JETP Letters, Vol. 78, No. 4, 2003, pp. 179–182. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 215–218.
Original English Text Copyright © 2003 by Arbuzov.

Virtual and Soft Pair Corrections to Polarized Muon Decay Spectrum¶

A. B. Arbuzov

Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, Dubna, Moscow region, 141980 Russia

e-mail: arbuzov@thsun1.jinr.ru

Received July 2, 2003

Radiative corrections to the muon decay spectrum due to soft and virtual electron–positron pairs are calculated within the next-to-next-to-leading approximation. The effect is relevant for modern precision experiments on the decay spectrum. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 13.35.Bv; 13.40.Hq; 12.20.Ds

1. Introduction. The TWIST experiment (TRIUMF Weak Interaction Symmetry Test) [1, 2] is currently running at Canada's National Laboratory TRIUMF. It is going to measure the muon decay spectrum [3, 4] with an accuracy level of about 1×10^{-4}. That will make a serious test of the space-time structure of the weak interaction. The experiment is able to put stringent limits on many parameters in models beyond the Standard Model (SM), e.g., on the mass and the mixing angle of a possible right-handed W boson. To confront the experimental results with the SM, sufficiently accurate theoretical predictions should be provided. This requires one to calculate radiative corrections within perturbative quantum electrodynamics (QED). Here we will present analytical results for two specific contributions, related to radiation of virtual and soft real electron–positron pairs. The corrections under consideration are of the order $\mathbb{O}(\alpha^2)$, where α is the fine structure constant.

The contributions of virtual $\mu^+\mu^-$, $\tau^+\tau^-$, and hadronic pairs were found [5] to be small compared with the 1×10^{-4} precision tag of the modern experiments. The contribution of e^+e^- pairs is enhanced by powers of the large logarithm $L = \ln(m_\mu^2/m_e^2) \approx 10.66$. Analysis of the leading and next-to-leading terms from this correction in [6, 7] has shown that the numerical effect not as small as for other leptonic flavors, and it should be taken into account. Comparison of the leading and next-to-leading contributions revealed a poor convergence of the series in L. Calculation of the terms without the large logarithm was found to be desirable.

Within the SM, the differential distribution of electrons (summed over electron spin states) in the polarized muon decay can be represented as

$$\frac{d^2\Gamma^{\mu^\mp \to e^\mp \nu\bar{\nu}}}{dx\,dc} = \Gamma_0[F(x) \pm cP_\mu G(x)],$$

$$\Gamma_0 = \frac{G_F^2 m_\mu^5}{192\pi^3}, \quad c = \cos\theta, \quad x = \frac{2m_\mu E_e}{m_\mu^2 + m_e^2}, \qquad (1)$$

$$x_0 \le x \le 1, \quad x_0 = \frac{2m_\mu m_e}{m_\mu^2 + m_e^2},$$

where m_μ and m_e are the muon and electron masses; G_F is the Fermi coupling constant; θ is the angle between the muon polarization vector $\mathbf{P}_\mu$ and the electron (or positron) momentum; and E_e and x are the energy and the energy fraction of $e^\pm$. Here we adopt the definition of the Fermi coupling constant following [8]. The functions $F(x)$ and $G(x)$ describe the isotropic and anisotropic parts of the spectrum, respectively. Within perturbative QED, they can be expanded in series in α,

$$F(x) = f_{\text{Born}}(x) + \frac{\alpha}{2\pi} f_1(x) + \left(\frac{\alpha}{2\pi}\right)^2 f_2(x) + \mathbb{O}(\alpha^3),$$

and in the same way for $G(x)$. Among the different contributions to the functions $F(x)$ and $G(x)$ (see [6] for details and discussion), there are ones related to electron–positron pair production. In this letter we will consider the effect of soft and virtual e^+e^- pairs.

2. Soft e^+e^- pairs. The process of real pair production does not reveal any infrared singularity, contrary to the case of photon radiation. Nevertheless, a separate consideration of soft pair emission can be of interest. In fact, e^+e^- pairs with energy below a certain threshold cannot be observed in experiments with muons decaying at rest. So, the corresponding contribution is a specific correction to the measured decay spectrum. More-

¶ This article was submitted by the author in English.

over, the behavior of the real pair emission in the soft limit is not smooth. An integration over the domain between the threshold of real pair production and a certain cut-off on the maximal energy of the soft pair is desirable.

The maximal energy of the soft pair is assumed to be large compared with the electron mass:

$$E^{\text{pair}} \le \Delta\frac{m_\mu}{2}, \quad \frac{m_e}{m_\mu} \ll \Delta \ll 1. \tag{2}$$

Due to the smallness of the pair component energies, the matrix element M of the process

$$\mu^-(p) \longrightarrow e^-(q) + \nu_\mu(r_1) + \bar\nu_e(r_2) + e^+(p_+) + e^-(p_-)$$

can be expressed as a product of the matrix element M_0 of the hard subprocess (the nonradiative muon decay) and the classic accompanying radiation factor:

$$M = M_0\frac{4\pi\alpha}{k^2}\bar v(p_+)\gamma^\mu u(p_-)J_\mu, \quad k = p_+ + p_-,$$

where $p_{+,-}$ are the momenta of the positron and electron from the created pair. The radiation factor reads

$$J_\mu = \frac{p_\mu}{pk - \frac{1}{2}k^2} - \frac{q_\mu}{qk + \frac{1}{2}k^2}.$$

Performing the covariant integration of the (summed over spin states) modulus of the matrix element over the pair components momenta, we obtain

$$\sum_{\text{spin}}\left|\bar v(p_+)\gamma^\mu u(p_-)\right|^2 = 4\left(p_+^\mu p_-^\nu + p_+^\nu p_-^\mu - \frac{k^2}{2}g^{\mu\nu}\right),$$

$$\int\frac{d^3\mathbf{p}_+ d^3\mathbf{p}_-}{p_+^0 p_-^0}\delta^4(p_+ + p_- - k)\left(p_+^\mu p_-^\nu + p_+^\nu p_-^\mu - \frac{k^2}{2}g^{\mu\nu}\right)$$

$$= \left(-\frac{2\pi}{3}(k^2 + 2m_e^2)\sqrt{1 - \frac{4m_e^2}{k^2}}\right)\left(g^{\mu\nu} - \frac{1}{k^2}k^\mu k^\nu\right).$$

It is convenient to parameterize the phase volume of the total pair momentum as

$$d^4k = dk_0\mathbf{k}^2 d|\mathbf{k}|d\Omega_k = \pi dk_0 dk^2\sqrt{k_0^2 - k^2}\,dc_k d\phi_k,$$

where a trivial integration over the azimuthal angle can be performed: $\int d\phi_k \longrightarrow 2\pi$. Now we integrate over the total pair momentum with the condition (2) ($k_0 \equiv E^{\text{pair}}$). In this way, we get the following result for the soft pair contribution:

$$\frac{d\Gamma^{\text{SP}}}{dcdx} = \frac{d\Gamma^{\text{Born}}}{dcdx}\delta^{\text{SP}},$$

$$\frac{d\Gamma^{\text{Born}}}{dcdx} = \Gamma_0[f_0(x) \pm cP_\mu g_0(x)] + \mathbb{O}\left(\frac{m_e^2}{m_\mu^2}\right),$$

$$f_0(x) = x^2(3-2x), \quad g_0(x) = x^2(1-2x),$$

$$\delta^{\text{SP}} = \frac{\alpha^2}{3\pi^2}\left[\frac{1}{12}\ln^3 A - \frac{2}{3}\ln^2 A + \ln A\left(\frac{61}{18} - \zeta(2)\right)\right. \tag{3}$$

$$\left. - \frac{223}{27} + \frac{8}{3}\zeta(2) + 2\zeta(3)\right],$$

$$\ln A = L + 2\ln\Delta, \quad \zeta(n) = \sum_{k=1}^{\infty}\frac{1}{k^n}, \quad \zeta(2) = \frac{\pi^2}{6}.$$

So, we have calculated explicitly all the terms in δ^{SP} except the ones suppressed by the small factors $(\alpha/\pi)^2 m_e^2/m_\mu^2$ and $(\alpha/\pi)^2\Delta$.

3. Virtual e^+e^- pair. We will use here the substitution suggested by Schwinger for the photon propagator (with 4-momentum k) corrected by a one-loop vacuum polarization insertion:

$$\frac{1}{k^2 - \lambda^2 + i0} \longrightarrow \frac{\alpha}{\pi}\int_0^1 dv\,\phi(v)\frac{1}{1-v^2}\frac{1}{k^2 - M^2 + i0}, \tag{4}$$

$$M^2 = \frac{4m_2^2}{1-v^2}, \quad \phi(v) = \frac{2}{3} - \frac{1}{3}(1-v^2)(2-v^2),$$

where m_2 is the mass of the fermion in the loop. Using the substitution, we can reproduce the known [9, 10] asymptotic expressions for the $\mathbb{O}(\alpha^2)$ virtual pair contributions into the Dirac form factor of the muon (see Appendix B in preprint [11]).

The standard technique of integration over Feynman parameters can be used here. We are interested in the region of electron energy fractions $x \gg m_e/m_\mu$. Analytical expressions for the relevant integrals in this region can be found in preprint [11]. As concerns the region of small electron energy fractions, it requires a more accurate treatment. But the differential width there decreases rapidly (see, e.g., the Born level functions in Eq. (3)), and the contribution of this region to the total width is also suppressed by the mass ratio.

Formally, we have an ultraviolet singularity in the virtual pair correction. The Fermi theory is not renormalizable in the general case. But for the muon decay everything is safe, since the standard renormalization of the electron and muon wave functions removes the singularity [12]. Note that we need to use only the pair

contribution into the renormalization constants. They can be found easily from the calculation of the virtual pair corrections to the Dirac form factors of the electron and muon and used as described in [13].

We get the following result for the virtual e^+e^- pair contribution:

$$\frac{\Gamma^{\mathrm{VP}}}{dcdx} = \Gamma_0\left(\frac{\alpha}{2\pi}\right)^2\left[f_{2,\,\mathrm{virt}}^{(e^+,e^-)}(x) \pm cP_\mu g_{2,\,\mathrm{virt}}^{(e^+e^-)}(x) + \mathbb{O}\left(\frac{m_e^2}{m_\mu^2}\right)\right], \tag{5}$$

where

$$f_{2,\,\mathrm{virt}}^{(e^+e^-)}(x) = f_0(x)W(x) - 2x^2\ln xL - 2x^2\ln^2 x - 2x^2\mathrm{Li}_2(1-x) - \frac{2}{3(1-x)}\ln x + \frac{2}{3}x\ln x + 7x^2\ln x + \frac{2}{3}\ln x,$$

$$g_{2,\,\mathrm{virt}}^{(e^+e^-)}(x) = g_0(x)W(x) - \frac{2}{3}x^2\ln xL - \frac{2}{3}x^2\ln^2 x - \frac{2}{3}x^2\mathrm{Li}_2(1-x) + \frac{2}{3(1-x)}\ln x - \frac{2}{3}x\ln x + \frac{13}{9}x^2\ln x - \frac{2}{3}\ln x,$$

$$W(x) = -\frac{1}{9}L^3 + \left(\frac{25}{18} - \frac{2}{3}\ln x\right)L^2 + \left(-\frac{397}{54} - \frac{4}{3}\zeta(2) + \frac{38}{9}\ln x - \frac{4}{3}\ln^2 x - \frac{4}{3}\mathrm{Li}_2(1-x)\right)L + \frac{517}{27} - \frac{8}{3}\zeta(2)\ln x + \frac{22}{9}\zeta(2) + \frac{4}{3}\zeta(3) - \frac{8}{3}\ln x\mathrm{Li}_2(1-x) - \frac{265}{27}\ln x + \frac{38}{9}\ln^2 x - \frac{8}{9}\ln^3 x + \frac{38}{9}\mathrm{Li}_2(1-x) - \frac{8}{3}\mathrm{S}_{1,2}(1-x) + \frac{4}{3}\mathrm{Li}_3(1-x),$$

$$\mathrm{Li}_2(x) \equiv -\int_0^x dy\frac{\ln(1-y)}{y}, \quad \mathrm{Li}_3(x) \equiv \int_0^x dy\frac{\mathrm{Li}_2(y)}{y},$$

$$\mathrm{S}_{1,2}(x) \equiv \frac{1}{2}\int_0^x dy\frac{\ln^2(1-y)}{y}.$$

It is worth noting that the subleading virtual corrections don't factorize before the Born functions $f_0(x)$ and $g_0(x)$.

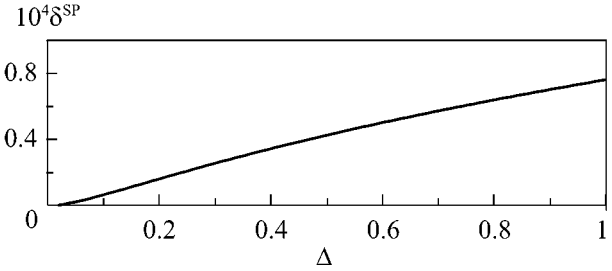


Fig. 1. The relative effect of soft pair corrections versus the cutoff value.

By integrating over the energy fraction and the angle, we receive the corresponding contribution to the total muon width:

$$\Gamma^{\mathrm{VP}} = \int_{-1}^{1} dc\int_0^1 dx\frac{\Gamma^{\mathrm{VP}}}{dcdx} = \Gamma_0\left(\frac{\alpha}{2\pi}\right)^2\left[-\frac{1}{9}L^3 + \frac{5}{3}L^2 - \left(\frac{265}{36} + \frac{8}{3}\zeta(2)\right)L + \frac{20063}{1296} + \frac{61}{9}\zeta(2) + \frac{16}{3}\zeta(3)\right] \approx -5.0497\times 10^{-5}\Gamma_0. \tag{6}$$

This quantity was calculated earlier in [14] by numerical integration using dispersion relations:

$$\Gamma^{\mathrm{VP}}([10]) \approx -5.1326\times 10^{-5}\Gamma_0, \tag{7}$$

which is close but different from my number (6). The main reason for this deviation seems to be in the different approaches to the renormalization of ultraviolet singularities. Subtraction at $q^2 = 0$ (q is the momentum transferred between the electron and muon) was used in [14] in analogy to the case of the Dirac form factor calculations. But in the case of muon decay, the subtraction of singularities should be done as described here (see details in preprint [11]). And for nonequal masses of the charged particles, the subtraction point does not correspond to $q^2 = 0$. A certain (small) part of the numerical deviation can also be due to terms proportional to $(\alpha/\pi)^2(m_e^2/m_\mu^2)L^n$, which were omitted in my calculation.

The correction to the forward–backward asymmetry of the decay can also be found:

$$\Gamma_{\mathrm{FB}}^{\mathrm{VP}} = \left[\int_0^1 dc - \int_{-1}^0 dc\right]\int_0^1 dx\frac{\Gamma^{\mathrm{VP}}}{dcdx} = \Gamma_0\left(\frac{\alpha}{2\pi}\right)^2\left[\frac{1}{54}L^3 - \frac{13}{54}L^2 + \left(\frac{647}{648} + \frac{4}{9}\zeta(2)\right)L - \frac{10339}{7776} - \frac{3}{2}\zeta(2) - \frac{8}{9}\zeta(3)\right] \approx -1.1676\times 10^{-5}\Gamma_0. \tag{8}$$

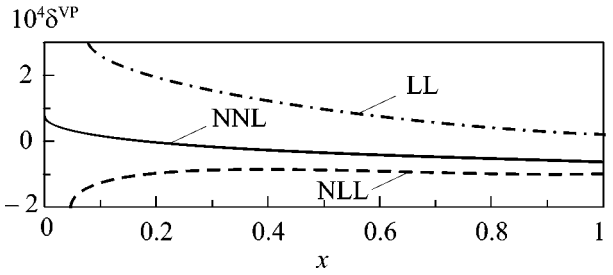


Fig. 2. The relative effect of virtual pair corrections versus electron energy fraction in different approximations.

4. NumericaL results and conclusions. The relative effect of the soft pair correction depends only on the cutoff value. It is shown in Fig. 1. The soft pair approximation (2) is not valid for values of Δ close to the threshold of real pair production and for large $\Delta \sim 1$. But it can be used there as a simple estimate. So, by taking $\Delta = 1$, we make an estimate of the order of magnitude of the total contribution due to real e^+e^- pairs (here the estimate is about two times the true value). For very small values of Δ, the correction should vanish in any case, so the approximation is very safe there.

Let us define the relative contribution of the virtual e^+e^- pair corrections in the form

$$\delta^{VP}(x) = \left(\frac{\alpha}{2\pi}\right)^2 \frac{f_{2,\,\mathrm{virt}}^{(e^+,e^-)}(x) + cP_\mu g_{2,\,\mathrm{virt}}^{(e^+e^-)}(x)}{f_0(x) + cP_\mu g_0(x)}. \qquad (9)$$

The dependence of this function on the electron energy fraction is shown in Fig. 2 in different approximations for $P_\mu = 1$, $c = 1$. The dependence on c is very weak, because the main part of the correction is factorized before the Born level functions. The leading logarithmic (LL) approximation takes into account only the terms of the order $\mathbb{O}(\alpha^2 L^{3,2})$, the next-to-leading logarithmic (NLL) approximation also includes the $\mathbb{O}(\alpha^2 L^1)$ terms, and the next-to-next-to-leading approximation (NNL) represents the complete result.

The third power of the large logarithm cancels out in the sum of the virtual and soft pair contributions:

$$\frac{\Gamma^{SVP}}{dcdx} = \Gamma_0\left(\frac{\alpha}{2\pi}\right)^2 [\, f_{2,\,SV}^{(e^+e^-)}(x) \pm cP_\mu g_{2,\,SV}^{(e^+e^-)}(x)], \qquad (10)$$

where

$$f_{2,\,SV}^{(e^+e^-)}(x) = \frac{4}{3}\delta^{SP} f_0(x) + f_{2,\,virt}^{(e^+e^-)}(x),$$

$$g_{2,\,SV}^{(e^+e^-)}(x) = \frac{4}{3}\delta^{SP} g_0(x) + g_{2,\,virt}^{(e^+e^-)}(x).$$

I checked that the leading and next-to-leading terms in these sums agree with the corresponding contribution obtained within the fragmentation function formalism in [6, 7].

Summing up the virtual and soft pair contributions, we simulate the experimental setup with a certain energy threshold for registration of pairs, while events with pair production above the threshold (with several visible charged particles in the final state) are rejected.

If the radiation of real pairs is completely forbidden by kinematics (or experimental conditions), only the virtual corrections (5) contribute. That happens, for instance, at large values of $x \gtrsim 0.99$.

Thus, two contributions to the total set of radiative corrections for the muon decay spectrum are presented. They are required to reach the level of the theoretical accuracy below 1×10^{-4}. The formulas can be used for semianalytical estimates and as a part of a Monte Carlo code to describe the pair production contribution to the decay spectrum. The formulas are also valid for pair corrections to leptonic τ decays.

I am grateful to V. Bytev and E. Kuraev for discussions. This work was supported by the Russian Foundation for Basic Research, grant no. 03-02-17077.

REFERENCES

1. N. L. Rodning, W. Andersson, Yu. Davydov, *et al.*, Nucl. Phys. (Proc. Suppl.) **98**, 247 (2001).
2. M. Quraan, A. Gaponenko, P. Green, *et al.*, Nucl. Phys. A **663**, 903 (2000).
3. W. Fetscher and H. J. Gerber, Eur. Phys. J. C **15**, 316 (2000).
4. Y. Kuno and Y. Okada, Rev. Mod. Phys. **73**, 151 (2001).
5. A. I. Davydychev, K. Schilcher, and H. Spiesberger, Eur. Phys. J. C **19**, 99 (2001).
6. A. Arbuzov, J. High Energy Phys. **0303**, 063 (2003).
7. A. Arbuzov and K. Melnikov, Phys. Rev. D **66**, 093003 (2002).
8. T. van Ritbergen and R. G. Stuart, Nucl. Phys. B **564**, 343 (2000).
9. G. J. Burgers, Phys. Lett. B **164B**, 167 (1985).
10. A. H. Hoang, J. H. Kuhn, and T. Teubner, Nucl. Phys. B **452**, 173 (1995).
11. A. B. Arbuzov, hep-ph/0301114.
12. S. M. Berman and A. Sirlin, Ann. Phys. (N.Y.) **20**, 20 (1962).
13. A. B. Arbuzov, Phys. Lett. B **524**, 99 (2002).
14. T. van Ritbergen and R. G. Stuart, Phys. Rev. Lett. **82**, 488 (1999).

JETP Letters, Vol. 78, No. 4, 2003, pp. 183–187. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 219–223.
Original Russian Text Copyright © 2003 by Gurov, Aleshkin, Lapushkin, Morokhov, Panin, Pechkurov, Poroshin, Sandukovsky, Tel'kushev, Chernyshev.

The Formation of the Superheavy Hydrogen Isotope ^{6}H in the Absorption of Stopped π^--Mesons by Nuclei

Yu. B. Gurov[1,*], D. V. Aleshkin[1], S. V. Lapushkin[1], P. V. Morokhov[1], A. V. Panin[1], V. A. Pechkurov[1], N. O. Poroshin[1], V. G. Sandukovsky[2], M. V. Tel'kushev[1], and B. A. Chernyshev[1]

[1] *Moscow Engineering Physics Institute (State University), Kashirskoe sh. 31, Moscow, 115409 Russia*
[2] *Joint Institute for Nuclear Research, Dubna, Moscow region, 141980 Russia*
** e-mail: Gurov@axpk40.mephi.ru*

Received July 17, 2003

An experimental search for the superheavy hydrogen isotope ^{6}H was conducted through studying the absorption of stopped π^--mesons by ^{9}Be and ^{11}B nuclei. A structure in the missing mass spectrum caused by the resonance states of ^{6}H was observed in three reaction channels, namely, ^{9}Be(π^-, pd)X, ^{11}B(π^-, d^3He)X, and ^{11}B(π^-, p^4He)X. The parameters of the lowest state $E_r = 6.6 \pm 0.7$ MeV and $\Gamma = 5.5 \pm 2.0$ MeV (E_r is the resonance energy with respect to the disintegration into the triton and three neutrons) are evidence that ^{6}H is a more weakly bound system than ^{4}H and ^{5}H. Three excited states of ^{6}H were observed. Their resonance levels ($E_{1r} = 10.7 \pm 0.7$ MeV, $\Gamma_{1r} = 4 \pm 2$ MeV, $E_{2r} = 15.3 \pm 0.7$ MeV, $\Gamma_{2r} = 3 \pm 2$ MeV, and $E_{3r} = 21.3 \pm 0.4$ MeV, $\Gamma_{3r} = 3.5 \pm 1.0$ MeV) are energetically capable of disintegrating into six free nucleons. *© 2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 25.80.Hp; 21.10.-k

Studies of light nuclei close to the nucleon stability boundary have been focused on the spectroscopy of superheavy hydrogen isotopes ^{n}H ($n \geq 4$), to a great extent because of the expectation that stable nucleon isotopes with a large number of neutrons can be observed, for theory does not forbid the existence of such nuclei. Note that the question of the applicability of the existing theoretical models to isotopes heavier than ^{5}H remains open, because such nuclear states are difficult to reduce to a three-particle system. Experimental information about hydrogen isotopes heavier than ^{4}H is scarce and contradictory. For instance, the recent data on the ground state of the ^{5}H isotope [1–4] are at variance with each other, and the question of the reasons for the discrepancies remains open [3]. In this situation, it is not clear whether or not the phenomenological rules obtained in studies of 2p-shell nuclei at the nucleon stability boundary can be extended to superheavy hydrogen isotopes [5].

Up to now, indications of the existence of the ^{6}H isotope have only been obtained in reactions on heavy ions [6]. The resonance state of ^{6}H with $E_r = 2.7 \pm 0.4$ MeV and $\Gamma = 1.8 \pm 0.5$ MeV (E_r is the resonance energy with respect to the disintegration to the triton and three neutrons) was recorded for the ^{7}Li(^{7}Li, ^{8}B)X reaction at a $E(^7\text{Li}) = 82$ MeV energy [7]. The resonance state of ^{6}H with $E_r = 2.6 \pm 0.5$ MeV and $\Gamma = 1.3 \pm 0.5$ MeV was observed in the reaction ^{9}Be(^{11}B, ^{14}O)X at a $E(^{11}\text{B}) = 88$ MeV energy [8]. The parameter values given above are in agreement with each other, but the measurement statistics were very poor in both works. No ^{6}H was observed in the ^{6}Li(π^-, π^+)X double recharging reaction at $E_{\pi^-} = 220$ MeV [9]. Nor were we able to record the formation of ^{6}H in two reaction channels, namely, ^{9}Be(π^-, pd)X [10] and ^{7}Li(π^-, p)X [11], in our early works. The measurement statistics and the energy resolution were, however, insufficient for obtaining reliable spectroscopic data on ^{6}H.

The use of shell models to describe ^{6}H can hardly be considered justified. It can, nevertheless, be noted that the existence of states with E_r= 1.78, 2.8, and 4.79 MeV has been predicted in [12]. The calculations performed in [13] by the method of angular potential functions gave $E_r = 6.3$ MeV for the ground state of ^{6}H.

In this work, the search for the ^{6}H isotope was made by recording missing mass spectra for the following reaction channels of the absorption of stopped π^--mesons by nuclei:

$$^9\text{Be}(\pi^-, pd)\text{X},\ ^{11}\text{B}(\pi^-, d^3\text{He})\text{X}\ ^{11}\text{B}(\pi^-, p^4\text{He})\text{X}.$$

Experiments were performed using a multilayer semiconductor spectrometer [13] with a beam of low-energy pions at the Los Alamos Meson Physics Facility. Secondary charged particles formed in the absorption

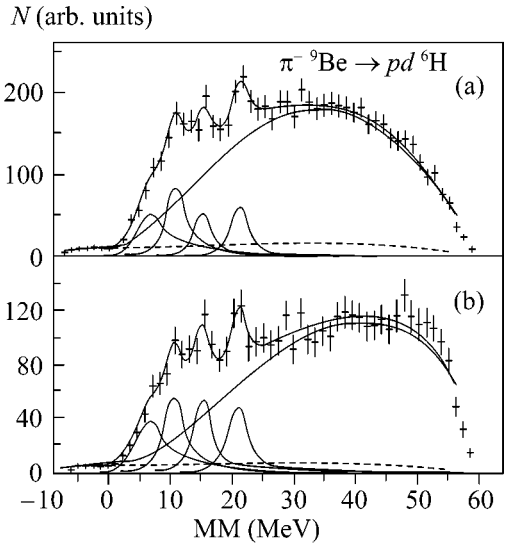


Fig. 1. Missing mass spectra for the reaction ^{9}Be(π^-, *pd*)X: (a) experimental spectrum and (b) experimental spectrum obtained under the $P_R \le 100$ MeV/*c* limitation on the nonrecordable residue momentum. The solid lines are the complete description and the Breit–Wigner distributions, (*1*) total distribution over the phase volume and (*2*) background of fortuitous coincidences.

of pions by nuclei were recorded with two silicon telescopes oriented at a 180° angle with respect to each other. Each telescope enabled charged particles to be identified and their energy to be measured up to the kinematic reaction boundaries. The energy resolution for singly charged particles (*p*, *d*, *t*) and helium ions (^{3,4}He) was 0.5 and 1.0 MeV, respectively.

The energy resolution of the missing mass spectra was determined by correlation measurements for the reactions ^{11}B(π^-, *pd*)^{8}He [14], ^{12}C(π^-, p^4He)^{7}He, and ^{12}C(π^-, d^3He)^{7}He (measurements for the ^{12}C target were performed in the same experimental run). The principal states and widths of the produced helium isotopes were reliably determined [6, 15]. An analysis of the results showed that the missing mass energy resolution for the events with the *pd*-, p^4He, and d^3He pairs was 1.0, 2.5, and 2.0 MeV. A decrease in resolution for the events with the ^{3,4}He ions was not caused by an increase in their ionization-induced energy loss in the target compared with the *p* and *d* ions. The error in absolute scale referencing was ≤0.1 MeV. The ^{8}He peak parameters were also used to control the time stability of the characteristics of the spectrometer in work with the ^{11}B target. Correlation measurements of *tt* pairs on the ^{9}Be target were performed for a similar purpose [3].

Quantitative determination of possible impurities in the targets was performed by identifying the peaks that corresponded to the known two-particle reactions on impurity nuclei. The major impurity in the ^{11}B target was ^{12}C (8%). The contribution of the other (uncontrolled) impurities in the ^{11}B and ^{9}Be targets did not exceed 1%.

The spectrometer and the procedure for measurements were described in more detail in [14, 16].

The missing mass (MM) spectrum for the reaction ^{9}Be(π^-, *pd*)X is shown in Fig. 1a. The sum of the masses of the triton and three neutrons was used as the origin. First note the absence of features corresponding to ^{6}H bound states in the region of negative missing masses. A weak background in this region is caused by fortuitous coincidences in correlation measurements. At the same time, the spectrum in the region of positive missing masses contains structures that can correspond to the formation of resonance states. We used the method of least squares to describe the experimental spectra by the sum of *n*-particle distributions over the phase volume (all possible final states with $n \ge 4$, including the reaction channels that yielded ^2n, ^{4}H, and ^{5}H, were taken into account) and Breit–Wigner distributions in order to identify these ^{6}H states and determine their parameters. The calculations were performed taking into account the angular and energy resolution of the spectrometer and the background of fortuitous coincidences.

As follows from Fig. 1a, the distributions over phase volumes cannot reproduce the structure observed at missing masses below 25 MeV. Note that the major contribution to the total distribution is made by the 5-particle phase volume with the dineutron in the final state ($d + p + {}^2n + t + n$). A satisfactory description (with a χ^2 value of 0.95 per degree of freedom) of the experimental spectrum can only be attained by introducing four ^{6}H resonance states, whose parameters are listed in the table. The Γ values are the total widths at half-heights of the peaks shown in the figures. The parameter uncertainties listed in the table are caused by both statistical and systematic measurement errors.

