrated by frame 14d and, especially, by frame 13d. This frame is given in the "η_2^2 representation" similar to that used in the TEM observations [19]. The microtwin displayed in this frame is very similar to those presented at the center of Fig. 11b.

The final, quasi-equilibrium, stage of evolution is illustrated by the last frames in Figs. 13–16 and Fig. 17. In particular, these frames demonstrate the peculiar phenomenon of the alignment of shift-APBs: within each twin band oriented along the (110) direction and containing the (100)-oriented domains, the APBs tend to align normally to a certain direction $\mathbf{n}_a = (\cos\alpha, \sin\alpha, 0)$ with a certain angle α, so that these directions alternate in the neighboring twin bands. The tilting angle α strongly depends on the type of chemical interaction, especially on its radius r_v, as well as on the concentration and temperature. Thus, for model 5" with a large r_v, the angle α is close to $\pi/4$; for the "realistic" model 4" with an intermediate r_v, the angle $\alpha < \pi/4$; that is, the plane of the APBs is tilted to the tetragonal axis; and $\alpha \simeq 0$ for models with short-range interaction; that is, the APBs are oriented along the tetragonal axis. This effect is naturally explained [15] by the competition between the anisotropy of the surface energy (which tends to orient APBs along the tetragonal axis, that is, to decrease α in the systems with small and intermediate r_v) and the tendency to minimize the total area of the APBs in the given twin band, which corresponds to $\alpha = \pi/4$. The phenomenon of the alignment of APBs within twin bands indicated in [15] was observed in the CuAu alloys, where interactions are short-ranged and $\alpha \simeq 0$ [23], and in the Co–Pt alloys, where the angles α are close to those shown in frames 14f and 15f [15, 24]. Let us also note that the phenomenon of wetting and splitting of shift-APBs by the adjacent twin bands predicted in [14] and illustrated by frames 15d–15f is also confirmed by recent TEM observations for the $Co_{0.4}Pt_{0.6}$ alloy [24].

Figure 16 illustrates the evolution for the model with short-range interaction as in Cu–Au alloys. It is

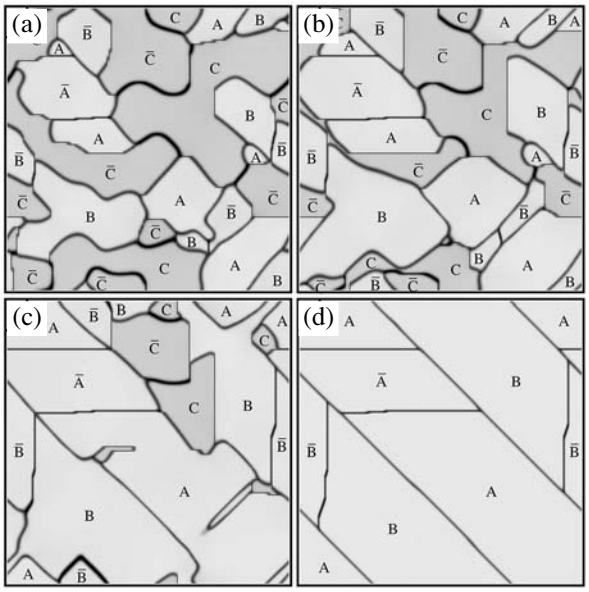


Fig. 16. As Fig. 14, but for model 1" at $T' = 0.9$ and $t' =$ (a) 30, (b) 40, (c) 60, (d) 120.

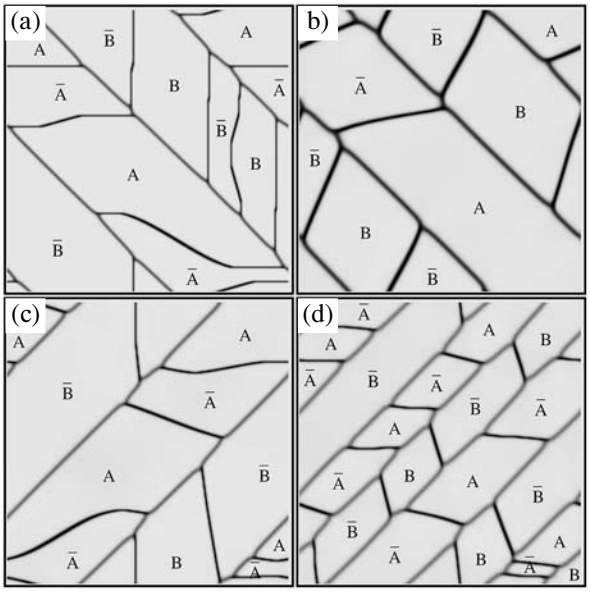


Fig. 17. As Fig. 12, but for model 2" at the following sets of values (c, T', $\eta^2_{\max}$, t'): (a) (0.5, 0.77, 0.24, 350), (b) (0.5, 0.95, 0.21, 300), (c) (0.46, 0.77, 0.20, 350), and (d) (0.44, 0.77, 0.19, 300).