The use of only distributions over the phase volume and the Breit–Wigner distributions to describe the missing mass spectrum leaves open the question of the significance of final state interactions, whose role is noticeable in the absorption of π^--mesons by light nuclei [17]. Final state interaction between nuclear residue particles was taken into account by including reac-

Resonance parameters of the ^{6}H isotope

Reaction channel			
^{9}Be(π^-, *pd*)^{6}H		^{11}B(π^-, p^4He)^{6}H	
E_r (MeV)	Γ (MeV)	E_r (MeV)	Γ (MeV)
6.6 ± 0.7	5.5 ± 2.0	7.3 ± 1.0	5.8 ± 2.0
10.7 ± 0.7	4 ± 2	–	–
15.3 ± 0.7	3 ± 2	14.5 ± 1.0	5.5 ± 2.0
21.3 ± 0.4	3.5 ± 1.0	22.0 ± 1.0	5.5 ± 2.0

tion channels with the formation of ^2n, ^{4}H, and ^{5}H. It is, however, unclear what contribution is made by reaction channels with final state interactions between one of the recorded particles and the neutron, namely, π^- + ^{9}Be $\longrightarrow$ p + $(dn)_{FSI}$ + R and π^- + ^{9}Be $\longrightarrow$ $(pn)_{FSI}$ + d + R, in which correlations of particle velocities arise in the "quasitriton" $(dn)_{FSI}$ and "quasideuteron" $(pn)_{FSI}$ unbound states. We can nevertheless suggest a method that allows the contribution of these channels to be strongly suppressed in the important case of quasifree absorption of pions.

The quasifree processes in which residual nucleus nucleons do not directly participate make a substantial contribution to the absorption of pions by light nuclei [17]. The residual nucleus momentum p_R is determined in these processes by intranuclear Fermi motion. For instance, for the reaction with the quasitriton, we have

$$p_{(dn)_{FSI}} \approx p_p \text{ and } T_{(dn)_{FSI}} + T_p = Q,$$

where Q is the total kinetic energy of the particles that are formed in the absorption of the pion and p and T are the momentum and kinetic energy of the particles, respectively.

The neutron momentum can be estimated in the nonrelativistic approximation on the assumption that the velocities of the neutron and deuteron are equal in the quasitriton. This gives

$$p_n = \frac{m_n}{m_n + m_d}\sqrt{2\frac{m_p(m_n + m_d)}{m_p + m_n + m_d}Q} \approx 120 \text{ MeV/s}.$$

Note that the neutron momentum in the reaction channel with the quasideuteron is much higher ($p_n \sim$ 200 MeV/c).

In our experiments, the energy of such neutrons was not recorded. The experimental data were therefore analyzed with the momentum $p_n(p_n)$ referred to the nuclear residue. The contribution of reaction channels with final state interactions can be substantially suppressed by introducing a limitation on the residual nucleus momentum.

The missing mass spectrum recorded with the $p_R <$ 100 MeV/c limitation is shown in Fig. 1b. This value cannot exceed the expected value for the Fermi momentum of the intranuclear cluster. Therefore, the introduced limitation also allows us to increase the relative contribution of quasifree absorption to the observed spectrum. The description of the spectrum with the Breit–Wigner distribution parameters listed in the table gives χ^2 equal to 1.01 per degree of freedom, which substantiates the hypothesis of the existence of four resonance states of the ^{6}H isotope.

With the ^{11}B target, the formation of ^{6}H can be observed in the missing mass spectra for two reaction channels, namely, ^{11}B(π^-, p^4He)X (Fig. 2) and ^{11}B(π^-, d^3He)X (Fig. 3). The ^{11}B target contained a ^{12}C impurity. For this reason, the contribution of the reactions ^{12}C(π^-, p^4He)X and ^{12}C(π^-, d^3He)X was subtracted from the experimental spectra (Figs. 2a, 3a). The contribution of these reactions (Figs. 2b, 3b) was determined by normalizing the spectra measured with the ^{12}C target in the same experimental session taking into account the relative fraction of the impurity (8%). The spectra obtained after the subtraction are shown in Figs. 2c and 3c.

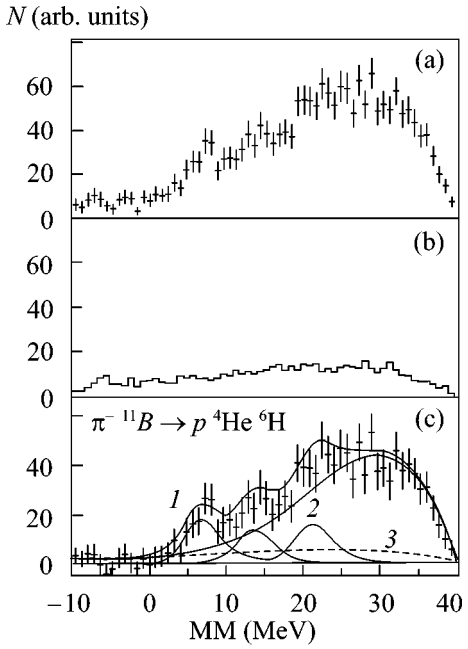


Fig. 2. Missing mass spectra for the ^{11}B(π^-, p^4He)X reaction: (a) spectrum measured on the ^{11}B target, (b) spectrum measured for the ^{12}C(π^-, p^4He)X reaction normalized taking into account the amount of the ^{12}C impurity in the ^{11}B target, and (c) spectrum obtained by subtracting the impurity contribution. The solid lines are the Breit–Wigner distributions; (*1*) complete description, (*2*) total distribution over the phase volume, and (*3*) background of fortuitous coincidences.

The spectra recorded with ^{11}B were analyzed by two methods. First, we described the spectra using the ^{6}H resonance state parameters obtained with ^{9}Be. This gave χ^2 of 0.88 and 0.97 per degree of freedom for the ^{11}B(π^-, p^4He)X and ^{11}B(π^-, d^3He)X reactions, respectively, which did not contradict the suggestion of the existence of four ^{6}H isotope levels. Next, the spectra were processed using level positions and widths and the number of levels as free parameters. The spectrum for the reaction ^{11}B(π^-, p^4He)X (Fig. 2c) can then be described (χ^2 of 0.87 per degree of freedom) by including only three resonance states with the parameters listed in the table. Note that these parameter values coincide with the results obtained for the ^{9}Be(π^-, pd)X

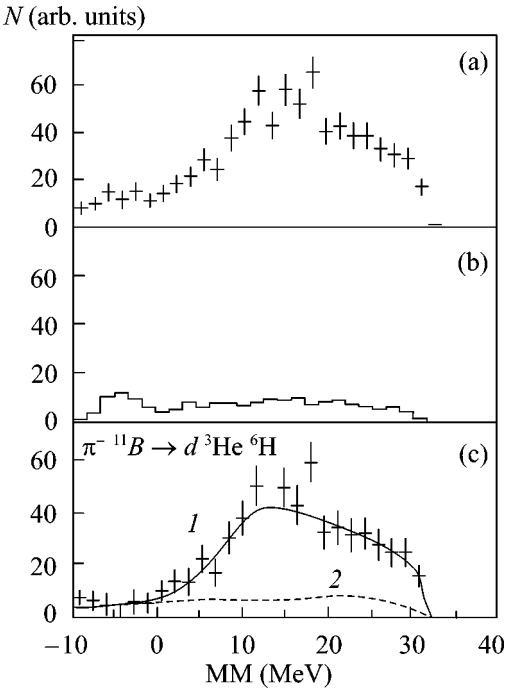


Fig. 3. Missing mass spectra for the $^{11}B(\pi^-, d^3He)X$ reaction: (a) spectrum measured on the ^{11}B target, (b) spectrum measured for the $^{12}C(\pi^-, d^3He)X$ reaction normalized taking into account the amount of the ^{12}C impurity in the ^{11}B target, and (c) spectrum obtained by subtracting the impurity contribution. The solid line is the Breit–Wigner distributions; (*1*) complete description and (*2*) background of fortuitous coincidences.

reaction to within the uncertainties. The spectrum for the $^{11}B(\pi^-, d^3He)X$ reaction (Fig. 3c) can be described ignoring the existence of ^6H, but a satisfactory χ^2 value per degree of freedom (1.0) can only be attained by including the phase volume of the $p + d + n + {}^5H$ channel into the total spectrum.

The analysis given above shows that the data obtained with two targets are not at variance with each other. However, note that the energy resolution of the spectra obtained with ^{11}B was worse and the measurement statistic poorer compared with ^9Be experiments. This impeded the observation of ^6H in the reaction with ^{11}B and led us to suggest that the data obtained for the $^9Be(\pi^-, pd)X$ reaction more correctly reproduced the structure of the ^6H isotope levels.

The resonance energy of the ^6H isotope ground state obtained in this work is substantially higher than those reported in the earlier works [9, 10]. It must be stressed that the measurement statistic in this work was more than an order of magnitude larger. Our results closely agree with the theoretical predictions made in [12]. Note that equally close agreement was obtained for the ground state of the ^5H isotope [3, 12]. At the same time, our data are at variance with the structure of ^6H levels predicted by the shell model [11].

Additional evidence for the existence of levels with $E_r = 10.7$ and 15.3 MeV can be obtained from the spectra of the ^6He isotope [6]. The missing mass spectrum measured in the $^7Li(^3He, p^3He)X$ reaction at $E(^3He) = 120$ MeV contained two comparatively narrow ($\Gamma \leq 2$ MeV) ^6H states with excitation energies $E_x \approx 32.0$ and 35.7 MeV [18]. The recalculation of these values to the binding energies of the corresponding states (the B value is positive for bound systems) gives $B(^6H) = -2.2 \pm 0.7$ and -6.8 ± 0.7 MeV and $B(^6He) \approx -2.7$ and -6.4 MeV, respectively. The Coulomb energy does not exceed 0.7 MeV in ^6He [19], which leads us to suggest that the observed levels are isobar analogue states.

To summarize, the following conclusions about superheavy hydrogen isotopes can be drawn from our study of pion absorption by light nuclei. The binding energy of the ground state gradually decreases as the number of neutrons increases; compare $B(^4H_{g.s}) = 6.5 \pm 0.2$ MeV [20], $B(^5H_{g.s}) = 3.0 \pm 0.2$ MeV [3], and $B(^6H_{g.s}) = 1.9 \pm 0.7$ MeV. The ^5H and ^6H superheavy hydrogen isotopes have several excited levels, which are energetically capable of disintegrating into free nucleons.

This work was financially supported by the Universities of Russia program (project no. UR.02.01.007) and CRDF (grant no. MO-011-0).

REFERENCES

1. D. V. Aleksandrov, E. Y. Nikol'skii, B. G. Novatskii, *et al.*, in *Proceedings of the International Conference on Exotic Nuclei and Atomic Masses (ENAM-95), Arles, France, 1995*, Ed. by Frontiers (Gif-sur-Yvette, France, 1995), p. 329.
2. A. Korsheninnikov, M. S. Golovkov, I. Tanihata, *et al.*, Phys. Rev. Lett. **87**, 092501 (2001).
3. M. G. Gornov, M. N. Ber, Yu. B. Gurov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 412 (2003) [JETP Lett. **77**, 344 (2003)].
4. L. V. Chulkov, in *Abstracts of VIII International Conference on Nucleus–Nucleus Collisions, Moscow, 2003* (JINR, Dubna, 2003), p. 10.
5. B. Jonson, in *Abstracts of VIII International Conference on Nucleus–Nucleus Collisions, Moscow, 2003* (JINR, Dubna, 2003), p. 17.
6. D. R. Tilley, C. M. Cheves, J. L. Godwin, *et al.*, Nucl. Phys. A **708**, 3 (2002).
7. D. V. Aleksandrov, E. A. Ganza, Yu. A. Glukhov, *et al.*, Yad. Fiz. **39**, 513 (1984) [Sov. J. Nucl. Phys. **39**, 323 (1984)].
8. A. V. Belozerov, C. Borcea, Z. Dlouhy, *et al.*, Nucl. Phys. A **460**, 352 (1986).
9. B. Parker, K. K. Seth, and R. Soundranayagam, Phys. Lett. B **251**, 483 (1990).

10. M. G. Gornov, Yu. B. Gurov, V. P. Koptev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **45**, 205 (1987) [JETP Lett. **45**, 252 (1987)].
11. A. I. Amelin, M. G. Gornov, Yu. B. Gurov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **51**, 607 (1990) [JETP Lett. **51**, 688 (1990)].
12. N. A. F. M. Poppelier, L. D. Wood, and P. W. M. Glaudemans, Phys. Lett. B **157B**, 120 (1985).
13. A. M. Gorbatov, V. L. Skopich, P. Yu. Nikishov, *et al.*, Yad. Fiz. **50**, 1551 (1989) [Sov. J. Nucl. Phys. **50**, 962 (1989)].
14. M. G. Gornov, Yu. B. Gurov, S. V. Lapushkin, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **62**, 1781 (1998).
15. F. Ajzenberg-Selove, Nucl. Phys. A **490**, 1 (1988).
16. M. G. Gornov, Yu. B. Gurov, P. V. Morokhov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **446**, 461 (2000).
17. H. Weyer, Phys. Rep. **195**, 295 (1990).
18. R. Franke, K. Koshskamper, B. Steinheuer, *et al.*, Nucl. Phys. A **433**, 351 (1985).
19. A. I. Baz', V. I. Gol'danskiĭ, V. Z. Gol'dberg, and Ya. B. Zel'dovich, *Light and Intermediate Nuclei near the Nuclear Stability Boundary* (Nauka, Moscow, 1972).
20. M. G. Gornov, Yu. B. Gurov, S. V. Lapushkin, *et al.*, in *LI Meeting on Nuclear Spectroscopy and Atomic Nucleus Structure* (Sarov, 2001), p. 142.

Translated by V. Sipachev

JETP Letters, Vol. 78, No. 4, 2003, pp. 188–192. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 224–228.
Original Russian Text Copyright © 2003 by Frank, Geltenbort, Kulin, Strepetov.

A Quantum Time Lens for Ultracold Neutrons

A. I. Frank[1,*], P. Geltenbort[2], G. V. Kulin[1], and A. N. Strepetov[3]

[1] *Joint Institute for Nuclear Research, Dubna, Moscow region, 141980 Russia*
[2] *Institute Laue–Langevin, BP 156-38042, Grenoble Cedex 9, France*
[3] *Kurchatov Institute, Russian State Scientific Center, Moscow, 123182 Russia*
* *e-mail: frank@nf.jinr.ru*

Received May 5, 2003; in final form, July 17, 2003

The possibility of creating a time lens, an analogue of the zone plate in X-ray optics, for ultracold neutrons is experimentally demonstrated. The neutron energy was changed by means of a purely quantum effect: the phase modulation of a neutron wave at a variable modulation frequency. The modulator was a phase grating with variable spatial period moving across the neutron beam. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 03.75.Be

Quantum experiments in neutron optics and the time focusing of neutrons. In recent years, a number of experiments were devoted to nonstationary action on slow neutrons [1–6] and cold atoms [7–9].[1] As is known, a distinctive feature of nonstationary action consists in changing the energy of a state. In the absence of fields, a change in the energy of a slow particle implies alteration of the classical velocity. The possibility of using this circumstance to create a neutron time lens and to focus neutrons in the time domain was considered in [10]. Below we present the results of our first experiments, demonstrating the possibility of implementing this idea.

The principle of time focusing is shown in Fig. 1. Let neutrons with velocities distributed within a certain interval be emitted at a time $t = 0$ from point $x = 0$ in the positive direction of the **X** axis. Accordingly, the time of arrival at the observation point $x = L$ is also distributed within a certain interval on the time scale. Time focusing means satisfying the condition of simultaneous arrival of all neutrons at the observation point. In order achieve this goal, it is necessary to act upon the neutron velocity by means of a certain device, called a time lens and situated at point $x = a$, so as to provide for the validity of the relation

$$\frac{a}{v_a} + \frac{b}{v_b} = t_0, \quad a + b = L. \tag{1}$$

Here and below, v_a and v_b are the velocities of a neutron before and after the time lens, respectively, and t_0 is the total time of flight (TOF).

Focusing condition (1) will be fulfilled if the time lens transfers to each neutron an energy

$$\Delta E(t) = \frac{m}{2} = \left|\left(\frac{b}{t_0 - t}\right)^2 - \left(\frac{a}{t}\right)^2\right|, \quad t = \frac{a}{v_a}. \tag{2}$$

Obviously, a pulse from any real source has a finite duration τ. The pulse duration Θ at the point of observation (or time imaging) is finite as well. By analogy with geometric optics, we can introduce the notion of time magnification M. Then, for a relatively small energy transfer $|\Delta E| \ll E$, we can use the well-known formula of a thin lens [10]:

$$M = \Theta/\tau = -b/a. \tag{3}$$

Similar results were obtained in [11] for the time focusing of atomic beams. As for the focusing of neutrons on the time scale, this possibility would be very

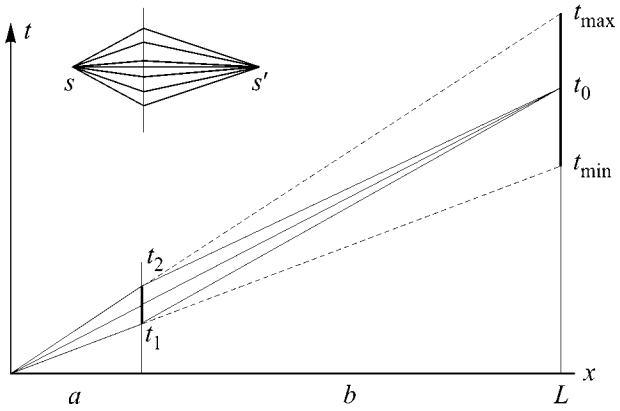


Fig. 1. Schematic diagram illustrating the idea of time focusing.

[1] A brief review of theoretical works on the subject under consideration can be found, e.g., in [6].

important in connection with the old problem of accumulating ultracold neutrons (UCNs) generated by a pulsed source [12, 13] (for more detail, see [10]).

A quantum modulator as a time lens. The function of a time lens can be performed by a quantum modulator. By this we imply a device capable of periodically (with a period of T) acting upon the amplitude or phase of the initial plane wave. In a small vicinity of this device, the wave function has the following form:

$$\Psi(x, t) \cong f(t)e^{i(k_0 x - \omega_0 t)}, \quad k_0^{-1} < x \ll vT. \tag{4}$$

In the entire right-hand half-space, the wave state is represented by a nonstationary superposition of waves with the energies $\hbar\omega_n$ and the wave vectors k_n [14]:

$$\Psi(x, t) = \sum_{n=-\infty}^{\infty} c_n e^{i(k_n x - \omega_n t)}, \quad x > 0, \tag{5}$$

where

$$\begin{aligned} \omega_n &= \omega_0 + n\Omega, \\ k_n &\cong k_0(1 + n(\gamma/2)), \quad \gamma = \Omega/\omega_0 \ll 1, \end{aligned} \tag{6}$$

$\Omega = 2\pi/T$, and the amplitudes c_n are given by the coefficients of expansion of the modulation function $f(t)$ into a Fourier series in frequencies $n\Omega$. In a certain sense, the action of the quantum modulator is analogous to that of the usual grating. The difference is that the usual diffraction is described by the Fourier transform of variables $x \longleftrightarrow k$, while the quantum modulator provides for a mutual transformation of the quantities $t \longleftrightarrow \omega$.

In order to provide for the time focusing effect, fast neutrons reaching the time lens within a time smaller than $t_c = t_0 a/L$ should be slowed down, while slow neutrons should be accelerated (Fig. 1). Since all waves in superposition (5) possess different energies and wave vectors $\mathbf{k_n}$, only one of these waves can meet condition (2) at any time. Thus, only a wave of one (nonzero) order is involved in the formation of a time image at each moment of time, the other waves forming a background. The same situation takes place in X-ray optics [15], where a zone plate is widely used as the lens.

The maximum intensity is usually inherent in waves of the lowest order. For first-order waves, the law of frequency variation with time is determined by the expression $\Omega(t) = \Delta E(t)/\hbar$, where $\Delta E(t)$ is given by formula (2). Within the first half of the time interval, waves of the –1 order are employed, while the rest of the time is devoted to waves of the +1 order. Preference should be given to the phase modulation, whereby $f(t) = \exp[i\varphi(t)]$. This modulator changes only the phase of a transmitted wave, not decreasing the total wave intensity. In practice, it is convenient to modulate the phase by π after each half period. In this case, the Fourier coefficients in expansion (5) are

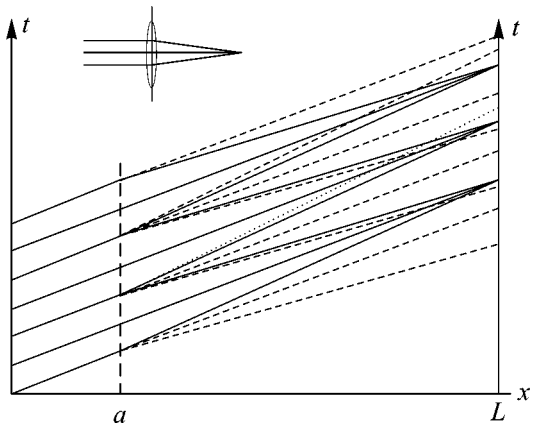


Fig. 2. A coordinate versus time scheme of the experiment with the time lens.

$$c_n = 2/i\pi n, \quad n = 2s - 1, \tag{7}$$

so that waves of all even orders (including zero order) are missing from the resulting spectrum. The intensity of the first-order waves is $|c_1|^2 \approx 0.4$.

Since we are interested only in a change of the particle energy, the phases of modulation at various points of the beam cross section can be different. For this reason, the possibility of using a modulator in the form of a phase grating moving across the beam can also be taken into consideration [10]. The related physical phenomena have been studied previously [14, 16]. It was demonstrated that a phase grating possessing sufficiently large velocity $\mathbf{V}$ at not too small a spatial grating period α ($\alpha k \gg 1$) acts quite similarly to a modulator with the modulation frequency V/α. The phenomenon of energy quantization upon UCN diffraction on a moving phase grating was recently observed in experiment [6].

Description of experiment and results. Our aim was to demonstrate the possibility of implementing the idea of a time lens based on the quantum modulation of a neutron wave. The formation of a time image of the point source is illustrated in Fig. 1. To demonstrate the operation of such a lens, we can use parallel rays from an infinitely remote source. In the spatiotemporal representation, this experimental scheme corresponds to monochromatic neutrons emitted from a stationary source (Fig. 2). In this case, a distribution uniform in time (but, naturally, statistical) of the initial neutrons is transformed so that a temporal group of neutrons appears at the detection point L. The lens can operate in a cyclic regime, focusing the neutrons entering the device within a certain period of time T_{cycl}. Figure 2 presents three such periods. This very scheme was implemented in the experiment.

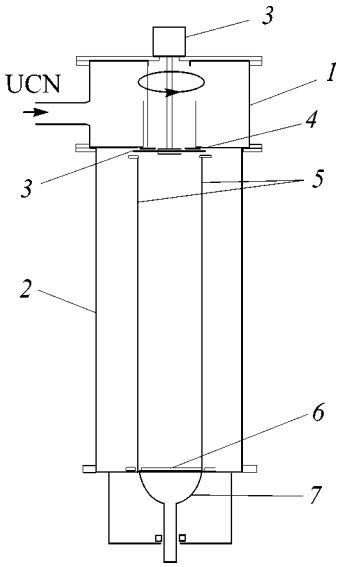


Fig. 3. A schematic diagram of the experimental setup: (*1*) UCN entrance chamber; (*2*) vacuum chamber; (*3*) rotating phase grating; (*4*) interference filter-monochromator; (*5*) neutron guide; (*6*) scintillator; (*7*) photomultiplier.

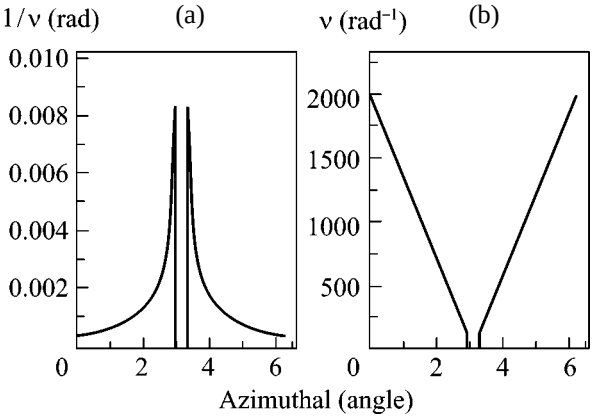


Fig. 4. Plots of (a) the angular distance between grooves of the phase grating and (b) the inverse quantity (spatial frequency) versus the azimuthal angle measured on the grating surface.

The action on the neutron energy was provided by diffraction on a rotating phase grating. The grating, manufactured by lithography on the surface of a silicon disc with a diameter of 152 mm and a thickness of 0.6 mm, had the form of radial grooves. Owing to the refraction in silicon, the phases of neutron waves passing through the adjacent elements (groove plus ridge) of the grating acquired a difference $\Delta\varphi$ determined by the formula

$$\Delta\varphi = k(1-n)d, \qquad (8)$$

where k is the wavenumber, n is the refractive index, and d is the groove depth. In our experiments, the groove depth was chosen to be $d = 0.14$ μm, which corresponded to $\Delta\varphi = \pi$. During rotation of the grating, the phase of a transmitted wave was modulated at each point of the beam cross section with a frequency of $\Omega = 2\pi V/\alpha$. As a result, the initial monochromatic spectrum was transformed into a discrete spectrum according to (5) and (6) with the partial wave amplitudes determined by formula (7).

The experiment was performed using the source of cold and ultracold neutrons at the Institute Laue–Langevin (Grenoble, France) [17]. A general schematic of the experimental setup, representing a modified UCN spectrometer [18, 19], is presented in Fig. 3. UCNs supplied from the source were fed to the device through a cylindrical channel with internal and external diameters of 110 and 130 mm, respectively, and monochromated by a neutron interference filter, a neutron analogue of the Fabry–Perot interferometer [18–20]. The filter transmitted neutrons with a narrow spectrum of vertical velocities representing a peak centered at 4.52 m/s with a width (FWHM) of 0.085 m/s.

Most of the output channel cross section was closed by a diaphragm with a 30° angular window. Past this window, the beam of monochromated neutrons struck the grating. Driven by an electric motor, the grating could rotate about the vertical axis at a rate of 5820 rpm, so that the linear velocity at the site of neutron beam passage was about 36 m/s.

The distance between grooves of the grating was chosen so as to provide that the spatial frequency ν (inverse to the grating period) would be a linear function of the azimuthal angle measured on the grating surface (see Fig. 4), varying from ~2000 rad^{-1} to ~130 rad^{-1}. In the region where the calculated angular frequency was smaller than 130 rad^{-1}, the structure was absent.

When the grating was rotated, neutrons were transmitted at each moment of time through a certain small region of the grating characterized by its period and the corresponding modulation frequency. At a constant rotation rate, the modulation frequency varied so as to ensure the focusing condition during a time equal to the period of rotation (representing the cycle of measurement). During the first half of this interval, the time lens focused neutrons losing their energy (with $\Delta E(t) = -\hbar\Omega(t)$), and in the second half it focused the accelerated neutrons (with $\Delta E(t) = +\hbar\Omega(t)$). Neutrons corresponding to the higher orders formed the background. The maximum modulation frequency amounted to $\Omega_{max} = 1.5 \times 10^7$ rad/s, which corresponded to an energy transfer of $\Delta E_{max} = 10$ neV.

A mirror neutron guide situated under the grating transported UCNs to a scintillation detector. The distance from the monochromator to the detector (TOF base) was 70 cm and the TOF of monochromatic neutrons was about 140 ms.[2] The measurements were per-

[2] This time and the neutron energy spectrum past the monochromator were measured in a separate experiment not described here.

formed in TOF mode. The motor driving the grating was equipped with a transducer generating master pulses at a certain phase of the rotation cycle. The TOF was measured between the master pulse and the moment of detection.

Figure 5 shows an experimental distribution of the neutron arrival times measured at the detector, plotted on a time scale equal to the grating rotation period. As is clearly seen, the distribution is by no means uniform and exhibits a pronounced peak, which is evidence of the focusing effect. It should be recalled that the moments when neutrons entered the instrument were absolutely random.

The measurements were performed for several grating rotation rates. As the rotation rate was decreased, the peak exhibited a shift. Simultaneously, the peak increased in width and decreased in height, completely vanishing at a grating rotation rate below 2000 rpm. The efficiency of focusing was about 17% (against the maximum possible value of 40% calculated for the intensity of the first-order waves). The decrease in the focusing efficiency is probably explained by insufficiently high quality of the grating, which was confirmed by special tests.

In forthcoming experiments, we hope to increase the efficiency of the time lens operating on the principle of π-phase modulation. Theoretically, it is possible to create a lens with efficiency approaching 100%. This would correspond to a modulation function of the type $f(t) = \exp[i\varphi(t)]$, where $\varphi = 2\pi t/T$, T is the period, and $0 < t < T$.

Brief discussion of results. We have described a new experiment in the field of quantum neutron optics. Such experiments usually have a character of demonstration. However, it was already pointed out [6] that the significance of demonstration experiments is not restricted to solving purely pedagogical tasks. It is also significant that such works can stimulated the development of new methods and devices based on quantum principles. We believe that the present work confirms the validity of this statement.

Previously [14, 16], the operation of a quantum modulator was theoretically studied and the effect of energy quantization upon diffraction on a moving grating was predicted. Later, the idea of creating a time lens for neutrons was formulated [10], while the possibility of using nonstationary quantum effects for this purpose was viewed upon rather as eccentric. Recently, however, the phenomenon of energy quantization upon diffraction on a moving grating was experimentally demonstrated in [6]. In conclusion to that paper, we suggested that "this effect, which has been observed for the first time, may be used for controlled changing of neutron energy in other experiments. The idea of neutron time-focusing … now looks more realistic." In the present paper, the possibility of realization of a quantum time lens for neutrons has been experimentally

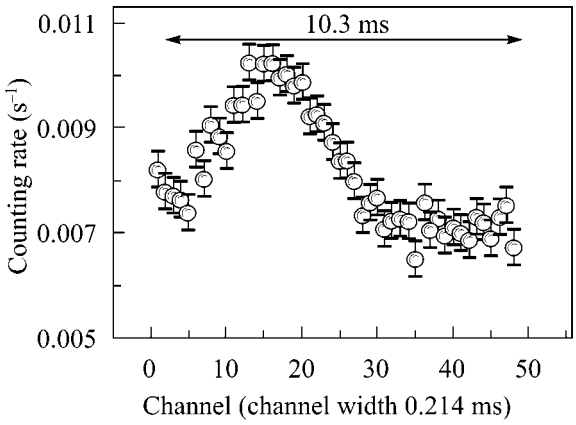


Fig. 5. A time focusing peak representing a distribution of the neutron arrival times plotted on a time scale equal to the grating rotation period.

demonstrated. Thus, we can move from prediction and demonstration of this effect to its practical application.

Evidently, in continuing development of the methods of time focusing, we may suggest using a time lens to create an intense source of ultracold neutrons [11, 12, 10].

The authors are grateful to V. G. Nosov and E. Kats for very fruitful discussions and to V. Yu. Kireev for his help in manufacturing gratings. This study was supported by the Russian Foundation for Basic Research (project no. 00-02-17172) and the INTAS Foundation (grant no. 00-00043).

REFERENCES

1. G. Badurek, H. Rauch, and D. Tuppinger, Phys. Rev. A **34**, 2600 (1986).
2. J. Summhammer, K. A. Hamacher, H. Kaiser, *et al.*, Phys. Rev. Lett. **75**, 3206 (1995); J. Summhammer, Phys. Rev. A **54**, 3155 (1996).
3. J. Felber, R. Gähler, C. Rausch, and R. Golub, Phys. Rev. A **53**, 319 (1996).
4. Th. Hils, J. Felber, R. Gähler, *et al.*, Phys. Rev. A **58**, 4784 (1998).
5. E. Iolin, B. Farago, F. Mezei, *et al.*, Physica B (Amsterdam) **241**–**243**, 1213 (1998).
6. A. I. Frank, S. N. Balashov, I. V. Bondarenko, *et al.*, Phys. Lett. A **311**, 6 (2003).
7. A. Stean, P. Szriftgiser, P. Desbiolles, and J. Dalibard, Phys. Rev. Lett. **74**, 4972 (1995).
8. P. Szriftgiser, D. Guéry-Odelin, M. Arndt, and J. Dalibard, Phys. Rev. Lett. **77**, 4 (1996).
9. S. Bernet, M. K. Oberthaler, R. Abfalterer, *et al.*, Phys. Rev. Lett. **77**, 5160 (1996).
10. A. Frank and R. Gaehler, in *Proceedings of IV International Seminar on Interaction of Neutrons with Nuclei, ISINN-4, JINR, Dubna* (1996), JINR E3-96-336, p. 308; A. I. Frank and R. Gaehler, Yad. Fiz. **63**, 605 (2000) [Phys. At. Nucl. **63**, 545 (2000)].

11. M. Arndt, P. Szriftgiser, J. Dalibard, and A. M. Steane, Phys. Rev. A **53**, 3369 (1996).

12. F. L. Shapiro, *Neutron Research* (Nauka, Moscow, 1976).

13. H. Rauch, Nucl. Instrum. Methods Phys. Res. A **491**, 478 (2002).

14. A. I. Frank and V. G. Nosov, Yad. Fiz. **57**, 1029 (1994) [Phys. At. Nucl. **57**, 968 (1994)].

15. *Proceedings of International Symposium on X-Ray Microscopy, Göttingen, 1983*, Ed. by G. Schmahl and D. Rudolph (Springer, Berlin, 1984), Springer Series in Optical Science, Vol. 43.

16. A. I. Frank and V. G. Nosov, Phys. Lett. A **188**, 120 (1994).

17. A. Steyerl, H. Nagel, F. Schriber, *et al.*, Phys. Lett. A **116**, 347 (1986).

18. I. V. Bondarenko, V. I. Bondarchuk, S. N. Balashov, *et al.*, Yad. Fiz. **62**, 775 (1999) [Phys. At. Nucl. **62**, 721 (1999)].

19. I. V. Bondarenko, A. I. Frank, S. N. Balashov, *et al.*, Nucl. Instrum. Methods Phys. Res. A **440**, 591 (2000).

20. A. Steyerl, W. Drexel, S. S. Malik, and E. Gutsmiedle, Physica B (Amsterdam) **151**, 36 (1988).

Translated by P. Pozdeev

JETP Letters, Vol. 78, No. 4, 2003, pp. 193–197. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 229–233.
Original Russian Text Copyright © 2003 by Golovan', Mel'nikov, Konorov, Fedotov, Gavrilov, Zheltikov, Kashkarov, Timoshenko, Petrov, Li, Yakovlev.

Efficient Second-Harmonic Generation by Scattering from Porous Gallium Phosphide

L. A. Golovan'[1,*], V. A. Mel'nikov[1], S. O. Konorov[1], A. B. Fedotov[1], S. A. Gavrilov[2], A. M. Zheltikov[1], P. K. Kashkarov[1], V. Yu. Timoshenko[1], G. I. Petrov[3], L. Li[3], and V. V. Yakovlev[3]

[1] *Physics Department, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
**e-mail: leo@vega.phys.msu.su*
[2] *Moscow State Institute of Electronic Technology, Moscow, 103498 Russia*
[3] *Physics Department, University of Wisconsin–Milwaukee, 53211 Milwaukee, WI, USA*
Received June 27, 2003; in final form, July 22, 2003

We experimentally study second-harmonic generation by femtosecond Cr: forsterite-laser radiation scattered on the surface of porous gallium phosphide with characteristic pore sizes and distances between the pores comparable with the second-harmonic wavelength. The intensity of the second-harmonic signal from samples with initial crystallographic surface orientations (110) and (111) is more than an order of magnitude higher than the intensity of the second harmonic generated in reflection from single-crystal gallium phosphide. The efficiency of second-harmonic generation by macroporous gallium phosphide substantially increases as the pump wavelength becomes shorter. The influence of light localization and scattering effects on the enhancement of second-harmonic generation and polarization properties of the second-harmonic is discussed. © *2003 MAIK "Nauka/Interperiodica"*.

PACS numbers: 42.65.Ky; 42.70.Nq; 78.67.-n

Advances in the nonlinear optics of nanostructures open the way for creation of new unique materials for laser physics and nonlinear optics, as well as for the development of highly efficient photonic devices with the use of these materials [1, 2]. Modern nanotechnologies, as demonstrated by recent experiments with mesoporous silicon [3–6], allow a radical modification of optical properties of materials and creation of structures with a strong artificial birefringence, offering new solutions for the phase matching of nonlinear-optical interactions [7].

Applications of porous silicon in optics, including photonics and nonlinear optics, are, however, limited, first, by the absorption of silicon, which restricts the efficient use of silicon to the infrared range; second, by the low quantum yield in the case of mesoporous structures; and, third, by the fact that porous silicon remains a centrosymmetric material, where some nonlinear-optical processes, such as second-harmonic generation (SHG), for example, are forbidden in the dipole approximation. In this respect, porous gallium phosphide holds much promise for optical applications. Due to a wider band gap, GaP can be employed in the yellow and red visible spectral ranges (for wavelengths exceeding 550 nm). The noncentrosymmetric crystal lattice of this material gives rise to a high dipole quadratic nonlinear susceptibility (~200 pm/V), which is two orders of magnitude higher than this susceptibility for most of the crystals employed for frequency doubling [8].

Porous gallium phosphide is of particular interest also in the context of light localization [9–11]. This effect, analogous to the Anderson localization of electrons in a medium with defects, is due to the interference of waves scattered by inhomogeneities [12]. The time of radiation–matter interaction increases under these conditions, leading to the enhancement of many optical processes, including nonlinear-optical phenomena. Nonlinear-optical interactions in disordered dielectric media have been studied earlier both theoretically [13–15] and experimentally. Experimental studies have been performed, in particular for layers of lithium niobate powder [15] and a powder laser [16]. However, all the aspects of the role of disordering in nonlinear-optical processes still have to be understood. In particular, no comparison with crystal media has yet been performed. Efficiencies of such processes have not been estimated either.

Thus, porous GaP offers much promise for the creation of new nonlinear-optical components. Nevertheless, many of the nonlinear-optical properties of this material are still to be understood. In earlier work [17, 18], SHG by Nd:YAG-laser radiation propagating through a porous GaP film was investigated. These studies have shown that the SHG efficiency in porous GaP with (100) and (111) surface orientations increases and the orientation dependence changes with respect to crystal

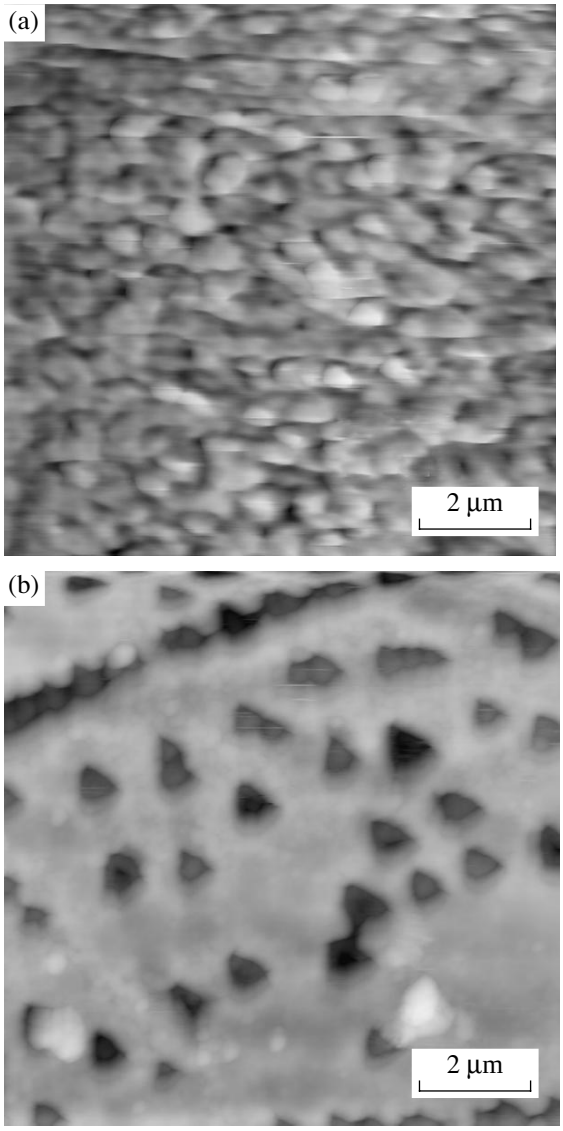


Fig. 1. Atomic-force-microscopy surface images of macroporous gallium phosphide produced on a substrate with (a) the (110) and (b) (111) surface orientation. The sizes of the images are 10×10 μm.

GaP. Little has been known until now about the influence of light scattering in macroporous GaP layers on the SHG efficiency.

In this paper, we study second-harmonic generation in near-surface layers of macroporous GaP. These layers were produced by electrochemically etching *n*-GaP doped with Te with a concentration of 3×10^{17} cm^{-3}, having initial surface orientations (110) and (111), in a 0.5-M water solution of H_2SO_4 and a 1 : 1 ethanol solution of HF, respectively. One of the films with the (110) surface orientation was cleaved from a substrate. The thicknesses of porous layers were 40 and 10 μm, respectively. Macroporous GaP layers employed in our experiments had a comparatively low optical quality of both surfaces. No characteristic interference fringes was observed in infrared and visible transmission spectra. Surface images of macroporous GaP made with the use of atomic-force microscopy are presented in Fig. 1. As can be seen from these images, the size of inhomogeneities [pores for the (111) surface and nanocrystals for the (110) surface] is on the order of 600 nm. Some anisotropy of crystal remainders is observed for the film with the (110) surface orientation.

The second harmonic was generated in our experiments with 60-fs pump pulses of Cr: forsterite-laser radiation with the wavelength $\lambda = 1250$ nm. The pulse repetition rate was 25 MHz, and the pulse energy was 6 nJ. The advantage of Cr: forsterite-laser radiation for our experiments is that neither fundamental radiation nor its second harmonic is absorbed by gallium phosphide. Pump radiation was focused on the sample along the normal to its surface with a lens having the focal length $F = 4$ mm and the numerical aperture $N.A. = 0.5$. With these parameters, the laser beam diameter on the sample was about 1.5 μm, and the confocal parameter $\lambda/\pi(N.A.)^2$ was estimated as 2 μm. Thus, an area with a length of 4 μm provided the main contribution to the second-harmonic signal. Scattered radiation collected with the same lens was directed to a detection system (a Jobin Yvon Triax 320 spectrometer) with the use of a dichroic mirror. To measure orientation dependences of the second-harmonic signal, the polarization of the pump field was rotated (with a half-wave plate) simultaneously with an analyzer (the Glan prism) for the second harmonic. Polarization rotation angles for the pump field and the second harmonic were equal to each other, with the analyzer being either parallel or perpendicular to the polarization of pump radiation.

Many important aspects related to the influence of light scattering on SHG can be understood by varying the ratio of the second-harmonic wavelength to the size of nanocrystals. Such SHG experiments were performed with the use of an optical parametric oscillator. We employed a nanosecond Solar Laser Systems laser, consisting of a Nd:YAG master oscillator, a frequency tripler for 1.06-μm radiation, and an optical parametric oscillator. An idler wave of the optical parametric oscillator with a wavelength smoothly tunable within the spectral range from 1 up to 1.5 μm was employed as a pump in our experiments. The laser system generated 3-ns pulses with an energy up to 10 mJ at a pulse repetition rate of 20 Hz. The *p*-polarized pump radiation was focused with a 5-cm-focal-length lens on a sample at an incidence angle of 45°. A porous GaP layer on a crystal substrate with the (110) orientation was studied in these experiments. Pump radiation was polarized perpendicular to the [001] crystallographic axis. Second-harmonic radiation generated through scattering and reflection from the studied samples was collected with a lens, analyzed with the use of an MDR-6 mono-

chromator, and detected with a photoelectric multiplier within the wavelength range of 0.5–0.75 μm.

The results of orientation measurements performed on the second harmonic are presented in Fig. 2. Clearly pronounced orientation dependences are observed for crystal GaP. This crystal belongs to the $\bar{4}3m$ point group, and its quadratic nonlinear susceptibility tensor is characterized by a single independent element $\chi^{(2)}_{123} = \chi^{(2)}_{231} = \chi^{(2)}_{213}$. The following relations can be easily derived in this case for the (110) and (111) surfaces in the regime of normal incidence [19]:

$$I_{\parallel}^{(11)} \sim \sin^4\phi\cos^2\phi, \quad (1)$$

$$I_{\perp}^{(11)} \sim \sin^2\phi(3\cos^2\phi - 1)^2, \quad (2)$$

$$I_{\parallel}^{(111)} \sim \cos^2 3\Phi, \quad (3)$$

$$I_{\perp}^{(111)} \sim \sin^2 3\Phi, \quad (4)$$

where the angles ϕ and Φ are measured from the [001] and $[\bar{1}\bar{1}2]$ crystallographic directions, respectively. As can be seen from Fig. 2, the orientation dependences of the second-harmonic intensity for crystal GaP agree well with Eqs. (1)–(4).

For macroporous GaP, the orientation dependence, however, becomes isotropic: the second-harmonic intensity is independent in this case of the relative orientation of the polarizer and the sample. The second-harmonic intensity is substantially higher under these conditions than in the case of crystal GaP. This SHG enhancement reaches an order of magnitude for macroporous GaP produced on the (110) surface (Fig. 1a) and nearly two orders of magnitude for macroporous GaP produced on the (111) surface (Fig. 1b). It should be noted that only some fraction of scattered second-harmonic radiation was collected in our experiments. Receiving the second harmonic within a larger solid angle would increase the ratio of second-harmonic intensities for macroporous and crystal GaP.

Figure 3 displays the ratio of the intensities of the second harmonic generated in macroporous and crystal GaP samples as a function of the pump wavelength. As can be seen from this dependence, the second-harmonic signal from a macroporous sample is higher than the second harmonic from crystal GaP for all the radiation wavelengths employed in our experiments. Importantly, the ratio of these signals substantially depends on the pump wavelength, decreasing by more than an order of magnitude with the increase in the wavelength.

Since the surface of the films employed in our experiments had a rather low optical quality, it is unlikely that the effects described above are due to the

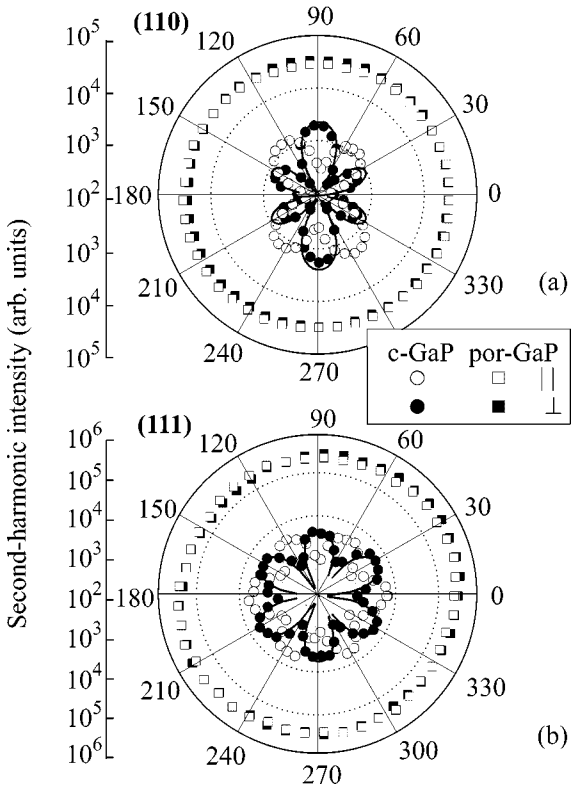


Fig. 2. Orientation dependences of the second-harmonic signal measured with parallel (open circles) and crossed (filled circles) polarizers for (circles) crystal and (squares) macroporous gallium phosphide with (a) the (110) and (b) (111) surface orientation. The results of calculations performed with (dotted line) Eqs. (1) and (3) and (solid line) Eqs. (2) and (4) are also shown.

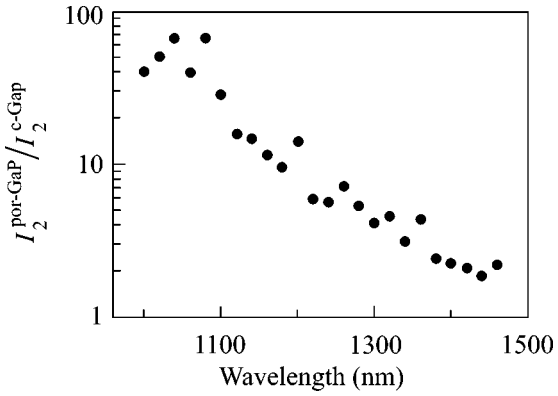


Fig. 3. The intensity of the second harmonic generated through scattering from macroporous GaP normalized to the intensity of the second harmonic generated in reflection from crystal GaP as a function of the pump wavelength.

reflection of pump or second-harmonic radiation from the porous GaP–crystal GaP interface for the porous layer on a substrate and from the porous GaP–air interface in the case of a free-standing film. In addition, tight

focusing reduced the influence of phase matching on SHG in experiments performed with a Cr : forsterite laser.

The results of these experiments suggest an important role of scattering in SHG. The isotropic character of orientation dependences obtained in our experiments may be mainly attributed to the regime of SHG where both pump radiation and the second harmonic propagate at different angles in a macroporous structure, interacting with crystal GaP surfaces having different orientations. Both of these factors change orientation dependences. The signal detected in our experiments is thus a sum of signals from different points in the entire volume contributing to second-harmonic generation. Since the signal from each point is characterized by its own orientation dependence, distorted by subsequent reflections, the total second-harmonic signal, collected within a large solid angle, is nearly completely depolarized. The increase in the interaction time of pump radiation with a macroporous structure due to multiple scattering is apparently responsible for the enhancement of SHG observed in our experiments.

The ratio of the characteristic size of spatial inhomogeneities to the radiation wavelength λ is the key parameter controlling the regime of linear and nonlinear scattering of electromagnetic radiation in a disordered medium. The characteristic size of inhomogeneities (the mean size of pores and crystal remainders in the case of porous materials) determines the photon free-path length l for a given radiation wavelength and influences the regime of radiation focusing. The condition $l/\lambda \sim 1$ defines the criterion of Anderson localization of light in disordered media [10]. The strong dependence of the SHG efficiency in macroporous GaP on the pump wavelength (Fig. 3) may, therefore, indicate the tendency of localization enhancement toward the Anderson localization regime. Analysis of surface images of porous gallium phosphide samples investigated in our experiments (Fig. 1) shows that the characteristic distance between nanocrystals and the typical pore size for these samples are close to the wavelength of the second harmonic of both Cr : forsterite-laser radiation and the idler wave of the optical parametric oscillator.

The above hypothesis of radiation localization enhancement in a macroporous material is consistent with the findings of the earlier experiments devoted to nonlinear-optical phenomena in mesoporous silicon [6]. These experiments have demonstrated, in particular, that the enhancement of nonlinear-optical processes in mesoporous silicon is sensitive to the parameters of the initial silicon wafer and the regime of electrochemical etching, which determine the sizes of the shape of pores and nanocrystals, eventually controlling the regime of radiation localization in the sample. These factors may be responsible for the noticeable difference in SHG enhancement observed for surfaces with the (110) and (111) orientations.

Our experiments on second-harmonic generation in scattering layers of porous gallium phosphide thus show that nanostructures where the characteristic sizes of inhomogeneities are comparable with the wavelength of second-harmonic radiation enhance second-harmonic generation by more than an order of magnitude relative to SHG in crystal gallium phosphide. Radiation at the frequency of the second harmonic is totally depolarized under these conditions. Enhancement of second-harmonic generation in scattering macroporous gallium phosphide substantially depends on the wavelength of pump radiation, with a decrease in the pump wavelength increasing the ratio of second-harmonic intensities for macroporous and crystal GaP. Enhancement of second-harmonic generation was observed in the above-described experiments in a disordered spatially nonuniform structure with parameters meeting the light-localization criterion. The results of our experiments indicate the possibility of creating efficient frequency converters and visualizers of infrared radiation based on porous gallium phosphide.

We are grateful to V. Shcheslavskiĭ for assistance in experimental work. This study was supported by the President of the Russian Federation Grant MD-42.2003.02, the Russian Foundation for Basic Research (project nos. 02-02-17259, 03-02-16929, and 02-02-17098), the Civilian Research and Development Foundation (project no. RE2-2369), NSF–ECS (project no. 9984225), the Volkswagen Foundation (project I/76 869), the European Research Office of the US Army (Contract no. 62558-03-M-0033), and NIH (project no. R21RR16282), as well as by programs of the Ministry of Industry, Science, and Technologies of the Russian Federation.

REFERENCES

1. A. G. Cullis, L. T. Canham, and P. D. J. Calcott, J. Appl. Phys. **82**, 909 (1997).
2. P. K. Kashkarov, L. A. Golovan, A. B. Fedotov, *et al.*, J. Opt. Soc. Am. B **19**, 2273 (2002).
3. O. G. Sarbeĭ, E. K. Frolova, R. D. Fedorovich, and D. B. Dan'ko, Fiz. Tverd. Tela (St. Petersburg) **42**, 1205 (2000) [Phys. Solid State **42**, 1240 (2000)].
4. N. Künzner, D. Kovalev, J. Diener, *et al.*, Opt. Lett. **26**, 1265 (2001).
5. L. P. Kuznetsova, A. I. Efimova, L. A. Osminkina, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **44**, 780 (2002) [Phys. Solid State **44**, 811 (2002)].
6. L. A. Golovan, L. P. Kuznetsova, A. B. Fedotov, *et al.*, Appl. Phys. B **76**, 429 (2003).
7. L. A. Golovan, V. Yu. Timoshenko, A. B. Fedotov, *et al.*, Appl. Phys. B **73**, 31 (2001).
8. R. L. Sutherland, *Handbook on Nonlinear Optics* (Marcel Dekker, New York, 1996).
9. F. J. P. Schuurmans, D. Vanmaekelbergh, J. van de Lagemaat, *et al.*, Science **284**, 141 (1999).

10. A. Lagendijk, J. Gómez Rivas, A. Imhof, *et al.*, in *Photonic Crystals and Light Localization in the 21st Century*, Ed. by C. M. Soukoulis (Kluwer, Dordrecht, 2001), p. 447.
11. J. Gómez Rivas, A. Lagendijk, R. W. Tjerkstra, *et al.*, Appl. Phys. Lett. **80**, 4498 (2002).
12. M. Rusek and A. Orlowski, in *Optics of Nanostructured Materials*, Ed. by V. A. Markel and T. F. George (Wiley, New York, 2001), p. 201.
13. V. M. Agranovich and V. E. Kravtsov, Zh. Éksp. Teor. Fiz. **95**, 484 (1989) [Sov. Phys. JETP **68**, 272 (1989)].
14. V. E. Kravtsov, V. M. Agranovich, and K. I. Grigorishin, Phys. Rev. B **44**, 4931 (1991).
15. J. F. de Boer, A. Lagendijk, R. Sprik, *et al.*, Phys. Rev. Lett. **71**, 3947 (1993).
16. M. A. Noginov, S. U. Egarievwe, N. Noginova, *et al.*, J. Opt. Soc. Am. B **15**, 2854 (1998).
17. I. M. Tiginyanu, I. V. Kravetsky, G. Marowsky, *et al.*, Phys. Status Solidi A **175**, R5 (1999).
18. I. M. Tiginyanu, I. V. Kravetsky, J. Monecke, *et al.*, Appl. Phys. Lett. **77**, 2415 (2000).
19. P.-F. Brevet, *Surface Second Harmonic Generation* (Presses Polytech. Univ. Romandes, Lausanne, 1997).

Translated by A. Zheltikov

JETP Letters, Vol. 78, No. 4, 2003, pp. 198–201. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 234–237.
Original English Text Copyright © 2003 by Chernyak, Chertkov, Kolokolov, Lebedev.

Periodic and Quasi-Periodic Compensation Strategies of Extreme Outages Caused by Polarization Mode Dispersion and Amplifier Noise¶

V. Chernyak[1], M. Chertkov[2], I. Kolokolov[2, 3], and V. Lebedev[2, 3, *]

[1]*Corning, NY 14831, USA*

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

[3]*Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117334 Russia*

**e-mail: lebede@landau.ac.ru*

Received July 16, 2003

The effect of birefringent disorder on the bit error rate (BER) in an optical fiber telecommunication system subject to amplifier noise may lead to extreme outages, related to anomalously large BER values. We analyze the probability distribution function of BER for various strategies of polarization mode dispersion compensation. A compensation method is proposed that is capable of more efficient extreme suppression of outages, which leads to substantial improvement of the fiber system performance. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 42.79.Sz; 42.81.Gs; 05.40.Ca

Polarization mode dispersion (PMD) is an essential impairment for modern optical fiber systems [1–3]. Therefore, dynamical PMD compensation has become an important subject in modern communication technology [4–7]. Optical noise generated in optical amplifiers represents another impairment that may not be reduced or compensated and, therefore, should also be considered in any evaluation of fiber system performance [8]. The bit error rate (BER) calculated for a given realization of birefringent disorder by means of averaging over the amplifier noise statistics constitutes an appropriate object for characterizing the joint effect of the two impairments. In two preceding papers [9, 10], we have demonstrated that the probability of extreme outages (BER values much higher than typical) is substantially larger than one could expect from naive Gaussian estimates singling out effects of either of the two impairments. The natural object of interest is the probability distribution function (PDF) of BER and, specifically, the PDF tail corresponding to anomalously large BER. In [9] we have developed a consistent theoretical approach to calculating this tail. The case when no compensation is applied and also the effect of the simplest "setting of the clock" compensation on the PDF tail suppression have been discussed in [9]. Our investigation was then extended to study effects of the standard first- and higher-order compensations on extreme outages [10]. In the present letter we propose a compensation scheme that appears to be more efficient in reducing the extreme outages compared to the traditional high-order compensation scheme with the same number of compensating degrees of freedom.

We consider the return-to-zero (RZ) modulation format, when optical pulses are well separated in time t and thus can be analyzed as individual objects. We represent the pulse intensity measured at the system output as

$$I = \int dt G(t) |\mathcal{H}\mathbf{\Psi}(t)|^2, \qquad (1)$$

where $G(t)$ is a convolution of the electrical (current) filter function with the sampling window function. The two-component complex field $\mathbf{\Psi}(t)$ describes the output optical signal (the components correspond to two polarization states of the signal). The linear operator $\mathcal{H}$ in Eq. (1) represents optical filtering; it may also account for a compensating device. The compensating part of the linear operator, $\mathcal{H}_c$, is applied first, i.e., before filtering described by $\mathcal{H}_f$, resulting in $\mathcal{H} = \mathcal{H}_f\mathcal{H}_c$. Ideally, I takes two distinct values depending on whether the information slot is vacant or filled. However, the impairments enforce deviations of I from those fixed values. If the output signal intensity exceeds the decision level I_d, then "1" is associated with the slot, otherwise the slot is labeled "0." Sometimes the information is lost; i.e., the initial "1" is detected as "0" at the output or vice versa. The BER is the probability of such events that naturally depends on a specific realization of birefringent disorder in the fiber. BER must be extremely small to guarantee successful system performance. It has been demonstrated in [9] that anomalously high BER originates solely from the "1 $\longrightarrow$ 0"

¶ This article was submitted by the authors in English.

events. We denote the probability of such events by B and study its sensitivity with respect to disorder.

In this letter we restrict ourselves to the linear propagation regime, when the output signal $\boldsymbol{\Psi}(t)$ can be represented as a sum of two contributions: $\boldsymbol{\varphi}$, related to the noiseless evolution of the initial pulse, and the noise-induced part $\boldsymbol{\phi}$. We consider the cases of distributed or, alternatively, lumped amplification with the fiber length Z substantially exceeding the interamplifier separation (span length) within the same scope. $\boldsymbol{\phi}$ becomes a zero-mean Gaussian variable, completely characterized by its two-point correlation function

$$\langle \phi_\alpha(t_1)\phi^*_\beta(t_2)\rangle = D_\xi Z\delta_{\alpha\beta}\delta(t_1 - t_2), \qquad (2)$$

insensitive to particular realizations of birefringent disorder and chromatic dispersion in the fiber. The product $D_\xi Z$ represents the amplified spontaneous emission (ASE) spectral density accumulated along the fiber. The coefficient D_ξ is introduced into Eq. (2) to reveal the linear growth of the ASE factor with Z [8]. The noise-independent part of the signal is governed by

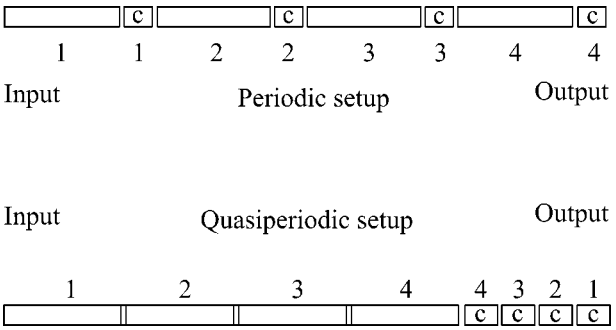
$$\partial_z\boldsymbol{\varphi} - \hat{m}(z)\partial_t\boldsymbol{\varphi} - id(z)\partial_t^2\boldsymbol{\varphi} = 0, \qquad (3)$$

z and d being the coordinate along the fiber and chromatic dispersion. The birefringence matrix can be represented as $\hat{m} = h_j\hat{\sigma}_j$, where h_j is a real three-component field and $\hat{\sigma}_j$ are the Pauli matrices. Averaging over many states of the birefringent disorder that any fiber goes through over time, or (alternatively) over the states of birefringence of different fibers, one finds that $h_j(z)$ is a zero-mean Gaussian field described by

$$\langle h_i(z_1)h_j(z_2)\rangle = D_m\delta_{ij}\delta(z_1 - z_2). \qquad (4)$$

If birefringent disorder is weak, the integral $\mathbf{H} = \int_0^Z dz\mathbf{h}(z)$ coincides with the PMD vector. Thus, $D_m = k^2/12$, where k is the so-called PMD coefficient.