evident that the general character of the evolution is similar to that considered above. However, the microstructures contain many conservative APBs, which can be of two types in the $L1_0$ phase: the flip-type and the shift-type [13]. Their presence leads to features similar to those discussed in connection with Figs. 6–8. Thus, all the shift-APBs in the final frame 16d are steplike,

and these APBs can be considered as "faceted" versions of tilted APBs in frames 14f, 15f, 17c, and 17d. With increasing temperature T or nonstoichiometry $\delta c = (0.5 - c)$, the anisotropy of the APBs rapidly drops. This leads to a sharp change in the morphology of the aligned APBs that can be described as the "faceting–tilting" phase transformation, as illustrated by frames 16d and 17a–17d. These morphological changes are realized via the local bends of faceted APBs, which are seen, in particular, in frames 17a and 17b. Therefore, this "morphological phase transition" is actually smeared over a certain interval of temperature or concentration. However, frames 17a–17d show that these "intervals of PT smearing" can be rather narrow.

6. Conclusions. In this review, I tried to show that the kinetics of phase transformations in alloys is a field of solid-state physics that is not only important practically but also rather interesting and rich in phenomena. Because of the limitations of the review volume, I could discuss only certain structural aspects of the kinetics of $L1_2$ and $L1_0$ orderings. At the same time, there are many other important problems here such as the structure and energy of antiphase and interphase boundaries at various temperatures and concentrations, nucleation of new phases in metastable states, fluctuation phenomena in the kinetics of transformations, and others. At present, all these problems are being intensively studied, and many new interesting results can be expected here.

I am deeply grateful to I.R. Pankratov, who made a great contribution to both the preparation of this review and the works described here; to K.D. Belashchenko, V.Yu. Dobretsov, and G.D. Samolyuk, who participated in these works; and also to G. Martin for valuable discussions.

This work was supported by the Russian Foundation for Basic Research, project nos. 00-02-17692, 00-15-96709, and 03-02-17204a.

REFERENCES

1. S. M. Allen and J. W. Cahn, Acta Metall. **27**, 1085 (1979).
2. V. Yu. Dobretsov, G. Martin, F. Soisson, and V. G. Vaks, Europhys. Lett. **31**, 417 (1995).
3. V. I. Gorentsveig, P. Fratzl, and J. L. Lebowitz, Phys. Rev. B **55**, 2912 (1997).
4. C. Frontera, E. Vives, T. Castan, and A. Planes, Phys. Rev. B **55**, 212 (1997).
5. L.-Q. Chen, Y. Z. Wang, and A. G. Khachaturyan, Philos. Mag. Lett. **65**, 15 (1992).
6. Y. Wang, D. Banerjee, C. C. Su, and A. G. Khachaturyan, Acta Mater. **46**, 2983 (1998).
7. A. G. Khachaturyan, in *Phase Transformations and Evolution in Materials*, Ed. by P. E. A. Turchi and A. Gonis (TMS, Warrendale, 2000), p. 3.
8. V. G. Vaks, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 105 (2002) [JETP Lett. **76**, 93 (2002)].
9. V. G. Vaks, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 447 (1996) [JETP Lett. **63**, 471 (1996)].
10. K. D. Belashchenko and V. G. Vaks, J. Phys.: Condens. Matter **10**, 1965 (1998).
11. K. D. Belashchenko, V. Yu. Dobretsov, G. D. Samolyuk, and V. G. Vaks, in *Phase Transformations and Evolution in Materials*, Ed. by P. E. A. Turchi and A. Gonis (TMS, Warrendale, 2000), p. 139.
12. K. D. Belashchenko, V. Yu. Dobretsov, I. R. Pankratov, *et al.*, J. Phys.: Condens. Matter **11**, 10593 (1999).
13. I. R. Pankratov and V. G. Vaks, J. Phys.: Condens. Matter **13**, 6031 (2001).
14. K. D. Belashchenko, I. R. Pankratov, G. D. Samolyuk, and V. G. Vaks, J. Phys.: Condens. Matter **14**, 565 (2002).
15. V. G. Vaks, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 274 (2001) [JETP Lett. **73**, 237 (2001)].
16. G. Martin, Phys. Rev. B **41**, 2279 (1990).
17. V. G. Vaks and I. R. Pankratov, Zh. Éksp. Teor. Fiz. **124** (2003) (in press).
18. F. Chassagne, M. Bessiere, Y. Calvayrac, *et al.*, Acta Metall. **37**, 2329 (1989).
19. A. Loiseau, C. Ricolleau, L. Potez, and F. Ducastelle, in *Solid-State Phase Transformations*, Ed. by W. C. Johnson *et al.* (TMS, Warrendale, 1994), p. 385.
20. L. Potez and A. Loiseau, J. Interface Sci. **2**, 91 (1994).
21. R. Kikuchi and J. W. Cahn, Acta Metall. **27**, 1337 (1979).
22. B. Kh. Kir, V Mire Nauki, No. 12, 99 (1986).
23. V. I. Syutkina, R. Z. Abdulov, N. D. Zemtsova, and L. P. Yasyreva, Fiz. Met. Metalloved. **58**, 745 (1984).
24. Y. LeBouar, in *Proceedings of III International Alloy Conference, Lisbon, 2002* (in press).

Translated by A. Bagatur'yants