In an operable communication system, typical damage caused by disorder and noise must be small, i.e., typically both impairments can cause only a small distortion to a pulse; thus, the optical signal-to-noise ratio (OSNR) and the ratio of the squared pulse width to the mean squared value of the PMD vector are both large. OSNR can be estimated as $I_0/D_\xi Z$, where $I_0 = \int dt|\Psi_0(t)|^2$ is the initial pulse intensity, the integration being performed over a single slot populated by an ideal pulse, encoding "1." Typically, B fluctuates around B_0, the zero-disorder ($h_j = 0$) value of B. A convenient auxiliary dimensionless object, $\Gamma = (D_\xi Z)\ln(B/B_0)/I_0$, depends on the birefringent disorder and the initial signal shape, as well as the details of the compensation and detection procedures; it is, however, insensitive to noise. Since the OSNR is large, even weak disorder can generate a strong increase in the value of B. This is why a perturbative (with respect to



Scheme of fiber-line-element installation corresponding to the periodic and quasi-periodic compensation strategies.

h) calculation of Γ gives the most essential part of the PDF $\mathcal{S}(B)$ of B. If no compensation is applied, one gets $\Gamma \sim H_3/b$, b being the pulse width, and the initial signal is assumed to be linearly polarized. In the simplest case of "setting of the clock" compensation, one arrives at $\Gamma \sim (H_1^2 + H_2^2)/b^2$. This yields the powerlike tail of the PDF of B [9]. Higher order compensation leads to $\Gamma \sim (H/b)^p$, where p is an integer exceeding by one the degree of compensation, provided no additional cancellations occur, and one gets the following asymptotic expression (tail) for the PDF $\mathcal{S}(B)$ of B [10]:

$$\ln\mathcal{S} = -\mu_p b^2[D_\xi Z\ln(B/B_0)/I_0]^{2/p}/[D_m Z], \qquad (5)$$

where μ_p is a dimensional coefficient. Therefore, as anticipated, compensation suppresses the PDF tail. However, applying high-order compensation is not very efficient, since the decrease in $\mathcal{S}$ is mild as p increases.

The main purpose of this letter is to introduce more efficient compensation strategies with the same number of compensating degrees of freedom. As a first example consider the following "periodic" scheme. One divides the optical line into N segments, each having a length $l = Z/N$, and apply the first-order compensation at the end of each segment (as schematically shown in the upper panel of figure, with "c" denoting the compensating elements). The noise-independent part of the compensated signal for the "periodic compensation" strategy is determined by

$$\mathcal{H}_c\boldsymbol{\varphi} = \exp(i\eta\partial_t^2)\mathcal{H}_{1N}\hat{U}_N\ldots\mathcal{H}_{11}\hat{U}_1\boldsymbol{\Psi}_0(t), \qquad (6)$$

$$\hat{U}_n = T\exp\left[\int_{(n-1)l}^{nl} dzh_j(z)\hat{\sigma}_j\partial_t\right], \qquad (7)$$

$$\mathcal{H}_{1n} = \exp\left[-\int_{(n-1)l}^{nl} dzh_j(z)\hat{\sigma}_j\partial_t\right], \qquad (8)$$

where $\mathbf{\Psi}_0(t)$ is the input signal profile, $\eta = \int_0^Z dz d(z)$ is the integral chromatic dispersion, and the ordered product on the right-hand side of Eq. (6) is taken over all the N segments (T exp is the standard notation for the so-called ordered exponential). The exponential factor $\mathcal{H}_{1n}$ represents the first-order compensation at the end of the nth segment.

This periodic compensation is not very convenient since it requires installation of compensating elements at multiple places along the fiber. However, one can naturally modify this scheme and have the same compensating elements inserted subsequently, but all at once, at the fiber output, as shown in the lower panel of the figure. If the disorder profile h_j is known (technically such end-point measurements are possible via the anti-Stokes refraction technique [11]) one can have an end-point, but multiple, compensation as $\mathcal{H}_c = \prod \mathcal{H}_{1n}$, leading to the following "quasi-periodic" modification of Eq. (6):

$$\mathcal{H}_c \boldsymbol{\varphi} = \exp(i\eta\partial_t^2)\mathcal{H}_{11}\ldots\mathcal{H}_{1N}\hat{U}_N\ldots\hat{U}_1\mathbf{\Psi}_0(t). \quad (9)$$

The natural idea behind this quasi-periodic compensation is obvious: to construct (in the compensating part) the best possible approximation (with a given number of compensating degrees of freedom) for the inverse of the ordered exponential $\hat{U}_N\ldots\hat{U}_1$.

Note that the (quasi-)periodic compensation does not influence the noise-dependent part of the signal; i.e., $\mathcal{H}_c\boldsymbol{\varphi}$ has the same correlation function (2) as $\boldsymbol{\varphi}$. Therefore, one arrives at the same expression $\ln(B/B_0) = \Gamma I_0/(D_\xi Z)$, with a new $\mathbf{h}$-dependent factor Γ. Furthermore, in the main region of interest Γ can be analyzed perturbatively, just as in [9, 10]. Expanding the factors in Eq. (6) up to the second order and making use of Eqs. (1), (2) one derives

$$\Gamma \approx \frac{\mu_2'}{b^2}\sum_{n=1}^{N}\int_{a_n}^{nl} dz \int_{a_n}^{z} dz'[h_1(z)h_1(z') - h_2(z)h_1(z')], \quad (10)$$

where $a_n = (n-1)l$. Here, the dimensionless coefficient μ_2' is related to the output signal chirp produced by an initial chirp and the nonzero integral chromatic dispersion η. As follows from Eq. (9), the same expression (10) is obtained in the second order for the quasi-periodic case. Substituting Eq. (10) into the expression for B and evaluating the PDF of B, with the Gaussian statistics of $\mathbf{h}$ described by Eq. (4), leads to the following expression for the tail of the PDF of B:

$$\mathcal{S}(B)dB \sim \frac{B_0^\alpha dB}{B^{1+\alpha}}, \quad \alpha = \frac{N\pi D_\xi b^2}{2|\mu_2'|D_m I_0}. \quad (11)$$

Equation (11) holds for $\ln(B/B_0) \gg \mu_2' D_m I_0/[D_\xi b^2]$. The exponent α in Eq. (11) contains an additional factor N compared to the expression for the first order end-point compensation, i.e., the (quasi-)periodic compensation makes the tail of $\mathcal{S}(B)$ steeper. It is instructive to compare the outage probability for the periodic case with the case of higher order end-point compensation described by Eq. (5). One finds that for higher order compensation, i.e., when $N \sim \mu_2' \ln(B_*/B_0) I_0/D_\xi Z$, the (quasi-)periodic scheme becomes more efficient compared to the straight Nth order compensation scheme. If the output signal is not chirped, $\mu_2' = 0$ and the leading term in the expansion of Γ in $\mathbf{h}/b$ is of the third order. Additional filtering efforts can be made to enforce the output pulse symmetry under the $t \longrightarrow -t$ transformation, thus removing the third-order term. Then the leading term in Γ will be of the fourth order in $\mathbf{h}/b$. Finally, even better compensation can be achieved if the standard high-order compensation approach and the (quasi-)periodic ones are combined, i.e., if in the (quasi-)periodic setting considered above, one uses higher order compensation instead of the first order one. Formally, this hybrid case means that the first-order compensation operators $\mathcal{H}_{1n}$ in Eqs. (6), (9) should be substituted by higher order compensation operators $\mathcal{H}_{cn}$. In the hybrid periodic case, Γ can be written as the sum of $\Gamma_n \sim (\int_{a_n}^{ln} dz\mathbf{h}/b)^p$, and, since $\mathbf{h}$ is short-correlated, Γ_n related to different segments are statistically independent. This leads to the following expression for the PDF tail:

$$\ln\mathcal{S}(B) \sim -\mu_p N^{2(p-1)/p}\frac{[D_\xi Z\ln(B/B_0)b^p/I_0]^{2/p}}{D_m Z}, \quad (12)$$

which is valid at $D_\xi Z/I_0\ln(B/B_0) \gg N^{1-p/2}(D_m Z/[\mu_p b^2])^{p/2}$.

Note that an important computational step, leading to our major results in Eqs. (11), (12), was the evaluation of Γ perturbatively in $\mathbf{h}$. In addition, in the periodic case, Γ is a direct sum of each segment contribution Γ_n, and the perturbative treatment applies separately to each Γ_n, requiring the weakness of the PMD effect at each segment only, i.e., $D_m Z/N \ll b^2$. Therefore, one concludes that even an optical line with practically inoperable (without compensation) characteristics ($D_m Z$, which is on the order of or larger than b^2) can still be used for transmission if N is sufficiently large. Moreover, this observation on the applicability of Eqs. (11), (12) also extends to the quasi-periodic case, in the sense that Eqs. (11), (12) provide an upper bound for the PDF of BER. This is due to an additional, oscillatory with $\mathbf{h}$, suppression of Γ_n in the quasi-periodic vs. periodic case. This suppression is especially important for segments strongly separated from their compensating segments.

For illustration purposes, let us briefly discuss the example of a fiber line with typical bit error probability, $B_0 = 10^{-12}$, and $\mu_2' = 0.14$. Assume also that the PMD coefficient, $k = \sqrt{12 D_m}$, is 1.5 ps/$\sqrt{\text{km}}$; the pulse

width, $b = 25$ ps; and fiber length, $Z = 2.500$ km. Then the dimensionless parameter $D_m Z/b^2$ measuring the relevant strength of the PMD effect is $O(1)$; i.e., without compensation, the PMD effect is large, pulses are destroyed, and no successful transmission is possible: $S(B) = 0(1)$ for any $B > B_0$. If, however, the (quasi-)periodic compensation with $N = 10$ compensation units is utilized, the relevant strength of the PMD effect is substantially reduced, so that $D_m Z/[b^2 N] \approx 0.1$, and $S(B)$ begins to decay with B at $B > B_0$. The system performance can be evaluated in terms of the outage probability $\mathbb{O}$, the probability of B being larger than B_*: $\mathbb{O} = \int_{B_*}^{1} dB\mathcal{P}(B)$. One derives from Eq. (11) that in the (quasi-)periodic case, $\mathbb{O} \approx 0.06$ for $B_* = 10^{-8}$ and $\mathbb{O} \approx 0.02$ for $B_* = 10^{-6}$; i.e., the system performance is substantially improved (to become not yet perfect, but already satisfactory).

To conclude, in this letter we have proposed a (quasi-)periodic compensation scheme which appears to be a strong alternative to the standard higher order compensations. The efficiency of the scheme has been demonstrated. Even though technical implementation of this procedure needs expensive equipment, we anticipate that if this compensation technique is implemented, the reduction in the probability of extreme outages will guarantee an essential overall benefit.

The authors thank I. Gabitov for numerous valuable discussions.

REFERENCES

1. C. D. Poole and J. A. Nagel, in *Optical Fiber Telecommunications*, Ed. by I. P. Kaminow and T. L. Koch (Academic, San Diego, 1997), Vol. 3A, p. 114.
2. F. Heismann, in *ECOC'98 Digest* (1998), Vol. 2, p. 51.
3. J. P. Gordon and H. Kogelnik, Proc. Natl. Acad. Sci. USA **97**, 4541 (2000).
4. T. Ono, S. Yamazaki, H. Shimizu, and H. Emura, J. Lightwave Technol. **12**, 891 (1994).
5. F. Heismann, D. Fishman, and D. Wilson, in *Proceedings of ECOC98, Madrid, Spain* (1998), p. 529.
6. L. Moller and H. Kogelnik, in *Proceedings of ECOC'99* (1999), p. 64.
7. H. Bülow, F. Buchali, W. Baumert, *et al.*, Electron. Lett. **36**, 163 (2000).
8. E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications* (Wiley, New York, 1994).
9. V. Chernyak, M. Chertkov, I. Kolokolov, and V. Lebedev, Opt. Lett. (in press); http://arXiv.org/abs/physics/0303013.
10. V. Chernyak, M. Chertkov, I. Kolokolov, and V. Lebedev, Opt. Express **11**, 1607 (2003).
11. B. Huttner, B. Gisin, and N. Gisin, J. Lightwave Technol. **17**, 1843 (1999).

JETP Letters, Vol. 78, No. 4, 2003, pp. 202–206. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 238–242.
Original English Text Copyright © 2003 by Dodin, Fisch, Fraiman.

Drift Lagrangian for a Relativistic Particle in an Intense Laser Field¶

I. Y. Dodin*, N. J. Fisch*, and G. M. Fraiman**

** Princeton Plasma Physics Laboratory, Princeton, NJ 08543, USA*

e-mail: idodin@pppl.gov

*** Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia*

Received February 3, 2003; in final form, June 30, 2003

The Lagrangian and Hamiltonian functions describing the average motion of a relativistic particle under the action of a slightly inhomogeneous intense laser field are obtained. In weak low-frequency background fields, such a particle on average drifts with an effective relativistically invariant mass, which depends on the laser intensity. The essence of the proposed ponderomotive formulation is presented in a physically intuitive and mathematically simple form yet represents a powerful tool for studying various nonlinear phenomena caused by the interaction of currently available smooth ultraintense laser pulses with plasmas. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 52.38.Kd; 52.20.Dq; 52.27.Ny; 52.35.Mw

The latest advances in high-power laser technology have resulted in the development of laser systems capable of delivering superstrong electromagnetic pulses, which can be focused to intensities as high as 10^{21} W/cm^2 [1], with even more powerful systems coming up in the near future [2]. The currently obtainable laser fields can accelerate electrons up to ultrarelativistic oscillatory velocities previously unachievable in experiments on laser-matter interaction. This revolutionary progress is now giving new life to theoretical studies on particle behavior under the action of intense electromagnetic radiation. The conventional models describing various nonlinear phenomena in plasmas illuminated by high-frequency radiation nowadays need to be revised, as new ultra-powerful laser systems are becoming available for laboratory experiments. To explain the already observed phenomena and predict the new effects that have taken place under the action of intense laser drive, an adequate description of single-particle motion under relativistically strong radiation must be developed first.

Currently, particle motion is well-understood when the only forces present are those from a wave of uniform intensity [3]. However, to study the guiding center dynamics in inhomogeneous laser radiation or drifts determined by the presence of low-frequency background fields, additional analysis is needed. Expanding the conventional understanding to this area would result in substantial progress in studying a number of plasma physics problems, such as, e.g., Coulomb collisions and energetic particle production in strong laser fields. Moreover, the hydrodynamics and the electrodynamics of laser-illuminated plasmas would be readily available for general revision.

Often, the dynamics of a particle moving in a high-frequency field is described in terms of the ponderomotive approach. In a nonrelativistic ponderomotive description, the effect of high-frequency electromagnetic forces on a particle is replaced by particle interaction with an average potential, linear on the intensity of laser radiation [4]. When ultraintense lasers are employed, this conventional description needs to be generalized to relativistic motion. Contrary to the degenerate case of a circularly polarized field, in which high-frequency variations of the relativistic mass can be neglected [5], the problem of particle motion in the case of a linear or elliptic polarization represents a certain challenge but can still be studied analytically.

To describe the drift particle dynamics in such fields, recently, multiple studies were performed. Under various approximations, it was shown that the oscillating particle guiding center drifts in a smooth laser field with an effective mass, which depends on the electromagnetic field intensity [6–9]. (In strongly nonuniform laser fields though, the particle dynamics is more complicated [6, 10].) The Hamiltonian treatment of the relativistic drift under intense laser drive has been proposed in [11], though the problem of interaction with low-frequency background fields has not been studied. The first steps towards developing the general formalism with such interaction were made in [8]. However, only smooth (compared to the amplitude of oscillations) low-frequency background fields were taken into consideration, and the relativistic drift motion equations were induced without proper justification.

¶ This article was submitted by the authors in English.

These shortcomings are overcome in our paper, the major emphasis of which is twofold. First, we propose a general fully relativistic Lagrangian formulation of ponderomotive description of particle motion under the action of a quasi-monochromatic slightly inhomogeneous laser field. The proposed approach is physically intuitive yet more systematic and simple in comparison with those discussed previously. After natural generalization, it allows including particle interaction with weak background forces, additional to those from the laser field. The discussion on that aspect of the guiding center motion constitutes the second emphasis of our work. We show the effective mass concept to be applicable to a ponderomotive description of relativistic particle motion in low-frequency background fields, including even ones of a small spatial scale compared to the amplitude of oscillations. In the end, we discuss the most promising applications of the proposed formulation and summarize our main ideas. To start, consider particle motion under the action of a plane laser wave propagating in vacuum, with the vector potential given by

$$\mathbf{A}(\mathbf{r}, t) = (mc^2/e)\mathbf{a}(\eta), \tag{1}$$

where $\eta = \omega t - \mathbf{k}\cdot\mathbf{r}$ stands for the phase of the wave, ω is the wave frequency, and $\mathbf{k} = \mathbf{z}^0\omega/c$ represents the wave vector. The polarization of the wave will be assumed fixed though arbitrary. The magnitude of $\mathbf{a}$, $a = eE/mc\omega$ (where E is the laser electric field), can be understood as the ratio of the momentum imparted by the wave field in a single oscillation to mc, meaning that relativistic effects become important at $a \gtrsim 1$. (For the wavelength $\lambda = 2\pi c/\omega = 1$ μm, the intensity corresponding to $a \sim 1$ for electrons is about 10^{18} W/cm^2.)

In a certain, unique, frame of reference, in such a field the particle undergoes stationary oscillatory "figure-eight" motion in a linearly polarized wave or circular motion in a wave with circular polarization [3]. Averaging over the oscillations, one comes to the concept of the guiding center motion, which we study below. First, let us consider the variational principle that states the minimum value of the action

$$S = \int_{t_1}^{t_2} L\,dt, \tag{2}$$

where L is the Lagrangian function of the particle motion to be realized on the true trajectory. On time scales $t_2 - t_1$, large compared to the oscillation period, the major contribution to the action S (linear on $t_2 - t_1$) is provided by the time-averaged part of the Lagrangian, $\langle L\rangle$ while the contribution of the oscillatory Lagrangian into the integral (2) remains small (for precise analysis, see [9]). Thus, the action S is approximately given by $S = \int_{t_1}^{t_2}\langle L\rangle dt$, from where it follows that $\langle L\rangle$ can be treated as the Lagrangian of the average, guiding center motion.

To obtain the form of the drift Lagrangian $L_0 \equiv \langle L\rangle$, let us consider the latter in the frame of reference where the guiding center rests. In the new frame, the guiding center Lagrangian $\mathcal{L}_0$ can be nothing but a constant, which we put in the form

$$\mathcal{L}_0 = -m_{\text{eff}}c^2, \tag{3}$$

by analogy with the Lagrangian of a true particle with zero velocity. The formally introduced quantity m_{eff}, playing the role of a new effective mass, is yet to be defined. The action (2) is relativistically invariant and can be written as $S = \int_{\tau_1}^{\tau_2}\mathcal{L}_0 d\tau$, where the time τ represents the proper time of the guiding center. Since $d\tau$ is invariant by definition (and thus, so is the Lagrangian $\mathcal{L}_0$), the quantity m_{eff} must also be relativistically invariant. Using

$$d\tau = dt\sqrt{1 - v_0^2/c^2}, \quad \mathcal{L}_0 d\tau = L_0 dt, \tag{4}$$

where $\mathbf{v}_0$ is the velocity of the guiding center in the original frame of reference, one gets the Lagrangian of the guiding center motion

$$L_0 = -m_{\text{eff}}c^2\sqrt{1 - v_0^2/c^2}, \tag{5}$$

which formally coincides with the Lagrangian of a relativistic particle with mass m_{eff} moving with velocity $\mathbf{v}_0$. Since the original frame was chosen arbitrarily, the above expression represents the general form of L_0, where m_{eff} is left to be expressed in terms of the parameters of the laser field.

Let us calculate L_0 in a laboratory frame of reference where the particle has a nonzero average velocity $\mathbf{v}_0$. Instructive in itself, the derivation to follow will also provide us with a number of useful relations connecting the parameters of the particle drift and those related to the actual motion. To proceed, consider the Lagrangian of the true particle motion given by

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} + \frac{e}{c}(\mathbf{v}\cdot\mathbf{A}(\eta)), \tag{6}$$

which is a known periodic function of the phase η rather than time t. Thus, in order to average L over time, one needs to derive a relation connecting time-averaged and phase-averaged quantities. For an arbitrary quantity f, its time and phase averaged values given by

$$\langle f\rangle = \frac{1}{\Delta}\int_{t}^{t+\Delta} f\,dt', \quad \bar{f} = \frac{1}{2\pi}\int_{\eta}^{\eta+2\pi} f\,d\eta', \tag{7}$$

where the limits of integration over the phase correspond to the limits of integration over the time (i.e., $\eta =$

$\eta(t)$), and the time interval Δ is defined as one on which the total phase change equals 2π:

$$\Delta = \int_{\eta}^{\eta + 2\pi} \frac{dt}{d\eta} d\eta. \quad (8)$$

The time interval Δ coincides with the wave period $2\pi/\omega$ only if particle motion is nonrelativistic. However, generally, the phase time-derivative is given by

$$\frac{d\eta}{dt} = \omega\left(1 - \frac{v_z}{c}\right) = \omega\left(\frac{\gamma - p_z/mc}{\gamma}\right), \quad (9)$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the normalized relativistic energy ($\mathscr{E} = m\gamma c^2$) and $\mathbf{p} = m\gamma\mathbf{v}$ is the particle kinetic momentum.

Since the original Lagrangian depends on η (that is, on the combination $z - ct$, rather than z and t separately), there exists an invariant of motion given by

$$u \equiv \gamma - p_z/mc = \text{const}. \quad (10)$$

Substituting the above expressions into Eq. (7), one gets

$$\langle f \rangle = \overline{\gamma f}/\bar{\gamma}. \quad (11)$$

Note that the formula obtained is valid only in the case when the electromagnetic wave (1) is propagating in vacuum. If the refraction index of the medium differs from unit, Eqs. (9) and (10) need to be modified, and the relation between the time- and phase-averaged quantities becomes more complicated [11].

From Eqs. (10) and (11), it follows that

$$\bar{\gamma} = \sqrt{1 + (\bar{p}/mc)^2 + \overline{a^2}}, \quad \bar{\gamma} = \gamma_0\sqrt{1 + \overline{a^2}}, \quad (12)$$

where $\gamma_0 = (1 - v_0^2/c^2)^{-1/2}$, and

$$\mathbf{v}_0 \equiv \langle \mathbf{v} \rangle = \bar{\mathbf{p}}/m\bar{\gamma} \quad (13)$$

is the drift velocity of the particle (compared with the inexact expression given in [6]). Thus, L_0 can be put in the form (5) with m_{eff} given by

$$m_{\text{eff}} = m\sqrt{1 + e^2\overline{A^2}/m^2c^4}. \quad (14)$$

The guiding-center Lagrangian (5) with the expression (14) for the effective mass was also obtained in [9] by a somewhat similar, yet complicated and not straightforward, procedure. In the cited work, Eq. (14) was supposed valid only in the frame of reference where $\mathbf{v}_0 = 0$. In fact, as shown above, it remains applicable for arbitrary $\mathbf{v}_0$, and, furthermore, the actual value of m_{eff} must be relativistically invariant. To express the effective mass in the invariant form, let us notice that, in the laboratory frame where we chose the electric potential $\phi = 0$ (see Eq. (6)), $\sqrt{A^2}$ coincides with the norm of the 4-vector potential $\sqrt{A_\alpha A^\alpha}$, $A^\alpha = (\phi, \mathbf{A})$. The latter is Lorentz-invariant [3], and remains such after being averaged over relativistically invariant phase η. Thus, the expression for m_{eff}, invariant to relativistic transformations, can be put in the following form:

$$m_{\text{eff}} = m\sqrt{1 + \frac{e^2}{m^2c^4}(\overline{A_\alpha A^\alpha})}. \quad (15)$$

Equation (15) was also given in [8], where the average particle motion was studied in a different way.

Reverting to the formula for the drift Lagrangian (5) with the effective mass given by (14), the canonical momentum of the guiding center motion $\mathbf{P}_0$ equals the phase-averaged kinetic momentum $\bar{\mathbf{p}} = m_{\text{eff}}\gamma_0\mathbf{v}_0$, and thus the Hamiltonian function of the guiding center motion can be put in the form

$$H_0 = \sqrt{m_{\text{eff}}^2c^4 + P_0^2c^2}. \quad (16)$$

Here m_{eff} may smoothly depend on the guiding center location $\mathbf{R}_0$ and time t if the wave envelope is slightly nonuniform or time-dependent. Precisely, that means that the laser intensity "seen" by the particle changes insignificantly on one period of particle oscillations, so that the averaging (7) still makes sense, i.e.,

$$l \gg r_\sim, \quad T \gg \Delta, \quad l/v_0 \gg \Delta, \quad (17)$$

where l and T are the spatial and the temporal scales of the wave envelope (for detailed analysis, see [6, 10, 12]).

An alternative derivation of Eq. (16) can be found in [11], where a sequence of canonical transformations of the original motion equations was shown to lead to a similar result. In the present paper, we showed this tedious procedure to be unnecessary for obtaining the expression for the drift Hamiltonian. Compared to the cited work, the distinguishing advantage of the formulation proposed in the present paper is that, because of its apparent mathematical simplicity, this formulation allows easy generalization of the drift Lagrangian and Hamiltonian formalism in the case when the oscillating particles undergo weak acceleration by large-scale low-frequency forces satisfying (17). Interaction with these forces enters the expression for L_0 *additively* and, what is most important, can still be considered within the framework of the effective mass concept.

To show this, consider an oscillating relativistic particle interacting with a field governed by the 4-vector potential $A_{\text{bg}}^\alpha = (\phi_{\text{bg}}, \mathbf{A}_{\text{bg}})$, where the subindex "bg" stands for a background field additional to that of the laser wave. Assume that the field is weak:

$$eE_{\text{bg}}/\gamma_0 m_{\text{eff}}c \ll \omega, \quad eB_{\text{bg}}/\gamma_0 m_{\text{eff}}c \ll \omega, \quad (18)$$

where $\mathbf{E}_{\text{bg}}$ and $\mathbf{B}_{\text{bg}}$ are the corresponding electric and magnetic fields. In this case, the background fields do not impact the oscillatory motion significantly. Thus, averaging the kinetic term mc^2/γ in the Lagrangian

leads to the same expression as in Eq. (5) with $m_{\rm eff}$ given by Eq. (15). In the zeroth-order approximation with respect to the small parameters (18), the average part of the Lagrangian corresponding to particle interaction with the background field can be expressed in terms of the quantity $A_0^\alpha = (\phi_0, \mathbf{A}_0)$ given by

$$A_0^\alpha = \langle A_{\rm bg}^0(\mathbf{R}_0 + \mathbf{r}_\sim)\rangle. \qquad (19)$$

The time-averaging procedure is invariant with respect to changing the drift frame of reference; i.e., it does not alter the Lorentz transformation properties of the quantity being averaged. Thus, A_0^α represents a true 4-vector and can be considered as a new effective electromagnetic field. In terms of this field's potentials, the drift Lagrangian can be put in the following form:

$$L_0 = -m_{\rm eff}c^2\sqrt{1-\frac{v_0^2}{c^2}} + \frac{e}{c}(\mathbf{v}_0\cdot\mathbf{A}_0) - e\phi_0. \qquad (20)$$

In certain applications, it is of interest to consider particle interaction with background fields having spatial scale $l_{\rm bg} \lesssim r_\sim$. If the drift velocity is small, so that the drift displacement on a single period $v_0\Delta$ is small compared to $l_{\rm bg}$, the ponderomotive description can still be applied. However, in this case the difference between the time-averaged potential A_0^α and the true potential $A_{\rm bg}^\alpha$ taken at the location of the guiding center $\mathbf{R}_0$, is crucial. For example, this situation is realized at Coulomb scattering in intense laser fields when $r_\sim$ exceeds the radius of effective interaction [13]. Note that, as follows from the above analysis, the characteristic amplitude of the effective potential remains unchanged as one generalizes the expression for ϕ_0 to the case of relativistic particle motion. In this case, the only difference in calculating ϕ_0 is provided by the change in the oscillatory trajectory $\mathbf{r}_\sim(t)$ to be averaged over.

In the context of the Coulomb scattering problem, the considered Lagrangian approach represents a unique tool for studying the ponderomotive and even the stochastic behavior of the particles being scattered. This problem deserves detailed consideration and will be discussed in future studies, though, briefly, the extension of the proposed formulation can be explained as follows. The stochastic behavior of a dynamical system with periodic coefficients is often convenient to describe in terms of mapping of the dynamical trajectory onto a subspace of the system phase space (for review, see [14]). For the Hamiltonian mapping $(\mathbf{R}_0, \mathbf{P}_0) \longrightarrow (\bar{\mathbf{R}}_0, \bar{\mathbf{P}}_0)$ connecting the particle locations and momenta before and after the time interval equal to the period of the laser field, the generating function is given by the action (2) with $t_1 = t$ and $t_2 = t + 2\pi/\omega$ [14]. Since the drift Lagrangian obtained is approximately proportional to S, it can readily be used for constructing the actual form of this mapping. As will be shown in our future publications, when studying the statistical properties of particle stochastic dynamics (rather than single particle motion) by means of such a mapping, the conditions (17) can be significantly relaxed, which substantially broadens the applicability of the proposed Lagrangian approach. That also allows significant progress in studying the problem of energetic particle production in strong laser fields [15].

Since, in the case of relativistic drift, $\mathbf{r}_\sim$ depends on $\mathbf{v}_0$, the expression for the canonical momentum $\mathbf{P}_0 = \partial L_0/\partial\mathbf{v}_0$, the drift motion equations become complicated. However, in two special cases of interest, these can be simplified. In a large-scale background field satisfying the conditions (17), locally, $A_{\rm bg}^\alpha$ can be treated as a linear function of $\mathbf{r}$. Therefore, the velocity-dependent part averages out when calculating the potential A_0^α, and one gets $A_0^\alpha \approx A_{\rm bg}^\alpha$. Thus, the drift canonical momentum equals $\mathbf{P}_0 = m_{\rm eff}\gamma_0\mathbf{v}_0 + (e/c)\mathbf{A}_{\rm bg}$, and the Hamiltonian function is given by

$$H_0 = \sqrt{m_{\rm eff}^2c^4 + \left(\mathbf{P}_0 - \frac{e}{c}\mathbf{A}_{\rm bg}\right)^2c^2} + e\phi_{\rm bg}, \qquad (21)$$

where the potentials are assumed to be slow functions of $\mathbf{R}_0$ and t. The guiding center motion equations can be put in the covariant form

$$\frac{dR_0^\alpha}{d\tau} = \frac{p_0^\alpha}{m_{\rm eff}}, \quad \frac{dp_0^\alpha}{d\tau} = \frac{e}{c}F_{\rm bg}^{\alpha\beta}U_\beta - c^2\frac{\partial m_{\rm eff}}{\partial R_0^\alpha}, \qquad (22)$$

where $R_0^\alpha = (ct, \mathbf{R}_0)$ is the 4-coordinate of the guiding center, $p_0^\alpha = (\mathscr{E}_0/c, m_{\rm eff}\gamma_0\mathbf{v}_0)$ is the drift kinetic 4-momentum, $\mathscr{E}_0 = m_{\rm eff}\gamma_0c^2$ is the energy of the guiding center motion, $F_{\rm bg}^{\alpha\beta}$ is the electromagnetic field tensor corresponding to the potential $A_{\rm bg}^\alpha$ [3], and $U^\alpha = \gamma_0(c, \mathbf{v}_0)$ is the guiding center 4-velocity. Covariant Eqs. (22) were also given in [8], though no strict derivation of these was proposed. Another expression for the relativistic ponderomotive force is given in [9].

From Eq. (21), it follows that, in a low-frequency large-scale background field, the guiding center of a relativistic particle moving under the action of intense laser radiation behaves as a particle with the effective mass $m_{\rm eff}$ drifting in the same background field. This conclusion is also supported by direct numerical computations given in [8]. For example, in static magnetic field $\mathbf{B}_{\rm bg}$, the guiding center undergoes Larmor motion with the frequency $\omega_B = eB_{\rm bg}/\gamma_0m_{\rm eff}c$. The conventional expression for the drift velocity in a nonuniform magnetic field [3] also applies to the average motion if the particle mass is replaced with that given by Eq. (15).

In addition to the case of large-scale background fields, the guiding center motion equations can also be put in a simple physically intuitive form in the case of nonrelativistic drift motion. Since the drift velocity enters the expression for A_0^α only through relativistic dependence of $\mathbf{r}_\sim$ on $\mathbf{v}_0/c$, then, in the case $\mathbf{v}_0 \ll c$, $\partial A_0^\alpha/\partial \mathbf{v}_0$ can be neglected. In this case, the drift canonical momentum is given by $\mathbf{P}_0 = m_{\text{eff}}\mathbf{v}_0 + (e/c)\mathbf{A}_0$, and the Hamiltonian can be put in the form

$$H_0 = \frac{1}{2m_{\text{eff}}}\left(\mathbf{P}_0 - \frac{e}{c}\mathbf{A}_0\right)^2 + m_{\text{eff}}c^2 + e\phi_0, \qquad (23)$$

where the effective mass m_{eff} and the potential energy $\psi_{\text{eff}} = m_{\text{eff}}c^2 + e\phi_0$ may slowly depend on the guiding center location $\mathbf{R}_0$ and time t. Note that even in a uniform laser field, $\nabla\psi_{\text{eff}}$ may differ significantly from $e\nabla\phi_{\text{bg}}$ when the amplitude of particle oscillations $\mathbf{r}_\sim$ exceeds the spatial scale of the background field l [13]. The regime of slow drift motion superimposed on relativistic oscillations is the one that is actually realized in many current experiments on intensive laser pulses interaction with rare plasmas. (By rare plasmas we mean those having a refraction index close to unity, as assumed for all the results obtained in the present paper.) This fact makes the above analysis especially useful from the practical point of view, as it represents a simple tool for studying actual experimental data. Finally, the well-known nonrelativistic ponderomotive potential [4] can be readily derived from Eq. (23) by keeping the correction to the effective mass, linear with respect to the wave intensity (see also [9]).

In summary, we showed that, in weak low-frequency background fields, a relativistic particle moving under the action of intense laser radiation drifts like a quasiparticle with an effective mass, which depends on the intensity of the laser field. The intuitive expectation that, by the order of magnitude, the drift motion equations must coincide with those without the laser field if the appropriate relativistic correction of particle mass is introduced, can now be considered proven for various types of background fields. The proposed formulation can be useful for studying numerous phenomena resulting from intense laser-plasma interaction, such as, e.g., the energetic particle production and Coulomb scattering in strong laser fields. Moreover, the mathematical simplicity of the proposed approach allows easy generalization of the rare plasma hydrodynamics and electrodynamics to the case of plasmas illuminated by ultraintense laser radiation. Replacing the electron mass with the effective mass (15), one can readily derive the generalized dispersion relations for various linear waves in plasmas, as well as revise the nonlinear plasma dynamics.

This work was supported by the US DOE, contract DE-AC02-76 CHO3073, and the Russian Foundation for Basic Research, grant nos. 02-02-17277 and 02-02-17275.

REFERENCES

1. M. D. Perry, D. Pennington, B. C. Stuart, *et al.*, Opt. Lett. **24**, 160 (1999); M. H. Key, M. D. Cable, T. E. Cowan, *et al.*, Phys. Plasmas **5**, 1966 (1998).
2. V. M. Malkin, G. Shvets, and N. J. Fisch, Phys. Rev. Lett. **82**, 4448 (1999); V. M. Malkin, G. Shvets, and N. J. Fisch, Phys. Plasmas **7**, 2232 (2000).
3. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 7th ed. (Nauka, Moscow, 1988; Pergamon Press, Oxford, 1971).
4. A. V. Gaponov and M. A. Miller, Sov. Phys. JETP **7**, 168 (1958); H. Motz and C. J. H. Watson, Adv. Electron. **23**, 153 (1967).
5. A. G. Litvak, *Reviews of Plasma Physics* (Énergoatomizdat, Moscow, 1980); V. A. Kozlov, A. G. Litvak, and E. V. Suvorov, Zh. Éksp. Teor. Fiz. **76**, 147 (1979) [Sov. Phys. JETP **49**, 75 (1979)].
6. B. Quesnel and P. Mora, Phys. Rev. E **58**, 3719 (1998); P. Mora and T. Antonsen, Jr., Phys. Plasmas **4**, 217 (1997).
7. T. W. B. Kibble, Phys. Rev. **150**, 1060 (1966); X. Li, Phys. Rev. E **59**, 6048 (1999).
8. J. E. Moore and N. J. Fisch, Phys. Plasmas **1**, 1105 (1994).
9. D. Bauer, P. Mulser, and W. H. Steeb, Phys. Rev. Lett. **75**, 4622 (1995).
10. F. V. Hartemann, S. N. Fochs, G. P. Lesage, *et al.*, Phys. Rev. E **51**, 4833 (1995); D. R. Bituk and M. V. Fedorov, JETP **89**, 640 (1999); V. D. Taranukhin, JETP **90**, 447 (2000); N. B. Narozhny and M. S. Fofanov, JETP **90**, 753 (2000); A. V. Serov, JETP **92**, 20 (2001); J. X. Wang, Y. K. Ho, and W. Scheid, Phys. Lett. A **234**, 415 (1997).
11. M. D. Tokman, Plasma Phys. Rep. **25**, 140 (1999).
12. I. Y. Dodin and N. J. Fisch, Phys. Rev. E (in press).
13. M. H. Mittleman, *Introduction to the Theory of Laser–Atom Interactions* (Plenum, New York, 1993); M. Pont, N. R. Walet, and M. Gavril, Phys. Rev. A **41**, 477 (1990).
14. A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, 2nd ed. (Springer, New York, 1992).
15. A. A. Balakin and G. M. Fraiman, Prikl. Fiz. **1**, 65 (2003).

JETP Letters, Vol. 78, No. 4, 2003, pp. 207–212. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 243–248.
Original Russian Text Copyright © 2003 by Batyev.

On the Superconductivity of a Wigner Liquid

É. G. Batyev

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia
e-mail: batyev@isp.nsc.ru

Received July 1, 2003

The disorder effect on the interaction of quasiparticles between each other is discussed. The occurrence of a soft mode is taken as the basic assumption. The interaction through the soft mode results in attraction between Fermi quasiparticles (this is apart from the repulsion that has remained from the initial Coulomb interaction between particles). This attraction (in the vicinity of the Fermi surface) is strengthened with increasing concentration of the scattering centers. Therefore, even if the pure system exhibits no superconductivity, superconductivity could appear in the impurity system. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 71.27.+a; 74.10.+v

A low-density system of two-dimensional electrons (holes) exhibits a number of unusual properties (see the review [1]). One of the main of these properties is the occurrence of a metal–insulator transition upon decreasing the carrier density below a certain value of n_c. The resistivity in the metallic phase (at higher densities $n > n_c$) decreases with decreasing temperature. However, it has been generally agreed that, in a disordered medium in the two-dimensional case, all particles are localized and the metallic state does not exist, and it has been expected that the resistivity must increase with decreasing temperature, as is the case at $n < n_c$. The transition mentioned above is observed at sufficiently low carrier densities, when the Coulomb interaction of particles becomes large as compared to the kinetic energy. Such a system is strongly correlated with short-range order as in a Wigner crystal (therefore, it is sometimes called Wigner liquid).

Though the problem of the transition will not be discussed here, it is still reasonable to give some consideration to the mechanism of this transition to clarify the formulation of the problem to be considered. It might be expected that the origin is in Coulomb centers, which act as though they pin the liquid simply because the liquid is adjusted to these centers to form short-range order in a certain vicinity due to the Coulomb interaction. In the subsequent discussion, the effect of such centers will not be taken into account; that is, it will be assumed that their density is sufficiently small. In addition, a simplified (idealized) situation will be considered in general, when the scattering centers are assumed to be short-ranged and sufficiently weak so that phenomena like the above-mentioned pinning can be neglected. Then, quasiparticle scattering by a center exists, whereas there is no significant rearrangement of the liquid in the vicinity of such a center, as apparently occurs at a charged center.

In this work, the possibility that the superconducting state forms in a Wigner liquid is discussed. The peculiar feature of a low-density two-dimensional electron system is strong correlations, which may result in the appearance of a so-called soft mode, that is, low-energy Bose excitations at finite momenta (like rotons in superfluid helium). It is suggested that Bose excitations arise apart from the usual Fermi excitations. This suggestion was made in [2] to explain the temperature dependence of the resistivity in the metallic phase. The physical picture here is simple: a soft mode can appear as a precursor of Wigner crystallization, the number of Bose excitations increases, and their contribution to the dissipation of the system momentum and to the resistivity increases as well.

The soft mode results in electron–electron interaction through boson exchange, which, as is known from the theory of superconductivity [3–5], leads to attraction between particles, and, if this attraction is larger than the Coulomb repulsion, Cooper pairs and a superconducting state can form. Attraction at the Fermi surface is most important for the formation of Cooper pairs. In our problem, this is possible only through the exchange of two bosons (because the boson momentum is larger than two Fermi momenta). In the impurity system, the momentum can be passed to an impurity; therefore, the exchange of one boson is possible; this additional interaction channel enhances attraction with increasing disorder. As a result, even though superconductivity is absent in the pure system, it can appear in an impurity system.

Previously, arguments were advanced in favor of the possibility of a superconducting state in the systems under consideration [6]. In this work, a mechanism that is peculiar to a Wigner liquid (and that can apparently lead to superconductivity in a system with disorder) is discussed.

0021-3640/03/7804-0207$24.00 © 2003 MAIK "Nauka/Interperiodica"

Model. The model uses the spectra of fermions and bosons and their interaction with each other and with impurities. The Hamiltonian of the system has the form

$$H = H_a + H_{ab} + H_b,$$

$$H_a = \sum_{\mathbf{p}} \xi_{\mathbf{p}} a_{\mathbf{p}}^+ a_{\mathbf{p}} + \frac{1}{2V} \sum_{\mathbf{p},\mathbf{p}',\mathbf{k}} W_a(\mathbf{k}) a_{\mathbf{p}}^+ a_{\mathbf{p}'}^+ a_{\mathbf{p}'+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}},$$

$$H_{ab} = \frac{\lambda}{V} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} a_{\mathbf{p}}^+ B_{-\mathbf{q}} B_{\mathbf{q}+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}}, \quad (1)$$

$$H_b = \sum_{\mathbf{q}} \Omega_{\mathbf{q}} b_{\mathbf{q}}^+ b_{\mathbf{q}} + \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} U(-\mathbf{q}) B_{\mathbf{q}}.$$

Here, $\Omega_{\mathbf{q}}$ is the boson energy, the operator $B_{\mathbf{q}}$ coincides (within a factor) with the coordinate of the corresponding oscillator and is expressed through the boson creation and annihilation operators

$$\Omega_{\mathbf{q}}^2 = \Omega_0^2 + v_0^2 (q - q_0)^2, \quad B_{\mathbf{q}} = \frac{b_{\mathbf{q}} + b_{-\mathbf{q}}^+}{\sqrt{2\Omega_{\mathbf{q}}}}.$$

Here, it is taken into account that an additional term due to the external field (the field of impurities) exists in the system energy for spinless bosons. This term is linear in the oscillator coordinate. The other designations pertain to fermions (the spin and valley indices are omitted): $\xi_{\mathbf{p}}$ is the fermion energy reckoned from the Fermi surface, the quantity $W_a(\mathbf{k})$ is the Fourier component of the interaction of the fermions with each other (the interaction of fermions with impurities will be unnecessary).

We emphasize that the interaction (attraction) of fermions at the Fermi surface is important for the superconducting state [3–5]. In the model Hamiltonian (1), the part that describes the interaction of fermions with bosons and leads to the attraction of fermions at the Fermi surface is written as the operator H_{ab}. Instead of this operator, it would appear reasonable to use another operator as follows:

$$H'_{ab} = \frac{\lambda'}{\sqrt{V}} \sum_{\mathbf{p},\mathbf{q}} a_{\mathbf{p}}^+ a_{\mathbf{p}-\mathbf{q}} B_{\mathbf{q}}.$$

However, if $q_0 > 2p_F$ (p_F is the Fermi momentum), which is suggested below (the validity of this suggestion is confirmed by the estimates [2]), this operator makes no direct contribution to fermion scattering at the Fermi surface; the contribution can arise in higher orders of perturbation theory. Because perturbation theory is inapplicable in our problem, it is appropriate to write the interaction with bosons directly, as is done in (1), using the phenomenological interaction parameter λ.

Consider first the interaction of bosons with impurities. The operator H_b is diagonalized by the transformation

$$b_{\mathbf{q}} \longrightarrow b_{\mathbf{q}} + C_{\mathbf{q}}, \quad C_{\mathbf{q}} = \frac{-1}{\sqrt{2V}} \frac{U(\mathbf{q})}{\Omega_{\mathbf{q}}^{3/2}}. \quad (2)$$

This transformation corresponds to a displacement of the equilibrium point of the oscillator with the boson spectrum remained unchanged.

Thus, the interaction of fermions with bosons changes in the field of impurities; therefore, instead of the term H_{ab}, one obtains

$$H_{ab} \longrightarrow \frac{\lambda}{V} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} a_{\mathbf{p}}^+ B_{-\mathbf{q}} B_{\mathbf{q}+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}} + \frac{2\lambda}{V} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} a_{\mathbf{p}}^+ \bar{B}_{-\mathbf{q}} B_{\mathbf{q}+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}} + \frac{\lambda}{V} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} a_{\mathbf{p}}^+ \bar{B}_{-\mathbf{q}} \bar{B}_{\mathbf{q}+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}}, \quad (3)$$

where, by definition,

$$\bar{B}_{\mathbf{q}} = \frac{C_{\mathbf{q}} + C_{-\mathbf{q}}^*}{\sqrt{2\Omega_{\mathbf{q}}}} \longrightarrow \frac{2C_{\mathbf{q}}}{\sqrt{2\Omega_{\mathbf{q}}}}. \quad (4)$$

The last term in expression (3) is the additional contribution to the interaction of fermions with impurities. It is instructive to elucidate the character of the interaction with an impurity resulting from this equation. The Fourier components entering into (2) depend on the coordinates of impurities

$$U(\mathbf{q}) = \sum_l u(\mathbf{q}) \exp(-i\mathbf{q}\mathbf{R}_l), \quad (5)$$

where $u(\mathbf{q})$ is the Fourier component of the interaction with a single impurity (for a short-range potential, it is real and constant, which is assumed hereafter), and the summation is performed over the coordinates of impurities. For a single impurity located at point $\mathbf{R}$, one obtains

$$\bar{B}_{\mathbf{q}} \longrightarrow -\frac{u(\mathbf{q})}{\sqrt{V}\Omega_{\mathbf{q}}^2} \exp(-i\mathbf{q}\mathbf{R}),$$

so that the last term in (3) for a single impurity (H_S) takes the form

$$H_S = \frac{1}{V} \sum_{\mathbf{k},\mathbf{p}} W_S(\mathbf{k}) a_{\mathbf{p}}^+ a_{\mathbf{p}-\mathbf{k}},$$

where the Fourier component of the interaction with an

impurity through the soft mode $W_S(\mathbf{k})$ has the form

$$W_S(\mathbf{k}) = \frac{\lambda u^2}{V}\sum_{\mathbf{q}}\frac{\exp(-i\mathbf{kR})}{\Omega_{\mathbf{q}}^2\Omega_{\mathbf{q}+\mathbf{k}}^2}. \quad (6)$$

Calculating the sum in (6) gives

$$\frac{1}{V}\sum_{\mathbf{q}}\frac{1}{\Omega_{\mathbf{q}+\mathbf{k}}^2\Omega_{\mathbf{q}}^2} \longrightarrow \frac{q_0}{2\Omega_0^2 v_0^2}J(k), \quad (7)$$

$$J(k) = \frac{1}{k\sqrt{1-k^2/(2q_0)^2}}. \quad (8)$$

Here, the special designation $J(k)$ has been introduced for the function that will also be used later (the value of $J(k)$ at $k \longrightarrow 0$ will be considered below). The calculation was carried out by passing from integration over the angle to integration over the quantity $q' = |\mathbf{q} + \mathbf{k}|$, namely,

$$d\varphi \longrightarrow 2J(k)dq'$$

(two regions, φ is the angle between vectors $\mathbf{q}$ and $\mathbf{k}$), and then passing to integration over the energy

$$dq \longrightarrow \frac{2\Omega d\Omega}{v_0\sqrt{\Omega^2-\Omega_0^2}}$$

(two regions as well). An analogous transformation was also made for q'. Because of a singularity in (8) at small values of k, the domain of applicability should be determined. For this purpose, it is necessary to calculate the sum in (7) at $k = 0$, which gives

$$\frac{1}{V}\sum_{\mathbf{q}}\frac{1}{\Omega_{\mathbf{q}}^4} = \frac{q_0}{4v_0\Omega_0^3}. \quad (9)$$

From here it is evident that expression (8) is true up to $k \sim \Omega_0/v_0$; that is, $J(0) \sim v_0/\Omega_0$.

Note that, because of the presence of a soft mode, the interaction of Fermi quasiparticles with a short-range impurity appears at long distances (at small transferred momenta) as an interaction with a repulsive charged center (with the screening radius of the order of v_0/Ω_0), that is, becomes long-ranged.

Further, the zero-temperature diagram technique will be used (see, for example, [4]). The boson propagator will be defined as follows:

$$D(t-t', \mathbf{q}) = -i\langle TB_{-\mathbf{q}}(t)B_{\mathbf{q}}(t')\rangle,$$

where the symbol T means time ordering. The Fourier component has the form

$$D(\omega, \mathbf{q}) = \frac{1}{\omega^2-(\Omega_{\mathbf{q}}-i\delta)^2}, \quad (10)$$

where, as usual, going around the poles for the causal Green's function is indicated.

Now, the interaction of fermions through bosons can be obtained from (3). In the second order of perturbation theory, the first term gives

$$W' \longrightarrow 2i\frac{\lambda^2}{V}\sum_{\mathbf{q}+\mathbf{q}'=\mathbf{k}}\int\frac{d\omega}{2\pi}D(\omega, \mathbf{q})D(\epsilon-\omega, \mathbf{q}')$$

(this interaction is additional to the direct interaction $W_a(\mathbf{k})$, see (1)). After calculating the integral over the frequency,

$$W' \longrightarrow \frac{\lambda^2}{\Lambda}\sum_{\mathbf{q}+\mathbf{q}'=\mathbf{k}}\frac{\Omega+\Omega'}{\Omega\Omega'}\frac{1}{\epsilon^2-(\Omega+\Omega')^2}. \quad (11)$$

Here, ϵ and $\mathbf{k}$ are the transferred frequency and momentum, respectively (going around the poles as in (10)), the quantities Ω and Ω' correspond to momenta $\mathbf{q}$ and $\mathbf{q}'$.

For the scattering of quasiparticles at the Fermi surface, one may assume that $\epsilon = 0$ in (11), so that the resulting fermion interaction appears as

$$W_a(\mathbf{k}) \longrightarrow \tilde{W}(\mathbf{k}) = W_a(\mathbf{k}) + W_1(\mathbf{k}),$$
$$W_1(\mathbf{k}) = \frac{-\lambda^2}{V}\sum_{\mathbf{q}+\mathbf{q}'=\mathbf{k}}\frac{1}{\Omega\Omega'(\Omega+\Omega')}. \quad (12)$$

After calculating the sum in (12), one obtains

$$W_1(\mathbf{k}) = \frac{-\lambda^2 q_0}{2v_0^2\Omega_0}J(k). \quad (13)$$

Estimates. The constants that determine the interaction can be estimated with the use of the following considerations. The initial (Coulomb) interaction gave rise to short-range order in the arrangement of electrons in the system (which, by hypothesis, is revealed in the existence of a soft mode). It appears reasonable that the system, as a Fermi liquid, contains Fermi quasiparticles (though, sometimes, doubt is cast on this suggestion). The fermion–fermion interaction has nothing to do with the initial Coulomb interaction, because the energy scales are different: for example, the Fermi energy $\epsilon_F \sim p_F^2/m^*$ is considerably smaller than the initial Coulomb energy; however, this interaction is not weak on these new scales. Therefore, it might be thought that the interaction of fermions has a characteristic value ϵ_F and a characteristic range equal to the interparticle distance; that is, the Fourier component of this interaction W_a (see (1)) can be estimated as follows:

$$W_a \sim 1/m^*.$$

This is allegedly the direct fermion interaction (repulsion). If a soft mode exists with the corresponding bosons, the interaction of fermions through the bosonic field appears additionally. This interaction corresponds to attraction at the Fermi surface and is represented by the contribution W_1 in expression (13). This part of

interaction can be described in much the same way; that is, at the characteristic momentum of order p_F, it should be expected that $|W_1| \sim W_a$ (these are different parts of the fermion–fermion interaction, and it is beyond reason to consider that these parts are strongly different). This leads to the following estimate for λ:

$$\lambda \sim \sqrt{\epsilon_F \Omega_0}/m^*. \tag{14}$$

As a result (on the average, that is, for the zero harmonic), this apparently gives repulsion.

Additional interaction. The second term in (3) is of interest. This term leads to an additional interaction of fermions at the Fermi surface due to the exchange of one boson accompanied by the transfer of part of the momentum to an impurity (this contribution is absent in the pure system).

Consider the process in which two fermions with momenta **p** and **p'** change their momenta in the following way:

$$\mathbf{p} \longrightarrow \mathbf{p} - \mathbf{k}, \quad \mathbf{p}' \longrightarrow \mathbf{p}' + \mathbf{k}'.$$

For this process, the contribution of the second term in (3) takes the form

$$W'' \longrightarrow \frac{(2\lambda)^2}{V}\sum_{\mathbf{q}} \bar{B}_{-\mathbf{q}-\mathbf{k}}\bar{B}_{\mathbf{q}+\mathbf{k}'}D(\epsilon, \mathbf{q}). \tag{15}$$

The momentum $\mathbf{k} - \mathbf{k}'$ is passed to the impurity.

To compare (11) and (15), the latter expression should be averaged over the arrangement of impurities, which makes sense for a sufficiently high concentration of the impurity centers. This can be made in the regular way [4]. After averaging the product $U(\mathbf{q})U(-\mathbf{q}')$ (see definition (5)), one obtains

$$\langle U(\mathbf{q})U(-\mathbf{q}')\rangle = N_i u^2 \delta_{\mathbf{q},\mathbf{q}'}, \tag{16}$$

where N_i is the number of impurities. For (15), one obtains

$$\langle W''\rangle \longrightarrow \frac{(2\lambda^2)}{V}\sum_{\mathbf{q}} \frac{u^2 n_i}{\Omega^4_{\mathbf{q}+\mathbf{k}}} D(\epsilon, \mathbf{q})\delta_{\mathbf{k},\mathbf{k}'}, \tag{17}$$

where $n_i = N_i/V$. Here, the symbol $\langle\ldots\rangle$ means averaging over the arrangement of impurities. As previously for quantity $W_1(\mathbf{k})$ (see (12) and (13)), expression (17) at $\epsilon \longrightarrow 0$ will be of interest. The corresponding additional term in the interaction $W_2(\mathbf{k})$ appears as follows:

$$W_2(\mathbf{k}) = -\frac{(2\lambda)^2}{V}\sum_{\mathbf{q}} \frac{u^2 n_i}{\Omega^4_{\mathbf{q}+\mathbf{k}}\Omega^2_{\mathbf{q}}}. \tag{18}$$

After calculating the sum, the following equation is obtained instead of (18):

$$W_2(\mathbf{k}) = -\frac{(\lambda u)^2 n_i q_0}{\Omega_0^4 v_0^2} J(k). \tag{19}$$

Now, different contributions can be compared:

$$W_2/W_1 \sim u^2 n_i/\Omega_0^3. \tag{20}$$

Little can be said about the quantity u, if only because this value depends on the type of impurities. However, it turns out that $u^2 n_i$ can be related to some observable quantities (see below). Let there be short-ranged neutral centers. A fermion feels such centers at distances of order v_0/Ω_0 because of the interaction through the soft mode (this was shown for a single impurity; see (7)–(9)). At sufficiently high concentrations of such centers, that is, at

$$n_i > \Omega_0^2/v_0^2,$$

the action regions of different centers overlap, and the averaging procedure seems natural. In this limit, the quantity $u^2 n_i$ can be related to the transport time τ_{tr}, which determines the conductivity σ by the conventional equation

$$\sigma = ne^2\tau_{\text{tr}}/m$$

(m is the band mass, as distinct from the effective mass m^* renormalized because of interaction). If it is considered that relaxation is determined by the interaction with impurities according to the last term in (3), two contributions to $1/\tau_{\text{tr}}$ should be taken into account: linear and quadratic ones in n_i. The second one dominates in the limit indicated above, so that the following estimate is obtained:

$$\frac{1}{\tau_{\text{tr}}} \sim (u^2 n_i)^2 \frac{\epsilon_F}{m v_0^2 \Omega_0^5}$$

(λ is estimated according to (14)). This gives just the sought relation. Apparently, $v_0 \sim v_F$; then, instead of (20), one obtains

$$\frac{W_2}{W_1} \sim \frac{\sqrt{m/m^*}}{\sqrt{\Omega_0\tau_{\text{tr}}}}.$$

This result gives an estimate of the wanted ratio.

Superconducting state. In the discussion of this subject, a number of simplifications are made. The main one is the consideration of the averaged (over the arrangement of impurities) interaction between fermions, which is valid only at a sufficiently high concentration of scattering centers (see the comment after (20)).

From the theory of superconductivity, it is known that the scattering from impurities substantially affects the Cooper pairs with nonzero angular momenta, which eventually leads to the breakdown of superconductivity. This does not take place for pairing with zero angular momentum. In our problem, it is apparently impossible to do without scattering from impurities, because it can be of decisive importance for the appearance of resulting attraction between fermions; hence, it is pairing with zero angular momentum that should be analyzed.

Therefore, the interaction of particles with different spins or different valleys, for which this kind of pairing is possible, will be considered. Denote the corresponding annihilation (creation) operators by symbols a, A and (a^+, A^+). The interaction between these particles H_{aA} can be written as (21)

$$H_{aA} = \frac{1}{V}\sum W(\mathbf{k})a^+_{\mathbf{p}-\mathbf{k}}A^+_{\mathbf{p}'+\mathbf{k}}A_{\mathbf{p}'}a_{\mathbf{p}}, \qquad W(\mathbf{k}) = W_a(\mathbf{k}) + W_1(\mathbf{k}) + W_2(\mathbf{k}). \tag{21}$$

Here, W_a is the direct interaction of particles (the value that remained from the initial Coulomb interaction, that is, repulsion, see (1)), and $W_1 + W_2$ is attraction due to the exchange of bosons (see (13) and (19)). This is a simplified equation, because bosons do not appear explicitly and their role is manifested in the attraction $W_1 + W_2$ induced by bosons. By analogy with the phonon model, the attraction of fermions here should be restricted to a certain vicinity $\pm\tilde{\Omega}$ at the Fermi surface, where $\tilde{\Omega} \sim \Omega_0$. As for W_a, according to the aforesaid, it is repulsion and it is substantial in the vicinity $\pm\tilde{\epsilon}$, where $\tilde{\epsilon} \sim \epsilon_F$. Presumably, $\tilde{\Omega} \ll \tilde{\epsilon}$.

Further, as usual, the problem can be simplified even more by retaining only opposite momenta:

$$H_{aA} \longrightarrow \frac{1}{V}\sum_{\mathbf{p},\,\mathbf{p}'} W(\xi, \xi')a^+_{\mathbf{p}}A^+_{-\mathbf{p}}A_{-\mathbf{p}'}a_{\mathbf{p}'}, \tag{22}$$

where $W(\xi, \xi') = W_a(\xi, \xi') + W_b(\xi, \xi')$. According to the aforesaid, it is implied that function $W_a(\xi, \xi')$ is a positive constant in a wide range of variables $|\xi|, |\xi'| < \tilde{\epsilon}$, whereas $W_b(\xi, \xi') \longleftrightarrow W_1 + W_2$ is a negative constant in the range $|\xi|, |\xi'| < \tilde{\Omega}$ (zero harmonics, that is, harmonics without angular dependence are implied for pairing with zero angular momentum). Such a model for taking into account Coulomb repulsion (apart from attraction through phonons) in the theory of superconductivity was considered previously (see the book [5], section 6.3). Therefore, the solution of the problem will be considered only briefly.

In the self-consistent field approximation, which works well for the model with interaction (22), one obtains

$$H_{aA} \longrightarrow \sum_{\mathbf{p}} \Delta(\xi)\{A_{-\mathbf{p}}a_{\mathbf{p}} + a^+_{\mathbf{p}}A^+_{-\mathbf{p}}\}$$

(an insignificant constant is omitted), where, by definition,

$$\Delta(\xi) = \frac{1}{V}\sum_{\mathbf{p}'} W(\xi, \xi')\langle\langle A_{-\mathbf{p}'}a_{\mathbf{p}'}\rangle\rangle \tag{23}$$

(the symbol $\langle\langle\ldots\rangle\rangle$ means averaging over the ground state at zero temperature). In this case, the order parameter Δ can be considered real. After adding the first term from (1) to H_{aA}, consider the part $h_{\mathbf{p}}$ containing momenta $\pm\mathbf{p}$

$$h_{\mathbf{p}} = \xi(a^+a + A^+A) + \Delta(Aa + a^+A^+) \tag{24}$$

(indices are omitted here). This operator is diagonalized by the Bogolyubov (u, v) transformation

$$a_{\mathbf{p}} = u_{\mathbf{p}}\alpha_{\mathbf{p}} + v_{\mathbf{p}}\beta^+_{-\mathbf{p}}, \quad A_{-\mathbf{p}} = u_{\mathbf{p}}\beta_{-\mathbf{p}} - v_{\mathbf{p}}\alpha^+_{\mathbf{p}} \tag{25}$$

with the coefficients

$$(u^2, v^2) = \frac{1}{2}\left(1 \pm \frac{\xi}{\sqrt{\xi^2 + \Delta^2}}\right), \qquad uv = \frac{-\Delta}{2\sqrt{\xi^2 + \Delta^2}}. \tag{26}$$

After that, instead of (24), one obtains

$$h_{\mathbf{p}} \longrightarrow \epsilon_{\mathbf{p}}(\alpha^+_{\mathbf{p}}\alpha_{\mathbf{p}} + \beta^+_{-\mathbf{p}}\beta_{-\mathbf{p}}), \quad \epsilon_{\mathbf{p}} = \sqrt{\xi^2_{\mathbf{p}} + \Delta^2} \tag{27}$$

(the constant part is omitted). Here, $\epsilon_{\mathbf{p}}$ is the quasiparticle energy in the superconducting phase. The equation for the order parameter (23) can be written now as

$$\Delta(\xi) = \frac{1}{V}\sum_{\xi'} W(\xi, \xi')\frac{-\Delta(\xi')}{2\sqrt{\xi'^2 + \Delta^2(\xi')}}. \tag{28}$$

With the properties of function $W(\xi, \xi')$ in mind, one can express the solution of Eq. (28) through two constants

$$\Delta(\xi) = \Delta_1 \quad (0 < |\xi| < \tilde{\Omega}),$$

$$\Delta(\xi) = \Delta_2, \quad (\tilde{\Omega} < |\xi| < \tilde{\epsilon}),$$

and, for the gap in the spectrum $\Delta = |\Delta_1|$, the following expression is obtained in the limit $\Delta \ll \tilde{\Omega}$:

$$\Delta = 2\tilde{\Omega}\exp(-1/g), \quad g = g_b - g_a\left(g_a\ln\frac{\tilde{\epsilon}}{\tilde{\Omega}} + 1\right)^{-1}.$$

Here, the following designations are introduced:

$$g_b = \gamma|W_b|, \quad g_a = \gamma W_a$$

(γ is the density of states at the Fermi surface, and $2\pi\gamma = m^*$). The ordered phase corresponds to $g > 0$. The role of repulsion is decreased because of the contribution with the logarithm [5].

According to my estimates, both interaction constants $g_{a,b}$ in the pure system are of order unity. With increasing disorder, the resistivity increases but the attraction of quasiparticles through the soft mode (that is the value of g_b) also increases; hence, there comes a point where the resistivity abruptly vanishes because of a transition to the superconducting state.

Thus, within the assumed soft-mode model, it is shown that a Wigner liquid can pass to a superconduct-

ing state at a sufficient amount of not very strongly scattering centers.

The author is grateful to M.V. Éntin and Z.D. Kvon for discussions.

This work was supported in particular by the Russian Foundation for Basic Research, project no. 02-02-16159 and INTAS, project no. 2212.

REFERENCES

1. E. Abrahams, S. V. Kravchenko, and M. P. Sarachik, Rev. Mod. Phys. **73**, 251 (2001).
2. É. G. Batyev, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 727 (2000) [JETP Lett. **72**, 506 (2000)]; Pis'ma Zh. Éksp. Teor. Fiz. **73**, 435 (2001) [JETP Lett. **73**, 566 (2001)]; Pis'ma Zh. Éksp. Teor. Fiz. **74**, 253 (2001) [JETP Lett. **74**, 231 (2001)].
3. P. G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966; Mir, Moscow, 1968).
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. N. N. Bogolyubov, V. V. Tolmachev, and D. V. Shirkov, *A New Method in the Theory of Superconductivity* (Akad. Nauk SSSR, Moscow, 1958; Consultants Bureau, New York, 1959).
6. P. Philips, Y. Wan, I. Martin, *et al.*, Nature **395**, 253 (1998).

Translated by A. Bagatur'yants

JETP Letters, Vol. 78, No. 4, 2003, pp. 213–217. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 249–252.
Original Russian Text Copyright © 2003 by Magarill, Éntin.

Optical and Photoelectric Properties of Helical Quantum Wires

L. I. Magarill and M. V. Éntin

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

Received July 7, 2003

The appearance of a steady current in a helical quantum wire under circularly polarized light is predicted. The effect is associated with the appearance of a travelling electromagnetic wave in the coordinate system related to the wire, the wavelength being determined by the period of the helix. The gyrotropy of a medium consisting of parallel quantum wires is considered. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 78.67.Lt; 73.63.Nm

In this paper, we consider helical quantum wires of the type shown in the figure and a homogeneous medium consisting of parallel quantum wires of this type. The symmetry of the helix allows the existence of an antisymmetric third-rank tensor. Physically, this tensor corresponds to the transformation of rotation to translational motion. If the rotation is understood as a circular polarization of an electromagnetic wave absorbed in the medium and the translational motion is interpreted as an electron current, one can expect the appearance of a steady current in a helical quantum wire under the effect of light, i.e., a circular photogalvanic effect [1]. A similar tensor determines the gyrotropy of a medium consisting of such wires.

In the optical and infrared frequency ranges, a photon's momentum is, as a rule, small compared to the electron momentum. Therefore, in the theory of the photogalvanic effect, this momentum is usually ignored. Exceptions are the effects directly related to the photon momentum, such as the electron–photon drag, including the resonance drag for which the photon momentum plays the role of a trigger while the electron receives its momentum from a third body.

The helical geometry of the quantum wire leads to an artificial increase in the wave vector of the wave. When the quantum wire is illuminated along its axis by circularly polarized light, for an electron moving through the wire the electromagnetic field represents a travelling wave with the wavelength equal to the period of the helix. The momentum transferred from light to an electron leads to allowed transitions between free electron states, which results in a collisionless absorption of electromagnetic field. As a result of the change in the electron momentum, a steady current appears in the system. In the classical limit, the steady current is explained by the acceleration of electrons moving with the phase velocity of the wave.

Below, we consider the circular photogalvanic effect in a single helical quantum wire and the geometrically induced gyrotropy of a homogeneous medium consisting of parallel quantum wires.

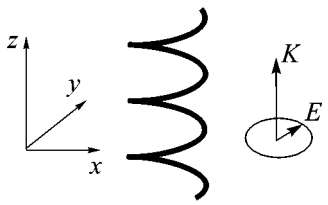
The gyrotropy of molecular systems, i.e., media consisting of chiral organic molecules, had been studied for more than hundred years. Examples of such media are protein and DNA solutions (see, e.g., [2]). Another well-known example of media with a geometrically induced gyrotropy are cholesteric liquid crystals (see, e.g., [3]).

However, it is only recently that the fabrication of nanohelices on the basis of ordinary semiconductors has become possible [4, 5]. The compatibility of these systems with elements of semiconductor electronics makes their study important for future practical applications.

Absorption of light in a helical quantum wire. Consider a thin helical quantum wire with a constant cross section. This wire can be described by the equation

$$\mathbf{r} \equiv \mathbf{a}(q) \equiv (R\cos kq,\ R\sin kq,\ \eta q). \qquad (1)$$

Here, R is the radius of the helix, q is the coordinate along the wire, the sign of k determines the helix direction $\xi = \pm 1$ (the plus sign corresponds to a left-hand helix), the length of the helix turn is equal to $2\pi/|k|$, η is the ratio of the pitch of the helix to the length of its turn, and the axis of the helix is directed along the z axis.



Helical quantum wire.

0021-3640/03/7804-0213$24.00 © 2003 MAIK "Nauka/Interperiodica"

The Hamiltonian of a curved quantum wire of constant cross section in the absence of external field was derived in the general form in [6, 7]. For the helical geometry determined by Eq. (1), when the spin and the band bottom displacement due to curvature and torsion are ignored, electrons are described by an ordinary one-dimensional quadratic Hamiltonian $H = p^2/2m$, where $p = -i\partial/\partial q$ is the operator of the electron momentum along the wire. In the presence of an external electromagnetic field **A**, the gradient transformation of the Hamiltonian leads to the replacement of the momentum by $p + e/cA_\mathbf{t}$, where $A_\mathbf{t}$ is the projection of the vector potential on the tangent to a quantum wire with the unit vector $\mathbf{t}(q) = (-kR\sin kq, kR\cos kq, \eta)$. We stress that, in our study, we neglect the transitions between transverse electron states in the wire. In the following calculations, we set $\hbar = 1$.

We consider a helical quantum wire in a homogeneous alternating external electric field $\mathbf{E}(t) = \mathrm{Re}(\mathbf{E}_0 e^{-i\omega t})$ and neglect the wave vector **K** of the wave. Only the tangential field component acts on the electrons in the wire:

$$E_\mathbf{t} = -\dot{A}_\mathbf{t}/c$$
$$= \mathrm{Re}(kR(-E_{0,x}\sin kq + E_{0,y}\cos kq) + \eta E_{0,z})e^{-i\omega t}.$$

Thus, with reference to the one-dimensional Hamiltonian, the fields E_x and E_y are waves acting on an electron and characterized by the wave vector k while the field E_z is a homogeneous field along the helix axis. For a free electromagnetic field of optical or lower frequencies, the initial wave vector of the wave is small. At the same time, the wave vector of the helix k can be made sufficiently large, in particular, comparable to the Fermi momentum of electrons. As a result, optical absorption is allowed even for free electrons.

The dynamic current caused by the field can be divided into contributions from corresponding spatial harmonics:

$$J(q, t) = \mathrm{Re}(kR\sigma(\omega, k)(-E_{0,x}\sin kq + E_{0,y}\cos kq) + \eta\sigma(\omega, 0)E_{0,z}e^{-i\omega t}). \quad (2)$$

Here, $\sigma(\omega, k) = \sigma(\omega, -k)$ is the dynamic conductivity of a one-dimensional system with allowance for the spatial dispersion, which is described by the Kubo formula:

$$\sigma(\omega, k) = \frac{2e^2}{m^2 L} \times \sum_{p, p'} \frac{f(\varepsilon) - f(\varepsilon')}{\varepsilon' - \varepsilon} |(\mathcal{U})_{p, p'}|^2 \frac{1}{\delta + i(\varepsilon - \varepsilon' - \omega)}, \quad (3)$$

where $\varepsilon = p^2/2m$ is the electron energy, $f(\varepsilon)$ is the Fermi function, $\mathcal{U} = \{p, e^{ikq}\}$ (the braces denote the symmetrization operation), L is the total length of the helix, and δ is the rate at which the field is turned on. Neglecting the scattering of electrons, for the real part of conductivity we obtain

$$\mathrm{Re}\sigma(\omega, k) = \frac{4e^2 m\omega}{|k|^3} \times \left[f\left(\frac{(k^2 - 2m\omega)^2}{8mk^2}\right) - f\left(\frac{(k^2 + 2m\omega)^2}{8mk^2}\right)\right]. \quad (4)$$

The time average power absorbed from a wave of amplitude E_0 propagating along the axis of the helix is equal to $\mathrm{Re}(\sigma)k^2R^2E_0^2/4$ in the case of linear polarization and to twice this value in the case of circular polarization. At a low temperature, the absorption in a degenerate electron gas differs from zero in the frequency band $|p_F|k| - k^2/2| < m\omega < (p_F|k| + k^2/2)$, which is determined by the energy and momentum conservation laws. For small k, the absorption band narrows to a single frequency $\omega = |k|v_F$.

Photogalvanic effect in a helical wire. Let us consider the formation of a steady current in the external field of a circularly polarized wave propagating along the helix axis (the z axis). In this section, as above, we neglect the wave vector of the wave. A steady current appears in a weak electromagnetic field as a second-order response to the external electric field.

The complex amplitude of the electric field can be represented as $\mathbf{E}_0 = (1, i\zeta, 0)E_0$, where $\zeta = \pm 1$, and

$$A_\mathbf{t} = \frac{cE_0R\zeta k}{\omega}\cos(\zeta kq - \omega t). \quad (5)$$

To calculate the current, we use the Kubo formula, which is quadratic in the external field. The problem is solved by analogy with the determination of the circular current in a quantum ring in an alternating external field [8]. Neglecting the electron scattering and retaining a small but finite rate δ of turning on the field, we obtain the steady current along the wire:

$$J_0 = \frac{e^3k^2R^2E_0^2}{2m^3\omega^2 L}\mathrm{Re}\sum_{p, p'}(p - p')\frac{f(\varepsilon) - f(\varepsilon')}{2\delta} \times |(\mathcal{U})_{p, p'}|^2 \frac{1}{\delta + i(\varepsilon - \varepsilon' - \omega)}. \quad (6)$$

The electron scattering can be simulated by the rate of turning on the field through the introduction of the

electron momentum relaxation time τ: $\delta \longrightarrow 1/2\tau$. Then, from Eq. (6), we obtain

$$J_0 = \frac{e^3\zeta k}{m^3\omega^2}(1-\eta^2)E_0^2\int_{-\infty}^{\infty} dp(p+k/2)^2 \times \tau(f(\varepsilon+\omega)-f(\varepsilon))\delta(kp/m+k^2/2m-\omega). \quad (7)$$

Using Eq. (4), for a constant τ, we derive

$$J_0 = -\frac{e}{4m\omega}\zeta\tau k\,\mathrm{Re}[\sigma(\omega,k)](1-\eta^2)E_0^2. \quad (8)$$

In the low-temperature limit, the frequency dependence of the current has the form of a rectangular peak $|p_F|k| - k^2/2| < m\omega < (p_F|k| + k^2/2)$. The sign of the current is determined by the product $\zeta\xi$; i.e., a change in the direction of polarization or helix changes the sign of the current. With $k \longrightarrow 0$, the width of the peak tends to zero faster than its position; i.e., the peak becomes infinitely narrow.

In the classical limit when $k \ll p_F$ and $\omega \ll \varepsilon_F$, a more accurate approximation taking into account, in particular, the energy dependence of the relaxation time can be obtained in terms of the kinetic equation. On the assumption that the electron scattering is elastic, the kinetic equation for the distribution function $f(p, q, t)$ has the form

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial q} - eE_t\frac{\partial f}{\partial p} = -\frac{f_-}{\tau}, \quad f_- = \frac{f(p)-f(-p)}{2}. \quad (9)$$

Here, $v = p/m$. Equation (9) represents a pair of coupled differential equations. The steady current is determined by the stationary distribution function component that is odd in the momentum and of the second order in the wave field. Solving Eq. (9) in the quadratic approximation in E_t, we obtain the following expression for the stationary part of the distribution function:

$$f_- = \frac{1}{2}\tau e^2(kRE_0)^2\zeta k v\frac{\partial}{\partial p} \times\left(\frac{\omega\tau}{\omega^2+\tau^2(\omega^2-(\zeta k v)^2)^2}v\frac{\partial f_0}{\partial\varepsilon}\right). \quad (10)$$

From this expression, for the steady current along the wire we obtain

$$J_0 = -\frac{e}{\pi}\int dp\, v f_-(p) = \frac{2e^3\zeta k}{\pi m^2}(1-\eta^2)E_0^2 \times\int dp\frac{\omega\tau}{(\omega^2-k^2v^2)^2\tau^2+\omega^2}\varepsilon\frac{\partial(\tau\varepsilon)}{\partial\varepsilon}\frac{\partial f_0}{\partial\varepsilon}. \quad (11)$$

With $\tau \longrightarrow \infty$, the current exhibits a resonance in frequency when the Fermi velocity of electrons v_F coincides with the phase velocity of the wave $s = \xi\zeta\omega/|k|$. The resonance is related to the fact that the uniform acceleration on the crest of the wave accelerates the electrons. This contribution, being combined with the contribution from the elastically scattered electrons, gives rise to a direct current along the phase velocity of the wave.

In this limit, at a constant relaxation time, Eq. (11) coincides with Eq. (8). Outside the resonance, we have

$J_0 \sim k/\omega^3$ for $\omega \gg |k|v_F$, $\omega\tau \gg 1$;

$J_0 \sim k\tau^2/\omega$ for $\omega \gg |k|v_F$, $\omega\tau \ll 1$ or for $\omega \ll |k|v_F$, $k^2v_F^2\tau \ll \omega$; and

$J_0 \sim \omega/k^3v_F^4$ for $\omega \ll |k|v_F$, $k^2v_F^2\tau \gg \omega$.

When $\omega \longrightarrow 0$ or $k \longrightarrow 0$, the current tends to zero.

We stress that our consideration of the photogalvanic effect was limited to the kinetic approach and neglected the localization of states in the one-dimensional system. This approach is valid when the localization is destroyed by a phase failure, in particular, when the phase failure time is comparable to τ (e.g., in a bounded system) or smaller than τ.

Strong field limit. Formulas presented above refer to a response quadratic in the field. Let us consider the case of a strong external field, when the potential produced by the field exceeds the Fermi energy while the phase velocity of the wave is smaller than the Fermi velocity or close to it in order of magnitude. In this limit, we can assume that the field forces the electrons to concentrate near the potential minima. Then, all electrons will move with the wave velocity. As a result, the direct current can be expressed as

$$J_0 = -en_e s, \quad (12)$$

where n_e is the mean linear density of electrons. Expression (12) depends on neither electric field nor electron relaxation.

This result also holds when the wave potential ceases being classical for electrons. Performing a gradient transformation, we find that the electron wave function satisfies the equation

$$\left[-i\frac{\partial}{\partial t} + \frac{\hat{p}^2}{2m} + eE_0R\cos(\zeta kq-\omega t)\right]\psi = 0. \quad (13)$$

In the moving coordinate system, an electron is described by the Bloch function $\phi_p(q) = u_p(q)e^{ipq}$ (with the quasi-momentum p and the energy ε_p corresponding to a certain energy band), which is a solution to the Mathieu equation

$$-\frac{1}{2m}\frac{d^2\phi_p}{dq^2} + eE_0R\cos(kq)\phi_p = \varepsilon_p\phi_p. \quad (14)$$

The functions ψ and u_p are related by the formula

$$\psi(q, t) = u_p(q - st)e^{imsq}e^{-it(\varepsilon_p + sp + ms^2/2)}. \quad (15)$$

In the laboratory coordinate system, the electron energy is equal to $\varepsilon_p + sp + ms^2/2$ and the electron velocity is $\partial\varepsilon_p/\partial p + s$. When the wave amplitude is large, equilibrium is established in the coordinate system related to the wave, or, at least, the distribution function becomes a function of only the energy ε_p. Then, for the steady current we again obtain Eq. (12).

Expression (12) corresponds to the quantization of the electron transport in an adiabatic electron pump [9], if each turn of the helix is considered as an individual electron pump. Indeed, according to Eq. (12), each occupied state in a helix turn should pump exactly one electron within the field variation period.

Gyrotropy of a medium consisting of helical quantum wires. In this section, we consider the gyrotropy of a three-dimensional medium consisting of identical parallel helical quantum wires. In formal terms, the symmetry of a homogeneous medium formed of helices allows the existence of an antisymmetric third-rank tensor $\gamma_{ijk} = -\gamma_{jik}$, which appears as a result of expanding the spatially dispersive dielectric constant $\varepsilon_{ij}(\mathbf{K})$ in the wave vector of the electromagnetic field $\mathbf{K}$:

$$\varepsilon_{ij}(\mathbf{K}) = \varepsilon_{ij}(0) + i\gamma_{ijk}K_k + \ldots.$$

A homogeneous medium consisting of helices possesses a symmetry D_∞, for which the tensor γ_{ijk} has only two independent components: γ_{xyz} and $\gamma_{yzx} = -\gamma_{xzy}$. To determine these components, it is sufficient to consider linearly polarized waves with the wave vectors directed along the helix axis and normally to it.

In the first case, for a wave linearly polarized along the x axis (this can be any chosen axis perpendicular to the helix axis), the tangential component of the electric field has the form

$$E_\mathbf{t} = -kRE_0\mathrm{Re}[e^{iK\eta q - i\omega t}\sin(kq)]. \quad (16)$$

The current J in a helix is determined by the superposition of the responses to two waves:

$$J = \mathrm{Re}\left[i\frac{kR}{2}E_0e^{i(-\omega t + \eta Kq)} \times (\sigma(\omega, k + \eta K)e^{ikq} - \sigma(\omega, -k + \eta K)e^{-ikq})\right]. \quad (17)$$

The volume current density $\mathbf{j}$ is determined by the current vector $\mathbf{t}J$ averaged over a helix turn and multiplied by the helix density n:

$$\mathbf{j} = n\langle \mathbf{t}J\rangle, \quad \langle\ldots\rangle = \frac{k}{2\pi}\int_0^{2\pi/k}(\ldots)dq. \quad (18)$$

For our purpose, it is sufficient to determine the y component of the current density

$$j_y = \mathrm{Re}[\sigma_{yx}(\omega, K)E_0e^{i(kZ - \omega t)}],$$
$$\sigma_{yx}(\omega, K) = in\frac{k^2R^2}{4}(\sigma(\omega, k + \eta K) - \sigma(\omega, k - \eta K)). \quad (19)$$

From Eq. (19), we obtain

$$\gamma_{yxz} = -\gamma_{xyz} = 2in\frac{\pi}{\omega}(1 - \eta^2)\eta\frac{\partial}{\partial k}\sigma(\omega, k). \quad (20)$$

In the second case, for a wave with $\mathbf{K} = (K, 0, 0)$, with a linear polarization along the z axis, the quantity $E_\mathbf{t}$ is determined by the expression (for a helix whose axis passes through the point $\boldsymbol{\rho}_l = (x_l, y_l)$)

$$E_t^{(l)} = \mathrm{Re}[e^{iK(x_l + R\cos(kq)) - i\omega t}\eta E_0] \approx \mathrm{Re}[e^{i(Kx_l - \omega t)}(1 + iKR\cos(kq))\eta E_0]. \quad (21)$$

Formula (21) is expanded with allowance for the smallness of the parameter KR. The volume current density in a medium consisting of helices has the form

$$j_y = \sum_l \langle t_y J^{(l)}\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_l - \mathbf{a}_\perp(q))\rangle, \quad \mathbf{a}_\perp(q) = (a_x(q), a_y(q)), \quad (22)$$

where

$$J^{(l)} = \mathrm{Re}\{E_0\eta e^{i(Kx_l - \omega t)}[\sigma(\omega, 0) + iKR\sigma(\omega, k)\cos kq]\}. \quad (23)$$

By averaging over the positions of the helices, we obtain

$$\gamma_{yzx} = -\gamma_{xzy} = 2in\frac{\pi}{\omega k}\eta(1 - \eta^2)(\sigma(\omega, k) - \sigma(\omega, 0)). \quad (24)$$

Formulas (20) and (24) together with expression (3) determine the gyrotropy of the system. Evidently, the circular dichroism and the rotation of the polarization plane are enhanced in the vicinity of the resonance.

This work was supported by the Russian Foundation for Basic Research, project no. 02-02-16377, and by the Programs of the Russian Academy of Sciences and

the Ministry of Science and Technology of the Russian Federation.

REFERENCES

1. V. I. Belinicher and B. I. Sturman, Usp. Fiz. Nauk **130**, 415 (1980) [Sov. Phys. Usp. **23**, 199 (1980)]; E. L. Ivchenko and G. E. Pikus, in *Semiconductor Physics*, Ed. by V. M. Tushkevich and V. Ya. Frenkel (Consultants Bureau, New York, 1986), p. 427.
2. M. V. Vol'kenshteĭn, *Biophysics* (Nauka, Moscow, 1981).
3. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, Cambridge, 1977; Mir, Moscow, 1980).
4. V. Ya. Prinz, V. A. Seleznev, A. K. Gutakovsky, *et al.*, Physica E (Amsterdam) **6**, 828 (2000).
5. V. Ya. Prinz, D. Grutzmacher, A. Beyer, *et al.*, Nanotechnology **12**, S1 (2001).
6. M. V. Entin and L. I. Magarill, Phys. Rev. B **66**, 205308 (2002).
7. L. I. Magarill and M. V. Éntin, Zh. Éksp. Teor. Fiz. **123**, 867 (2003) [JETP **96**, 766 (2003)].
8. L. I. Magarill and A. V. Chaplik, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 607 (1999) [JETP Lett. **70**, 615 (1999)].
9. D. J. Tauless, Phys. Rev. B **27**, 6083 (1983).

Translated by E. Golyamina

JETP Letters, Vol. 78, No. 4, 2003, pp. 218–222. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 253–257.
Original Russian Text Copyright © 2003 by Rodionov, Kravtsova, Mandel'.

On the Influence of Strong Electric and Magnetic Fields on Spatial Dispersion and Anisotropy of the Optical Properties of a Semiconductor

V. N. Rodionov, G. A. Kravtsova, and A. M. Mandel'

Moscow State Geological Prospecting University, ul. Miklukho-Maklaya 23, Moscow, 118873 Russia

Received July 10, 2003

Nonperturbative methods were used to construct a single-integral equation for the photon polarization operator in a semiconductor in strong electric and magnetic fields of arbitrary orientations. Spatial dispersion was taken into account. The external field configurations at which its influence could be ignored were found. Asymptotic equations for the field dependences of the absorption coefficient and refractive index were obtained. The neglect of the spin of particles was shown to lead to a fundamentally false idea of magnetic field effects on the optical properties of nonconducting crystals. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 78.20.Bh; 78.20.Ci; 78.20.Jq; 78.20.Ls

Studies of photon propagation in the vacuum in strong electromagnetic fields initiated more than 30 years ago [1, 2] still remain a topical research area [3, 4]. At the same time, optical fiber technologies and many related developments stimulate investigation of the influence of strong electromagnetic fields on the optical properties of nonconducting crystals. These problems are discussed, for instance, in monograph [5] (also see [6]). The most general and consistent approach to solving them is through studying the photon polarization operator in a semiconductor that experiences the action of external electric and magnetic fields of arbitrary orientations. We assume that the intensity of the fields is only limited by the character of the nonrelativistic approximation accepted in solid-state physics, and the only photoabsorption mechanism is direct allowed interband transitions. A scheme for one-loop approximation calculations of the photon polarization operator in an isotropic semiconductor with a simple band in given in monograph [7] for an external magnetic field without taking into account the spin of particles. In our problem, the virtual electron and hole wave functions are given by solutions to the Schrödinger equation with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m_{c,v}}\Delta - i\frac{\hbar eHx}{m_{c,v}c}\frac{\partial}{\partial y} + e\mathbf{E}\cdot\mathbf{r} + \frac{\hbar eH}{m_{c,v}c}\hat{S}_z,$$

where $\hbar$, e, and c are the Planck constant, the charge of the electron, and the velocity of light, respectively; m_c (m_v) is the effective mass of the electron (hole); Δ is the Laplace operator; magnetic field **H** is directed along the z axis; **E** is the electric field vector; and $\hat{S}_z$ is the particle spin operator. These wave functions are easy to obtain using relativistic solutions to the equations of motion of charged particles in electromagnetic fields of a complex configuration; such solutions were studied in detail in monograph [8] (also see [9]). In this approach, the influence of external fields is taken into account exactly, and perturbation theory is only used to describe interparticle interactions.

An equation in the form of a double series in the squares of Laguerre polynomials, similar to the equation for the polarization operator in a purely magnetic field, can be obtained following the procedure described in [7]. Such an equation is, however, difficult to deal with and use for drawing qualitative conclusions from it. In particular, because this representation is cumbersome, its regularization is impeded.

At the same time, using the method for summing the contributions of all Landau levels developed in [10] allows the photon polarization operator in superimposed electric and magnetic fields to be described by the single-integral equation

$$\Pi = \Pi_0\left\{\frac{\exp(-i\pi/4)}{(8\pi)^{1/2}} \times \int_0^\infty \frac{dx}{x^{3/2}}\left[\frac{hx f_\sigma(hx)}{\sin(hx)}\exp(iS) - 1\right] + 2\Theta(-I)\right\}. \quad (1)$$

Here, Π_0 is the modulus of the photon polarization operator in the absence of external fields, which contains the convolution of the polarization vector over the Bloch functions for electron and hole band edges [7],

and I is the energy released in the decay of the photon to the electron–hole pair,

$$I = \hbar\omega - \mathscr{E} - (\hbar k)^2/2M, \tag{2}$$

where $\hbar\omega$ is the photon energy, $\mathscr{E}$ is the forbidden band width, $M = m_c + m_v$, and k is the modulus of the photon wave vector.

The f_σ functions depend on the spin of charged particles $\sigma = 0$; 1/2 and are $f_0 = 1$ for scalar particles and $f_{1/2} = \cos(m^*hx/m_c)\cos(m^*hx/m_v)$ for the real electron and hole, $m^* = m_cm_v/M$ is the reduced mass of the pair,[1] and $\Theta(x)$ in (1) is the unit Heaviside step function. Lastly, the S value has the form

$$S = 2\,\mathrm{sgn}(I)x + \frac{2M^2\varepsilon^2}{m^{*2}h^2} \times\left[x\frac{m^*}{M} - \frac{\sin(xhm^*/m_c)\sin(xhm^*/m_v)}{h\sin(xh)}\right], \tag{3}$$

where ε and h are

$$\varepsilon = |\mathbf{E} + \boldsymbol{\kappa}\times\mathbf{H}|/E_k, \quad h = H/H_k.$$

In turn,

$$E_k = (2m^*|I|^3)^{1/2}/e\hbar, \quad H_k = m^*c|I|/e\hbar,$$
$$\boldsymbol{\kappa} = \hbar\mathbf{k}/Mc.$$

If $I > 0$, interband transitions can also occur in the absence of external fields; we then have $\Pi = i\Pi_0$. If $I < 0$, pairs cannot be created if no external field is applied, and the polarization operator is real, $\Pi = \Pi_0$.[2] The last term in (1) describes *spatial dispersion* manifestations [5, 7] without taking temperature effects into account.

Note that the contribution of the k_z component oriented along the magnetic field is only characteristic of certain separate polarization operator terms and is related to the representation of the operator by the sum over Landau levels [6, 7]. After the summation over all levels, the dependence of the polarization operator also on the $k_\perp$ transverse component [see (1) and (3)] is restored. The h parameter is the relative magnetic field strength in the natural critical field H_k units. The H_k field is determined from the condition of the equality of the cyclotron quantum value and the energy release in the reaction. The numerator ε is simply the modulus of the classical *Lorentz force* divided by the charge of the electron. Its second term is responsible for the dependence of the polarization operator on the direction of the motion of the photon and, therefore, *magnetic-field-induced anisotropy* of the optical properties of semiconductors. Such problems are as a rule solved in the approximation ignoring the contribution determined by the photon momentum (e.g., see [9, 12]). This corresponds to the absence of the second term in ε and the last term in I, which results in the neglect of spatial dispersion in the polarization operator.

For optical photons, $\kappa \sim 10^{-6}$ at $m_c \sim m_v \sim m$ (where m is the mass of the vacuum electron); that is, such an approximation is seemingly justified. It has, however, been shown in [7] for a purely magnetic external field that this approximation is inapplicable to real semiconductors, especially in near-threshold regions. In our problem, spatial dispersion and induced anisotropy can only be ignored if $E \geq H$.[3]

Consider the asymptotic estimates of the photon polarization operator [Eq. (1)] and the equations for the refractive index and extinction ratio that follow from (1).[4]

1. $E \ll E_k$, $H \ll H_k$, and $H \leq E$; or $h \ll \varepsilon \ll 1$. This corresponds to weak fields with the predominance of the electric field. Spatial dispersion and anisotropy can be ignored on the assumption $\varepsilon = E/E_k$. If the energy release is positive, $I > 0$, the polarization operator is estimated as

$$\Pi \approx \Pi_0 i\left[1 + \frac{\varepsilon^2}{32} - \frac{\varepsilon}{4}\phi_\sigma^+(h,\varepsilon)\exp\left(i\frac{4}{3\varepsilon}\right)\right], \tag{4}$$

where

$$\phi_\sigma^+(h,\varepsilon) = 1 + r_\sigma\frac{h^2}{\varepsilon^2}$$

and the spin-dependent r_σ constant coefficients are

$$r_0 = 1/6 \text{ and } r_{1/2} = -(m_c^2 + m_v^2 - m_cm_v)/3M^2.$$

Following [18], let us use (1) to determine the observed differences of the refractive index and extinction ratio from their values when the fields are "switched off," $\Delta n = n - n_0$ and $\Delta q = q - q_0$. The usual

[1] "Spinless" charged particles are included because the electron and hole spins are usually ignored in such calculations. It will be shown below that this is a fundamentally incorrect approach if a magnetic field is present.

[2] Note that, in a purely electric external field, different energy release signs correspond to the oscillation and tunneling modes thoroughly studied for the well-known Keldysh–Frants effect [11].

[3] The ε value then becomes the relative electric field strength in critical field E_k units; the critical field is determined from the condition of the equality of the work done by the field at the de Broglie wavelength $\hbar/(2m^*|I|)^{1/2}$ and the energy release in the reaction. Note that ε and h only differ in the scale of energy parameters from the similar values that arise in the theory of the field ionization of atoms [13].

[4] The methods for obtaining these estimates were described in detail in [10–17].

equation that relates the optical characteristics to the polarization operator (see [7, 9]),

$$(n + iq)^2 = 1 + \frac{4\pi c^2}{\omega^2}\Pi,$$

yields

$$\Delta n = \frac{\varepsilon n_0 (n_0^2 - 1)^{1/2}}{4(2n_0^2 - 1)}\left\{\frac{\varepsilon}{8}(n_0^2 - 1)^{1/2} + \phi_\sigma^+(h, \varepsilon) \times \left[n_0 \sin\left(\frac{4}{3\varepsilon}\right) - (n_0^2 - 1)^{1/2}\cos\left(\frac{4}{3\varepsilon}\right)\right]\right\}, \quad (5)$$

$$\Delta q = \frac{\varepsilon q_0 (q_0^2 + 1)^{1/2}}{4(2q_0^2 + 1)}\left\{\frac{\varepsilon}{8}(q_0^2 + 1)^{1/2} - \phi_\sigma^+(h, \varepsilon)\left[q_0 \sin\left(\frac{4}{3\varepsilon}\right) + (q_0^2 + 1)^{1/2}\cos\left(\frac{4}{3\varepsilon}\right)\right]\right\}. \quad (6)$$

Clearly, these equations contain both terms that monotonically grow as the electric field increases and terms that oscillate at a frequency determined by the ε parameter. The magnetic field gives a correction to the amplitude of these oscillations.

If the energy release is negative, $I < 0$, that is, under tunneling conditions, (1) gives

$$\Pi \approx \Pi_0\left[1 + \frac{\varepsilon^3}{32} + i\frac{\varepsilon}{8}\phi_\sigma^-(h, \varepsilon)\right], \quad (7)$$

where

$$\phi_\sigma^-(h, \varepsilon) = \left(1 - r_\sigma\frac{h^2}{\varepsilon^2}\right)\exp\left(-\frac{4}{3\varepsilon}\right),$$

and external field-induced changes in the optical properties of a semiconductor are[5]

$$\Delta n = \frac{\varepsilon^2}{64}\left(n_0 - \frac{1}{n_0}\right), \quad (8)$$

$$\Delta q = q = \frac{\varepsilon}{16}\left(n_0 - \frac{1}{n_0}\right)\phi_\sigma^-(h, \varepsilon). \quad (9)$$

It follows from (8) that the refractive index monotonically grows as the electric field increases, whereas absorption coefficient (9) is exponentially suppressed. The magnetic field gives a preexponential correction. Note that the signs of r_σ are different for scalar and spinor particles even in the limit of comparatively weak fields. Indeed, the magnetic field decreases and slightly increases photoabsorption for scalar and spinor particles, respectively. While this is not very important under oscillation conditions, the neglect of particle spins in the situation under consideration leads to an *incorrect idea of the role played by the magnetic field* in the processes under consideration. The situation in the theory of atomic field ionization is fully analogues, as it was established recently in [15, 19]. Therefore, the particle spins should be taken into account, despite the relativistic character of calculation and rather weak magnetic fields.

2. $E \ll E_k$, $H \ll H_k$, and $E \ll H$; or $\varepsilon \ll h \ll 1$. The fields are again weak, but the magnetic field now predominates. If $I > 0$, the photon polarization operator can be written in the form

$$\Pi \approx \Pi_0 i(1 + \chi_\sigma^+/16), \quad (10)$$

where

$$\chi_\sigma^+ = r_\sigma h^2 + \varepsilon^2/2,$$

and external field-induced corrections to the optical characteristics are given by the equations

$$\Delta n = \frac{(n_0^2 - 1)^2}{64 n_0 (2n_0^2 - 1)}\chi_\sigma^+, \quad (11)$$

$$\Delta q = \frac{q_0(q_0^2 + 1)}{16(2q_0^2 + 1)}\chi_\sigma^+. \quad (12)$$

It follows from (10) that the polarization operator is purely imaginary and oscillating terms are absent. Conversely, if $I < 0$, the polarization operator becomes real,

$$\Pi \approx \Pi_0(1 - \chi_\sigma^-/16), \quad (13)$$

where

$$\chi_\sigma^- = r_\sigma h^2 - \varepsilon^2/2.$$

In the situation that we are considering, field corrections to the refractive index can be written as

$$\Delta n = -\frac{n_0^2 - 1}{32 n_0}\chi_\sigma^-, \quad (14)$$

and the absorption coefficient q is negligibly small. The magnetic and electric Lorentz force components of ε are commensurate, and semiconductors can therefore become optically anisotropic. Indeed, the optical characteristics [Eqs. (11), (12), and (14)] depend on both

[5] It is here taken into account that there can be no photoabsorption in the absence of an external field under these conditions and therefore $q_0 = 0$.

the direction of photon motion and the mutual orientation of the **E** and **H** vectors.

Consider the simplest mutual orientations of **E**, **H**, and **k**. If the **E** and **H** vectors are parallel, a certain direction in space becomes distinguished. The largest difference in the optical properties is then observed for photons that move along this direction and normally to it. For crossed fields, that is, for $\mathbf{E} \perp \mathbf{H}$, the smallest Δn and Δq values are observed for photons that move in the direction of the $\mathbf{E} \times \mathbf{H}$ vector product, and the maximum differences, for photons that move in the opposite direction. Roughly, this resembles the interaction of a photon with a running electromagnetic wave [8].

Formally, the situation that we are considering admits the passage to the limit of a purely magnetic field. It is, however, well known (e.g., see [10, 20]) that, in such a field in the $h \longrightarrow 0$ limit, oscillations with an unlimited amplitude appear in the equations for the probabilities of pair photogeneration. Note that precisely these divergences were used in [7] to prove the importance of taking spatial dispersion into account. Indeed, for a photon that moves along the magnetic field ($\mathbf{H} \times \mathbf{k} = 0$) in the absence of an electric field, ε vanishes. It follows from the representation of $\sin^{-1}(h)$ by the exponential series

$$\sin^{-1}(hx) = 2i\sum_{n=0}^{\infty}\exp[-ihx(1+2n)]$$

that the exponent for S in (3) for spinless particles is representable in the form of a linear function that vanishes at $h = 2/(1 + 2n)$. If the spin is taken into account, the equation becomes more cumbersome, but the situation does not change radically.

If we leave a weak electric field in (3), then the amplitude of polarization operator (1) oscillations at the specified h values is limited by $\Pi \sim h\varepsilon^{-1/3}$. It follows that, generally speaking, the electric field plays the role of a *regularizing factor*. This in essence follows from the pioneer result due to Keldysh [11], namely, the introduction of an external electric field removes the root singularity at the threshold of the photoabsorption cross section of nonconducting crystals, which is present in the absence of a field. The same property was mentioned by Ritus [21] for the relativistic reaction of the photocreation of electron–positron pairs in crossed external fields ($\mathbf{E} \perp \mathbf{H}$, $E = H$). In addition, the method for removing the abovementioned divergences in a magnetic field was related in [20] to including Coulomb interaction of charged particles.

It follows that the regularizing role played by an electric field manifests itself in a fairly wide range of photoprocesses and persists at various coordinate dependences of **E**. Also note that, if a magnetic field is switched off while an electric field is retained, no nonphysical singularities arise [14].

3. $E \gg E_k$ and $H \leq H_k$, or $\varepsilon \gg 1$ and $h \leq 1$. We here have a strong electric and a moderate magnetic field. Importantly, the field can turn strong not only because of its large amplitude, but also because the energy release in the reaction is small [see the explanatory notes to (3)]. In other words, an arbitrary field is "strong" in the vicinity of the threshold [10]. In this limit, it follows from (1) irrespective of the spin and the energy release sign (cf. [15, 16]) that

$$\begin{aligned}\Pi &\approx \Pi_0\frac{1}{2\pi^{1/2}}\left(\frac{3}{2}\right)^{1/3}\Gamma\left(\frac{5}{6}\right)\left(-\frac{1}{3^{1/2}}+i\right)\varepsilon^{1/3}\\ &\approx 0.365\Pi_0\left(-\frac{1}{3^{1/2}}+i\right)\varepsilon^{1/3}.\end{aligned} \tag{15}$$

The absorption in this limit is strong,

$$\begin{aligned}q &\approx \frac{3^{5/12}}{2^{2/3}\pi^{1/4}}\Gamma^{1/2}\left(\frac{5}{6}\right)q_0^{1/2}(1+q_0^2)^{1/4}\varepsilon^{1/6}\\ &\approx 0.795q_0^{1/2}(1+q_0^2)^{1/4}\varepsilon^{1/6},\end{aligned} \tag{16}$$

and the determination of the refractive index becomes meaningless.

Note in conclusion that, according to monograph [9], photoabsorption in crossed fields becomes exponentially suppressed as electric field E increases. This conclusion is drawn from the representation of the absorption coefficient in the form of a double series in Laguerre polynomials, see above. Our results give evidence to the contrary. In addition, the author himself notes (see [9], p. 452) that this conclusion is at variance with the Keldysh–Frants equations, verified many times experimentally. In the approach developed in this work, switching a magnetic field on and the passage to the Keldysh–Frants equations do not require additional transformations (also see [14]). If the magnetic field is strong ($h \gg 1$), the dependence of the polarization operator on the spin of particles is most significant and requires a separate consideration.

The authors thank the participants of the seminar headed by I.M. Ternov for useful discussions. This work was financially supported by the Russian Foundation for Basic Research (project no. 02-02-16784).

REFERENCES

1. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **46**, 776 (1964) [Sov. Phys. JETP **19**, 529 (1964)]; Zh. Éksp. Teor. Fiz. **46**, 1768 (1964) [Sov. Phys. JETP **19**, 1191 (1964)].
2. A. E. Shabad, Ann. Phys. (N.Y.) **90**, 166 (1975).
3. N. V. Mikheev and M. V. Chistyakov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 726 (2001) [JETP Lett. **73**, 642 (2001)].
4. A. E. Lobanov and A. R. Muratov, Zh. Éksp. Teor. Fiz. **123**, 757 (2003) [JETP **96**, 669 (2003)].

5. V. M. Agranovich and V. L. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons*, 2nd ed. (Nauka, Moscow, 1979; Springer, New York, 1984).
6. L. I. Korovina and A. E. Shabad, Zh. Éksp. Teor. Fiz. **67**, 1032 (1974) [Sov. Phys. JETP **40**, 512 (1975)].
7. A. E. Shabad, Tr. Fiz. Inst. im. P.N. Lebedeva, Akad. Nauk SSSR **192**, 5 (1988).
8. I. M. Ternov, V. R. Khalilov, and V. N. Rodionov, *Interaction of Charged Particles with Strong Electromagnetic Field* (Mosk. Gos. Univ., Moscow, 1982).
9. A. I. Ansel'm, *Introduction to the Theory of Semiconductors* (Nauka, Moscow, 1978).
10. V. N. Rodionov, Zh. Éksp. Teor. Fiz. **113**, 23 (1998) [JETP **86**, 11 (1998)].
11. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1964)].
12. I. J. Berson, J. Phys. B: At. Mol. Phys. **8**, 3078 (1975).
13. V. S. Popov, B. M. Karnakov, and V. D. Mur, Zh. Éksp. Teor. Fiz. **113**, 1579 (1998) [JETP **86**, 860 (1998)].
14. V. N. Rodionov and A. M. Mandel', Vestn. Mosk. Univ., Fiz., Astron., No. 3, 25 (2001).
15. V. N. Rodionov, G. A. Kravtsova, and A. M. Mandel', Pis'ma Zh. Éksp. Teor. Fiz. **75**, 435 (2002) [JETP Lett. **75**, 363 (2002)].
16. V. N. Rodionov, G. A. Kravtsova, and A. M. Mandel', Vestn. Mosk. Univ., Fiz., Astron., No. 5, 6 (2002).
17. V. G. Kadyshevskiĭ, V. N. Rodionov, G. A. Kravtsova, and A. M. Mandel', Teor. Mat. Fiz. **134**, 227 (2002).
18. D. E. Aspnes, Phys. Rev. **147**, 554 (1966).
19. V. N. Rodionov, G. A. Kravtsova, and A. M. Mandel', Dokl. Akad. Nauk **386**, 753 (2002) [Dokl. Phys. **47**, 725 (2002)].
20. A. E. Lobanov and A. R. Muratov, Zh. Éksp. Teor. Fiz. **87**, 1140 (1984) [Sov. Phys. JETP **60**, 651 (1984)].
21. A. I. Nikishov and V. I. Ritus, Tr. Fiz. Inst. im. P.N. Lebedeva, Akad. Nauk SSSR **111**, 84 (1979).

Translated by V. Sipachev

JETP Letters, Vol. 78, No. 4, 2003, pp. 223–229. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 258–264.
Original Russian Text Copyright © 2003 by Perekrestov, Kosminskaya.

The Phase and Morphological Inhomogeneity of the Ti–C System Deposits Formed at the Plasma–Titanium Substrate Interface

V. I. Perekrestov and Yu. A. Kosminskaya

Sumy State University, Sumy, 40007 Ukraine

e-mail: kpe@sumdu.edu.ua

Received March 12, 2003; in final form, July 14, 2003

We have studied the mechanism of structure formation and the local chemical composition of the Ti–C system deposits formed at the interface between a carbon–argon–titanium plasma and a titanium substrate under conditions of the titanium content in the plasma decreasing with time. The deposits were studied in a scanning electron microscope (SEM) equipped with an energy-dispersive X-ray (EDX) analyzer. The spatially distributed growth of various structural modifications of carbon in the deposit is related to the mass transfer redistribution at the substrate surface and to the selective deposition of adatoms forming sufficiently stable chemical bonds to the growth surface in the thermal accommodation stage. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 81.15.Kk; 68.37.Hk; 68.55.Jk; 68.55.Nq

As is well known [1, 2], a necessary condition for the formation of α-C layers is that the process should be conducted at a high level of supersaturation. This provides for increased probability of the formation of sp^3-hybridized bonds, thus determining the growth of homogeneous diamondlike structures on the substrate surface. For this reason, processes taking place at low supersaturations did not attract much attention and remained almost unstudied.

However, our first experiments on the condensation of a weakly supersaturated vapor of a separate metal, or carbon, or a carbon–titanium mixture in a high-purity inert medium revealed a number of interesting features. The first to mention is the absence of growth coalescence and secondary nucleation, which accounts for the formation of a deposit in the form of relatively large (up to 500 nm in size) crystals weakly bound to each other [3]. Under the conditions of low supersaturation, an increase in the energy of deposited metal atoms reveals their tendency to form layers with coexisting amorphous and crystalline phases [3–5]. These peculiarities of structure formation can be explained by selective conditions favoring the deposition of adatoms forming sufficiently stable chemical bonds to the growth surface in the thermal accommodation stage, without taking into account the metastability of the growing phase. Therefore, while the structure formation in the deposit under high supersaturation conditions is determined to a significant extent by the collective processes in the vapor–deposit thermodynamic system, the regime of highly selective condensation at low supersaturations makes the formation of sufficiently strong bonds between individual adatoms and the growth surface the major factor.

Another important preliminary result was the synthesis of a new compound with the composition TiC_2 and, probably, of a diamondlike phase under low supersaturation conditions [6]. Of special interest is the metastability of the TiC_2 phase relative to the diamond phase, which was manifested in the experiment by the evident transformation of TiC_2 into diamond immediately under the action of a low-power electron beam in the electron diffraction setup. In view of the low intensity of the electron beam, such a conversion could take place only provided that the carbon sublattice of TiC_2 was susceptible to the transformation into diamond. We believe that the transition from TiC to the diamond phase under the conditions studied proceeds as follows. In the first stage, deposited carbon atoms with the energies of several tens of electronvolts penetrate into a subsurface layer of TiC and occupy interstitial sites, thus forming a TiC_2 phase. Under the action of an external perturbation (in our case, electron beam), Ti atoms also pass into interstitial sites and diffuse to the grain boundaries, while vacancies in the lattice sites exhibit collapse and the carbon sublattice contracts to form a diamond configuration.

Therefore, upon creating a Ti–C layer with the carbon concentration increasing in depth up to 66 at. % and providing the necessary technological conditions, we may expect the appearance of diamond nuclei. Should even local $TiC_2 \longrightarrow$ diamond transitions take place, it would be of interest to continue the growth of carbon under the aforementioned highly selective conditions, since a higher energy of the sp^3-hybridized

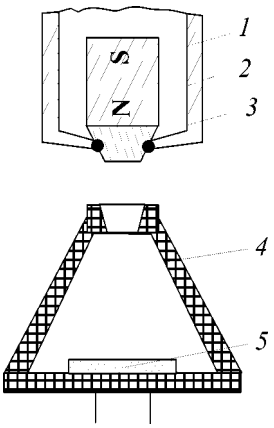


Fig. 1. Schematic diagram of the sputtering device: (*1*) water-cooled anode; (*2*) magnet; (*3*) magnetic circuit; (*4*) graphite hollow cathode; (*5*) titanium substrate.

bonds as compared to that of the π bonds in graphite can provide for the advancing layer-by-layer growth of diamond crystals. Evidently, such a process requires additional kinematic conditions to be provided favoring sp^3 hybridization, for example, by acting upon the growth surface with ion or inert gas atomic beams.

Although the possibility of such layer growth of diamond in the absence of hydrogen was absolutely rejected by some researchers [7], there are promising experimental results. In particular, investigation of the state of carbon atoms on the surface of a growth surface of α-C layer [8] showed that the surface atoms are linked primarily by the sp^3-hybridized bonds. From this we infer that the diamondlike state is formed due to the sp^3 hybridization, rather than as a result of implantation into a subsurface layer with sp^2-hybridized bonds [9]; at least, the aforementioned result does not exclude the possibility of a hydrogenless diamond growth process. However, this process must apparently take place at temperatures significantly higher than the temperature of formation of a diamondlike α-C deposit.

In this context, our study was aimed at elucidating the mechanism of structure formation in Ti–C deposits under the action of a carbon–argon–titanium plasma in which the metal content decreases with time. The process was conducted so as to provide that the carbon deposit would form under highly selective conditions mentioned above.

EXPERIMENTAL

An analysis of the conventional deposition techniques showed that no one of these meet the condition of high selectivity. For this reason, we have developed a principally new device based on a hollow cathode (Fig. 1). In this device, a Ti–C system deposit is formed immediately inside the hollow cathode. The process is conducted at a pressure of 20 Pa in the atmosphere of argon purified according to a special procedure [10]. In our experimental setup, the discharge current is amplified and additionally stabilized both due to the hollow cathode effect [11] and by an applied magnetic field (Fig. 1).

The internal surface of cylindrical hollow cathodes is usually subject to ion sputtering, which excludes the deposition process. However, if a hollow cathode has the shape of a truncated cone (Fig. 1), a significant sputtering of the material at the input hole and drawing of the sputtered particles into the cathode volume provide conditions for the deposition. At a discharge current of 1–3 A, a high-density plasma is generated inside the cathode, uniformly covering the entire cathode surface [11]. Under these conditions, the deposit structure formation is determined by the mass transfer in the region of interaction between the plasma and the growth surface.

The conditions of deposition inside the cathode are characterized by a number of peculiarities. At the initial moment, the internal cathode surface and the substrate bombarded by ions and atoms are heated to a temperature within T_c = 1050–1140°C and are subject to sputtering and sublimation. A part of the resulting vapor phase leaves the cathode space, while another part is converted into plasma and redeposited (atoms can be involved into multiply repeated sublimation–deposition cycles) [12]. Thus, adatoms weakly bound to the growth surface will be resputtered and revaporized with a greater probability. A combination of these conditions is just what ensures the required selectivity, accounting for the deposition of adatoms characterized by a sufficiently strong binding to the growth surface. The mass transfer processes at the substrate surface and in the adjacent plasma layer being closely related, the plasma and growing deposit should be considered as a unified system [12].

Since the area of the titanium substrate surface exposed to the plasma is significantly smaller than the graphite cathode surface (Fig. 1), the carbon component of the plasma layer inside the cathode will predominate over the titanium component. For this reason, the surface of titanium is gradually saturated with carbon and a TiC layer is formed. As the carbon concentration on the growth surface increases, the titanium plasma component drops, thus still more favoring the deposition of carbon and creating the required concentration profile. Evidently, by changing the system geometry, it is possible to control the carbon content in the plasma. The $TiC_2 \longrightarrow$ diamond transition is possible only at a certain stage of the concentration profile formation, provided that some perturbative factor (e.g., a local microdischarge) is acting upon the TiC_2 phase. Apparently, the probability of such events is rather small and, hence, the local $TiC_2 \longrightarrow$ diamond transitions are most probable.

In order to determine the deposition temperature, we placed the plates made of metals possessing various

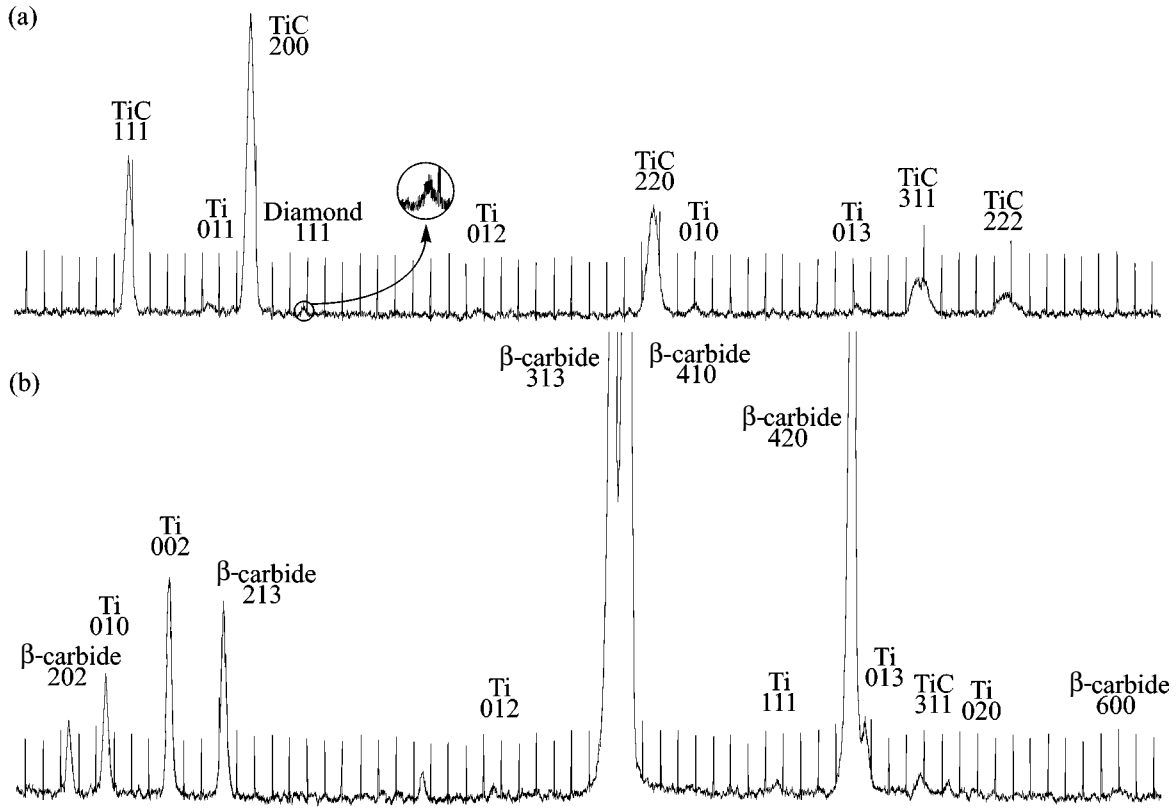


Fig. 2. X-ray diffractograms of deposits containing (a) dielectric inclusions and (b) β-carbyne.

melting temperatures inside the hollow cathode and gradually increased the discharge power. At the instant of melting of the metal surface, the intensity of its vaporization sharply increased, which was manifested by a change in the character of glow. At this moment, the discharge power was measured and the sputtering setup was switched off. A correlation between the discharge power and the temperature was considered as reliably determined only for the samples exhibiting partial fusion. The discharge power–deposition temperature relationship was determined using the known melting temperatures of the metal samples. During the Ti–C system deposition, a power supplied to the sputtering setup amounted to 430–470 W.

The microstructure and chemical composition of deposits were studied using a scanning electron microscope (SEM) equipped with an energy-dispersive X-ray (EDX) analyzer (X-ray microprobe) capable of probing a local sample region with a diameter of 1 μm and determining the composition to within a 4% accuracy. The chemical composition was calculated using data on the characteristic X-ray emission intensities of the samples and the Ti–C standards. In addition, the phase compositions of the deposit and the substrate were studied by X-ray diffraction.

RESULTS AND DISCUSSION

In the first series of experiments, we varied the geometry of the sputtering system and the discharge power in order to determine conditions under which the substrate was heated to ~1070°C and the deposition of carbon during an initial period of $t \sim 30$ min took place only in the form of implantation followed by thermal diffusion inward the titanium substrate. In this stage, according to the X-ray diffraction data (Fig. 2a), a TiC layer is formed. This layer had a maximum microhardness of ~1400 kgf/mm^2 and a thickness of about 1 μm. The substrate surface was partly sputtered, as confirmed by the formation of a characteristic microrelief with protruding crystal grains (Fig. 3a). As the deposition continued, nucleation took place only on some local areas and resulted in the growth of pyramids (Figs. 3b and 3c) containing 60 at. % C. Growing in the lateral directions, pyramids gradually covered the whole surface of the substrate (see Fig. 3d).

In the course of subsequent deposition, dielectric inclusions appear on the surface of the deposit (Fig. 3e). These inclusions consist of 99.8 at. % C and 0.2 at. % Ti. During the SEM measurements, these dielectric inclusions accumulate electrons and exhibit negative charging (thus being readily revealed by a dark background on the SEM micrographs). We believe that the dielectric properties and the chemical composition of

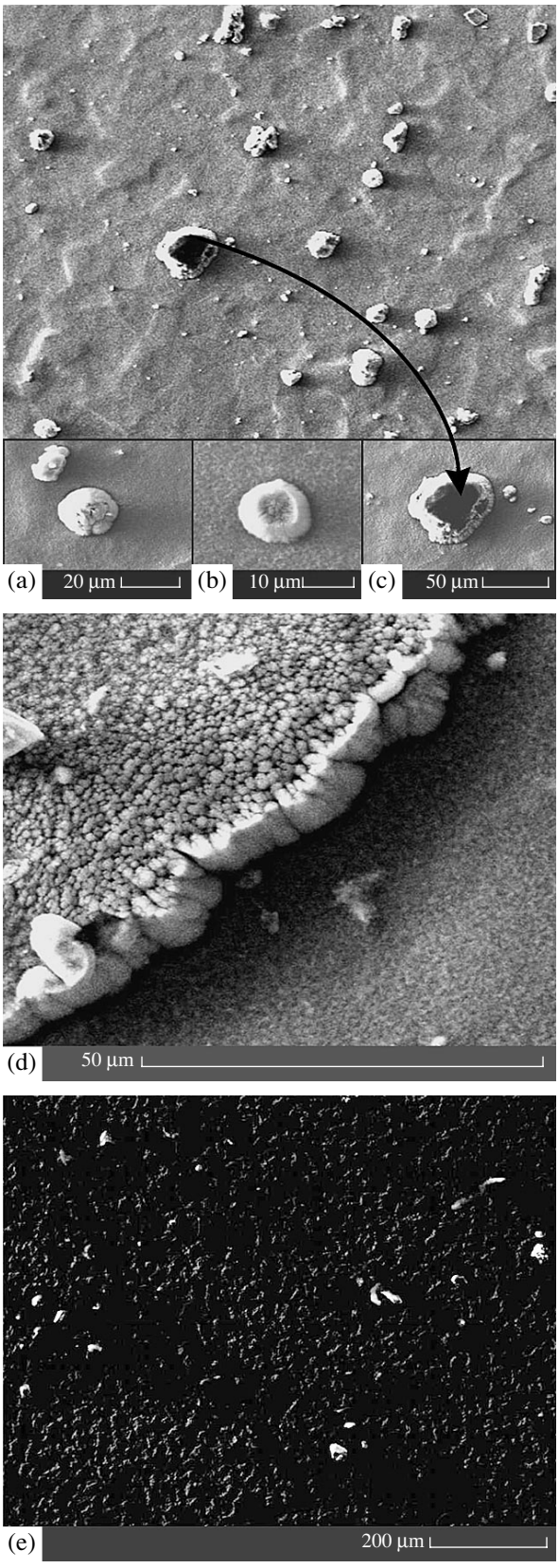


Fig. 3. SEM micrographs illustrating the nucleation and growth of deposit from a plasma with relatively low carbon content: t = 2 h (a–c), 4 h (d), and 8 h (e).

inclusions, together with the fact that weak (111) diamond reflections are observed on the X-ray diffractograms (Fig. 2a), provide sufficient ground to identify these inclusions as belonging to the diamond phase. However, attempts to obtain the deposit with dielectric inclusions of greater size in the first series of experiments were unsuccessful.

In the second series of experiments, we gradually decreased the size of titanium substrates. Under otherwise equal conditions, this increased the relative carbon component in the plasma. In this way, we have established conditions under which the growth rate of the dielectric inclusions exceeded that of the remaining deposit representing an amorphous carbon phase with a characteristic black color and the macroscopic properties typical of graphite. The dielectric inclusions appeared in the early stage of deposition, in the range of carbon concentrations within $C \sim 60$–80 at. % (Fig. 4a). When the carbon content on the growth surface of the graphitelike deposit approached 100 at. %, the dielectric inclusions in the form of polycrystals appeared only on separate carbide fibers protruding out of the growth surface (Figs. 4b and 4c). These observations confirm the important role of carbides in the possible process of diamond formation.

The dielectric inclusions formed from the nuclei appearing in the earlier stage of deposition frequently possess a single crystal structure (Fig. 4d). The formation of somewhat rounded edges is explained by sputtering, most probably, with neutrals. The content of carbon in these inclusions reaches almost 100 at. %, while the dark background around such crystals observed at small magnifications (Fig. 4e) is convincing evidence of their dielectric properties. Therefore, despite the absence of diamond reflections on the diffractograms (evidently, the amount of this substance is below the sensitivity threshold of X-ray diffraction), the crystal presented in Fig. 4d can be identified with a large probability as the diamond phase. The growth of relatively large dielectric crystals is apparently hindered by the accumulation of a critical positive charge, which leads to microdischarge, rapid heating, and detachment of the crystal from the substrate surface. It was experimentally established that only about 1/8 fraction of the nuclei grow to a size of several dozens of microns, and no more than a dozen of such crystals can be observed on a sample area of 1 cm^2.

It should be noted that, in addition to dielectric crystals, the deposit frequently contains the clusters of β-carbide (carbin). Being a semiconductor, β-carbyne more effectively reflects the electron beam during the SEM measurements, thus being revealed on the background of a graphitelike deposit. As can be seen from the SEM micrographs (Figs. 5a and 5b), β-carbyne grows more intensively on the areas of artificial roughness. This is probably explained by closure of the "mass-transfer cycles" on protruding parts of the substrate, where the electric field strength is increased. The

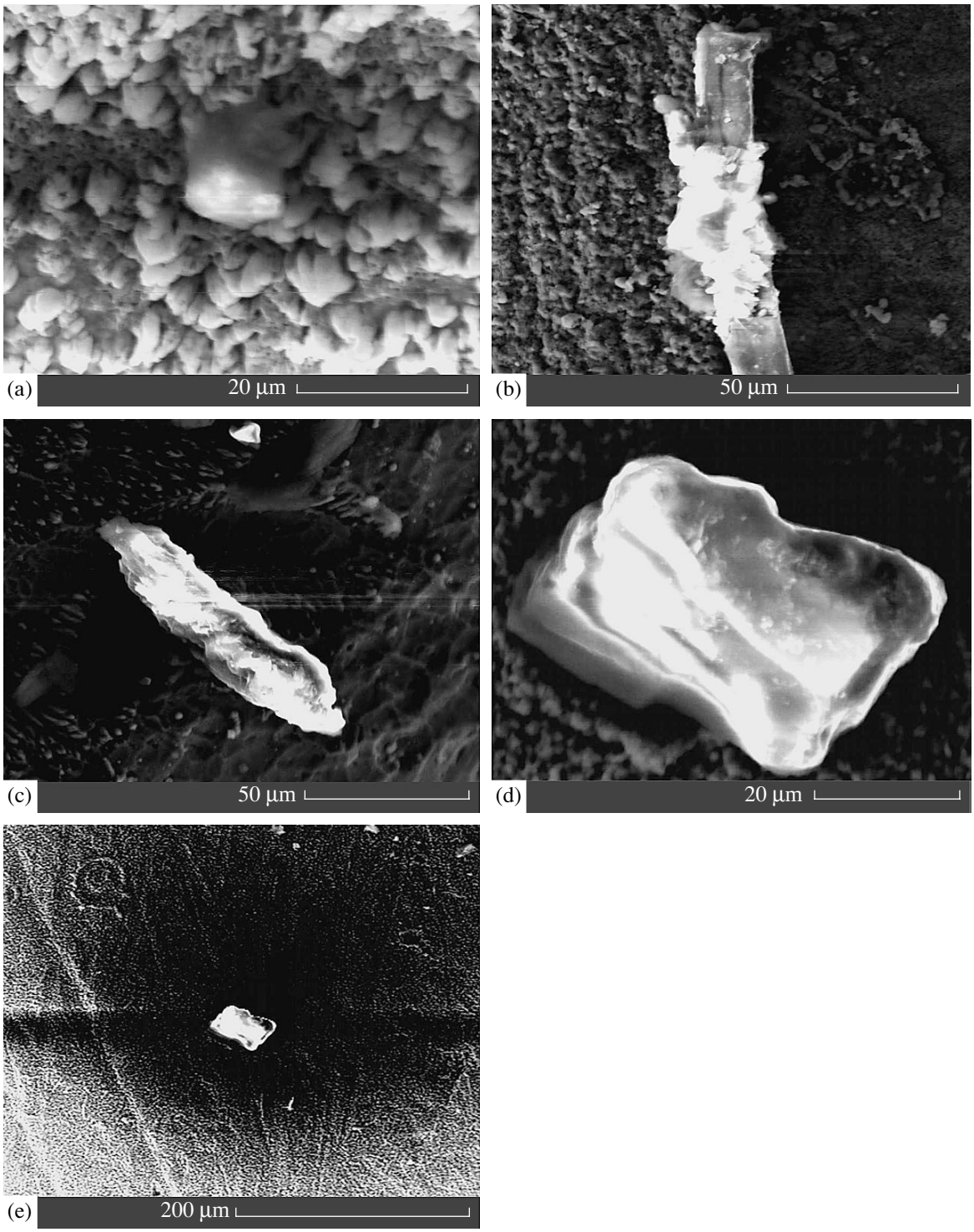


Fig. 4. SEM micrographs illustrating the formation of dielectric inclusions: t = 0.5 h (a), 1.5 h (b), 2.5 h (c), and 3.5 h (d, e).

presence of β-carbyne with a certain growth texture was confirmed by X-ray diffraction (Fig. 2b). Therefore, the transition from graphitelike structure to β-carbyne requires localization of the ion flows. Apparently, such a localization favors accelerated growth of β-carbyne inclusions, in the vicinity of which the electric field strength exhibits additional increase. Therefore, localization of the mass transfer has a self-sustained character. However, attaining a certain critical value of the local ion flow most probably results in strong heating of the β-carbyne inclusion, leading to its breakage or detachment from the growth surface.

Fig. 5. SEM micrographs illustrating the formation of β-carbide inclusions: (a) on polished substrate (t = 2.5 h); (b) on a substrate with artificial roughness (t = 2.5 h).

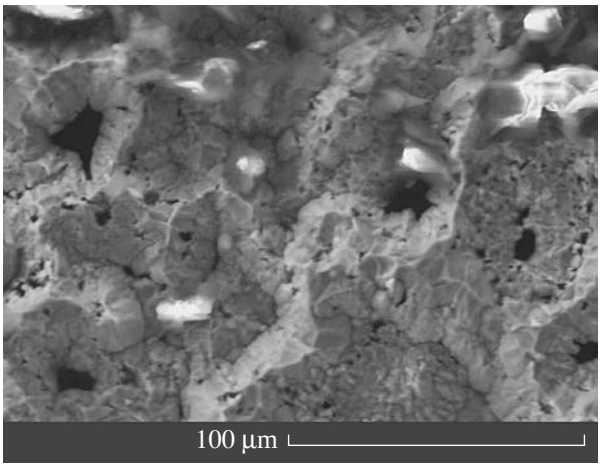


Fig. 6. A SEM micrograph showing the crater deposit structure.

A still greater proportion of carbon in the plasma was accompanied by the formation of a characteristic crater structure (Fig. 6). The results of the local chemical analysis showed that the carbon concentration amounts to several at. % on the bottom of a crater and approaches the maximum possible value on the top ridge. As can be seen from Fig. 6, some regions of the concentration profiles reflect transitions to a dielectric state. It is interesting to note that such transitions were observed in the regions where the carbon concentration is about 70 at. %. However, we failed to detect TiC_2 by X-ray diffraction measurements. This is probably explained by metastability of this phase and by its decomposition at the large deposition temperatures employed. On the other hand, the fact that the supposedly diamond phase is predominantly nucleated on the TiC fibers indicates that TiC_2 may participate in this process in an amount of several atomic monolayers. Note that, in the case under consideration, the preferential growth of the supposedly diamond phase is rather weakly pronounced. Further increase in the proportion of carbon in the plasma leads to rapid formation of a black carbon deposit free of diamondlike inclusions.

If the temperature falls out of the aforementioned T_c interval, the sputtering process is accompanied by the formation of morphologically homogeneous carbide or graphitelike deposits without dielectric carbon inclusions. Therefore, the spatially distributed selective growth of the desired carbon modifications takes place only within a rather narrow interval of the process parameters. It is experimentally established that the most important technological conditions determining the formation of inclusions supposedly identified as microdiamonds must (i) ensure the absolute stability and continuity of the selective deposition process and (ii) maintain the process temperature within the interval of $T_c \sim 1050$–1140°C. Reproducibility of the process depends on all geometric parameters of the sputtering system, including the substrate geometry, surface morphology, and arrangement in the working chamber. With all the aforementioned factors taken into account, the reproducibility of experiments reaches 60%.

It should be emphasized that the mechanisms responsible for the preferential growth of dielectric inclusions and β-carbyne are principally different. In the former case, the equipotential character of the growth surface is violated and the accumulation of positive charge by dielectric inclusions can both reduce the local electric field strength and change the local field direction. Therefore, there is ground to believe that the growth of possible diamond inclusions is related primarily to the weak flow of neutrals formed in a sufficient amount due to the resonance recharge [13] or due to the secondary electron–ion interactions.

In order to check for the role of titanium in the formation of dielectric carbon inclusions, we performed a series of control experiments with graphite and tantalum substrates. The results were negative. Using a mix-

ture of 10 at. % CH_4 + 90 at. % Ar as the working gas or increasing the partial pressures of chemically active gases (O_2 and N_2) violated stable operation of the sputtering device and impeded carbidization (by binding titanium in oxides, hydrides, and nitrides). These factors significantly alter the process of the deposit structure formation, whereby the cases of dielectric carbon crystal formation are seldom and the number of such inclusions is very small.

To summarize, we have found a new, extremely nonequilibrium mechanism of the transition of a substance from vapor into a condensed state. This mechanism accounts for the spatially distributed, selective growth of carbon deposits with various structures and morphologies at different rates.

REFERENCES

1. D. V. Fedoseev, *Diamond in Electron Technology* (Énergoatomizdat, Moscow, 1990).
2. A. S. Bakaĭ and V. E. Strel'nitskiĭ, *Structural and Physical Properties of Carbon Condensates Obtained by Sedimentation of Fast Particle Flows* (TsNIIatominform, Moscow, 1984).
3. V. I. Perekrestov, A. V. Koropov, and S. N. Kravchenko, Fiz. Tverd. Tela (St. Petersburg) **44**, 1131 (2002) [Phys. Solid State **44**, 1181 (2002)].
4. V. I. Perekrestov, S. N. Kravchenko, and A. V. Pavlov, Fiz. Met. Metalloved. **88**, 72 (1999).
5. V. I. Perekrestov, V. A. Kvorost, and S. N. Kravchenko, Vestn. Sum. Gos. Univ., Ser. Fiz. Mat. Mekh. **38**–**39**, 72 (2002).
6. V. I. Perekrestov and A. V. Pavlov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 17 (2001) [JETP Lett. **73**, 13 (2001)].
7. B. V. Spitsyn and A. E. Alekseenko, in *Proceedings of 5th International Symposium on Diamond Films and Films of Gognate Materials* (Khar'kov, 2002).
8. A. Dement'ev, D. Dolgikh, K. Maslakov, *et al.*, in *Proceedings of 5th International Symposium on Diamond Films and Films of Gognate Materials* (Khar'kov, 2002).
9. Y. Lifshits, Diamond Relat. Mater. **8**, 1659 (1999).
10. V. I. Perekrestov and S. N. Kravchenko, Prib. Tekh. Éksp., No. 2, 11 (2002).
11. B. I. Moskalev, *Discharge with Full Cathode* (Énergiya, Moscow, 1969).
12. A. G. Zhiglinskiĭ and V. V. Kuchinskiĭ, *Mass Transfer at an Interaction of Plasma with Surface* (Énergoizdat, Moscow, 1991).
13. Yu. M. Kagan and V. I. Perel', Dokl. Akad. Nauk SSSR **108**, 222 (1956) [Sov. Phys. Dokl. **1**, 289 (1956)].

Translated by P. Pozdeev

JETP Letters, Vol. 78, No. 4, 2003, pp. 230–235. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 265–270.
Original English Text Copyright © 2003 by Chtchelkatchev.

Superconducting Spin Filter¶

N. M. Chtchelkatchev

Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117940 Russia
Institute for High Pressure Physics, Russian Academy of Sciences, Troitsk, Moscow region, 142092 Russia
e-mail: nms@landau.ac.ru

Received July 14, 2003

Consider two normal leads coupled to a superconductor; the first lead is biased while the second one and the superconductor are grounded. In general, a finite current $I_2(V_1, 0)$ is induced in grounded lead 2; its magnitude depends on the competition between processes of Andreev and normal quasiparticle transmission from lead 1 to lead 2. It is known that, in the tunneling limit, when normal leads are weakly coupled to the superconductor, $I_2(V_1, 0) = 0$ if $|V_1| < \Delta$, and the system is in the clean limit. In other words, Andreev and normal tunneling processes compensate each other. We consider the general case: the voltages are below the gap, the system is either dirty or clean. It is shown that $I_2(V_1, 0) = 0$ for general configuration of the normal leads; if the first lead injects spin-polarized current then $I_2 = 0$, but spin current in lead 2 is finite. A XISIN structure, where X is a source of the spin-polarized current, could be applied as a filter separating spin current from charge current. We do an analytical progress calculating $I_1(V_1, V_2)$, $I_2(V_1, V_2)$. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 74.50.+r

Hybrid systems consisting of a superconductor (S) and two or more normal metal (N) or ferromagnetic (F) probes have recently started to attract great attention [1–5]. Among the most striking new results is the prediction that NSN (FSF) devices can play the role of an entangler, producing Einstein–Podosky–Rosen (EPR) pairs [4] having potential applications, for example, in quantum cryptography [6]. Not long ago, a rather unusual effect was described in a normal metal–tunnel barrier (I)–superconductor–tunnel barrier–normal metal (NISIN) junction (see, e.g., Fig. 1b) [2, 3]. It was shown that when N_1 is biased and N_2 and S are grounded, there is no current injection from N_1 to N_2 at subgap biases. The main assumptions were that 1) the superconductor is clean and 2) a large number of conducting channels are involved in electron tunneling through NS interfaces [2, 3]. In other words, the subgap cross conductance $G_{12} \equiv \partial_{V_1} I_2(V_1, 0)|_{|V_1| < \Delta} = 0$, where the current I_1 flows in N_1 and V_1 is the bias between N_1 and S and V_2, that between N_2 and S. The suppression of G_{12} was attributed to the compensation of the contributions to the current from Andreev and normal quasiparticle tunneling processes between N_1 and N_2 [2]. It was also noted that $G_{12} \neq 0$ in FISIF junctions: G_{12} decays exponentially as $\exp(-r/\xi)$ with the characteristic distance r between the normal terminals (see, e.g., Fig. 1b), where ξ is the superconductor coherence length; at small r/ξ, G_{12} also decays rather quickly (at atomic scales): as $1/(k_F r)^2$ (k_F in the superconductor) [2]. Thus, with clean superconductors, measurement of G_{12} may become difficult.

In this letter, we first of all generalize the results of [2, 3] and get rid of assumption 1 (i.e., S is not restricted to be clean). We show that when the superconductor is dirty (the mean free path is smaller than ξ) Andreev and normal transmission rates [as well as G_{12} in FISIF junctions] slowly decay with the characteristic distance r between the normal (ferromagnetic) terminals (at $r < \xi$), in contrast to the clean regime (see [2, 3]). For example, in FISIF with a superconducting layer thinner than ξ, see Fig. 1b, $G_{12} \sim \ln(r/\xi)$; when the superconductor is bulk, then $G_{12} \sim \xi/r$ ($r > \lambda_F$ is supposed). Measurements of effects related to electron tunneling through a superconductor (e.g., G_{12}) in the dirty superconductor case can be more easily realized experimentally than in the clean case, because then r is restricted not to atomic scales but by $\xi \gg \lambda_F$. We

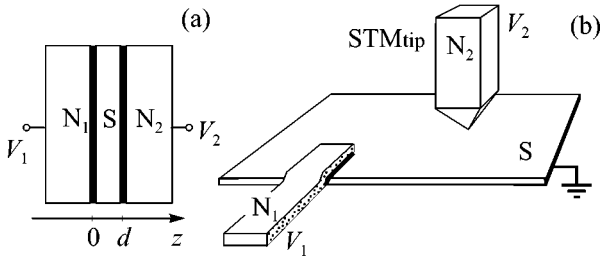


Fig. 1. The outline of the setup. $N_{1,2}$ are normal metals or ferromagnets.

¶ This article was submitted by the author in English.

show that contributions to the current from Andreev and normal quasiparticle tunneling processes always compensate each other in NISIN junctions (so, e.g., $I_2(V_1, V_2 = 0) = 0$ for $|V_1| < \Delta$ in the first nonvanishing order over the transparencies of the layers I) for any amount of disorder in the S-layer. If one prepares a NISIN junction with layers I having large transparency, then normal tunneling starts dominating Andreev tunneling (and $I_2(V_1, V_2 = 0) \neq 0$). We also considered a FISIN junction, in particular with $V_F \neq 0$ and $V_N = 0$. Then the ferromagnet F plays the role of the spin-polarized current injector. In this case, $I_2(V_1, V_2 = 0) = 0$ also, but spin current in N is finite: the charge component of the current converts into the supercurrent, spin accumulates in N. So the XISIN structure, where X is a source of the spin-polarized current, could be applied in spintronics [7] as a filter separating spin current from charge current. We find Andreev T_{he} and normal transmission probabilities T_{ee} of a NISIN sandwich for subgap energies $|E| < \Delta$ and different angles θ between incident quasiparticle trajectory and the normal to NS interface. It is shown that the probabilities have resonances where $T_{he} \sim T_{ee}$; the averages of T_{he} and T_{ee} over incident channels (over θ) are equal. This is the reason why $I_2(V_1, V_2 = 0)$ is suppressed and the spin current $I_2^{(s)}(V_1, V_2 = 0)$ is finite.

We start investigation of NISIN structures from the sandwich sketched in Fig. 1a: barriers at NS boundaries provide spectacular reflection; electrons in N and S move ballistically; the number of channels at both NS boundaries is much larger than unity. The transmission probabilities $T_{he}(E, \theta)$ and $T_{ee}(E, \theta)$ (see Fig. 2) describe Andreev and normal tunneling of an electron incident on the NS boundary with the angle θ and the energy E, respectively, into a hole and an electron in lead 2. Following the Landauer–Büttiker approach [8–10],

$$I_2(V_1, V_2 = 0) = \frac{e}{\hbar}\int dE \sum_{\text{channels}} (T_{ee} - T_{he})(f^{(1)} - f^{(2)}), \quad (1)$$

where the sum is taken over channels (spin degrees of freedom are included into channel definition); $f^{(1,2)}$ are distribution functions in the leads 1, 2; e.g., $f^{(2)} = n_F(E) = 1/[1 + \exp(\beta E)]$, $f^{(1)}$ is not necessary a Fermi function. We calculate the transmission and reflection probabilities using Boguliubov–de Gennes (BdG) equations. The layers I are approximated by δ barriers. Quasiparticle motion parallel and perpendicular to the NS interfaces can be decoupled [11, 12]. Matching appropriate wave functions in the normal region and the superconductor, we get an 8×8 linear system of equations for the transmission amplitudes. Analytical progress can be made. It follows that, if there is no barrier at NS boundaries (except Δ), $T_{he}/T_{ee} \lesssim (\Delta/E_F)^2$ for any thickness d of the superconducting layer. This result is intuitively quite clear, because $\Delta \ll E_F$ can hardly reverse the direction of the quasiparticle momentum, being about k_F [11, 13]. However if there are barriers at NS boundaries in addition to Δ (e.g., insulating layers I), then the situation changes: at certain θ the transmission probabilities have resonances where $T_{he} \sim T_{ee}$. When the transmission probabilities of the layers I, $T_{NS}^{(1,2)} \ll 1$, the areas under the resonance peaks of $T_{he}(\theta)$ and $T_{ee}(\theta)$ are nearly the same and

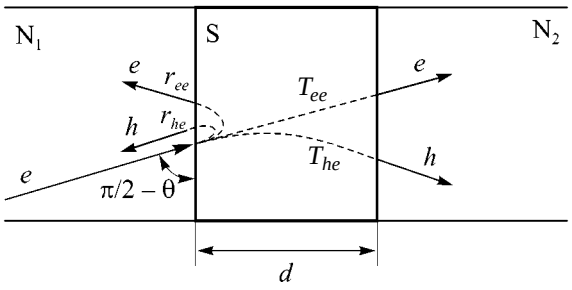


Fig. 2. Electron scattering from a NSN junction.

$$\langle T_{he}\rangle \approx \langle T_{ee}\rangle, \quad (2)$$

where $\langle \ldots \rangle = \sum_{\text{channels}} (\ldots)/N_{\text{channel}} \approx \int_0^1 (\ldots)\, d\cos^2\theta$. Equation (2) is exact in the first nonvanishing order over T_{NS}. The resonances appear at $k_F d\cos(\theta_n) = \pi n$, $n = 1, 2 \ldots$, give the leading contributions to $\langle T_{he}\rangle$, $\langle T_{ee}\rangle$ and are responsible for Eq. (2). The resonance width $\Gamma \sim \min\{1, T_{NS}, d/\xi\}$. Typical dependences of $T_{he}(\theta)$ and $T_{ee}(\theta)$ from θ and T_{NS} are illustrated in Fig. 3.

In fact, θ is a discrete variable; its particular value is determined by the channel of the incident particle. Equation (2) is applicable when 1) $T_{NS}(\theta)$ slightly change when θ changes from one channel to an adjacent one and 2) the change of θ from one channel to another is smaller than the resonance width. Condition 1 is fulfilled typically when $T_{NS}(\theta) \ll 1$; condition 2 requires $\lambda_F/\sqrt{A} \ll \min\{1, T_{NS}(0), d/\xi\}$, where A is the junction surface area.

It follows from Eqs. (1), (2) that subgap charge injection from lead 1 into lead 2 in the weak coupling regime ($T_{NS} \ll 1$) is suppressed: $I_2(V_1, V_2 = 0) = 0$, because charge currents of transmitted hole and electron quasiparticles compensate each other in lead 2; all the electron current converts into Cooper pair supercurrent in S. However, if spin-polarized current is injected from lead 1, finite spin current appears in lead 2; transmitted electron and hole quasiparticles contribute the spin current. An XISIN structure with $T_{NS}^{(1)} \ll T_{NS}^{(2)} \ll 1$ (this condition allows one to neglect the contribution to the charge current going in S from Andreev reflection at

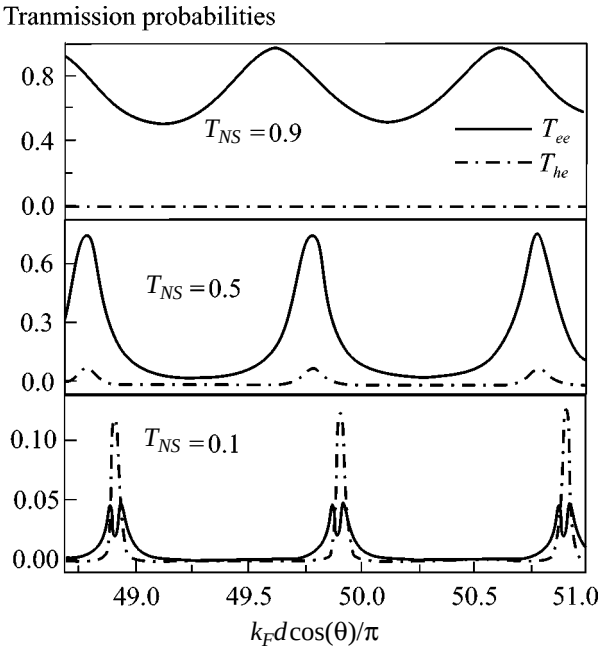


Fig. 3. Resonances of Andreev and normal transmission probabilities $T_{ee}(\theta)$, $T_{he}(\theta)$ of a NISIN junction Fig. 1a at different transparencies $T_{NS}(\theta)$ of the layers I. Parameters: $d/\xi = 0.1$, $E_F/\Delta = 1000$, the energy $E = 0$. Resonances correspond to $k_F d\cos(\theta)/\pi = n$, $n = 1, 2, \ldots$. If $0 < E < \Delta$ then the shape of the graphs slightly change: the resonance peaks become slightly asymmetric [the same applies to the case when T_{NS} are different but small]. It can be checked (even analytically) that the areas under the corresponding resonance peaks of T_{ee} and T_{he} become equal at small T_{NS}.

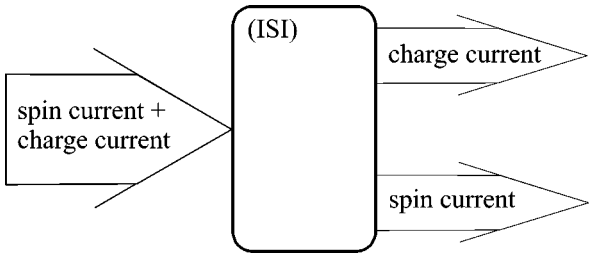


Fig. 4. Isolator–superconductor–isolator–normal metal structure can help to spatially separate spin and charge components of the current.

N_1S surface), where X is the current "injector," can play the role of the filter of spin and charge currents; see Fig. 4. The equation for the spin current follows from Eq. (1):

$$I_2^{(s)}(V_1, V_2 = 0) = \frac{e}{2\hbar}\int dE \sum_{\text{channels}} \sigma_1 (T_{ee} + T_{he})(f^{(1)} - f^{(2)}), \quad (3)$$

where $\sigma_1 = \pm 1$ labels spin degrees of freedom in X. The general feature of transmission probabilities T and the current is their exponential suppression with d/ξ when $d \gg \xi$ (ξ is the superconductor coherence length).

We show below that all the results discussed above remain true in the general NISIN structure with a shape more complicated than in Fig. 1a (e.g., as in Fig. 1b), no matter dirty or clean.

In general, a system of weakly coupled normal (ferromagnetic) and superconducting layers can be described by the Hamiltonian $\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_S + \hat{H}_T$, where $\hat{H}_{1,2}$ refer to the electrodes N_1 and N_2, and $\hat{H}_S$ refers to the superconductor. The tunnel Hamiltonian $\hat{H}_T$, which we consider as a perturbation, is given by two terms, $\hat{H}_T = \hat{H}_T^{(1)} + \hat{H}_T^{(2)}$, corresponding to one-particle tunneling through each tunnel junction:

$$\hat{H}_T^{(i)} = \sum_{k,p} \{\hat{a}_k^{(i)\dagger} t_{kp}^{(i)} \hat{b}_p + \text{h.c.}\}, \quad (4)$$

where the indices $i = 1, 2$ refer to normal (ferro) electrodes and $t_{kp}^{(i)}$ is the matrix element for tunneling from the state $k = (\mathbf{k}, \sigma)$ in normal lead N_i to the state $p = (\mathbf{p}, \sigma')$ in the superconductor. The operators $\hat{a}_k^{(i)}$ and $\hat{b}_p$ correspond to quasiparticles in the leads and in the superconductor, respectively.

The current can be expressed through the quasiparticle scattering probabilities within the Landauer–Büttiker approach. It is possible to calculate the scattering probabilities within the tight-binding model (4), but it is more convenient to describe the current in the language of electrons only: Andreev transmission probability T_{he} (1) is closely related to the crossed Andreev (CA) tunneling rate $\Gamma_{CA}^{S\leftarrow}(V_1, V_2)$, which shows how many electron pairs tunnel per second from leads 1 and 2 into the condensate of the superconductor (each lead gives one electron into a pair) and vice versa; see Fig. 5b and [2]. Elastic cotunneling e rate $\Gamma_{EC}^{2\leftarrow 1}$ corresponds to T_{ee}. The direct Andreev (DA) tunneling rates $\Gamma_{DA}^{S\to 1(2)}$ and $\Gamma_{DA}^{S\leftarrow 1(2)}$ describe Andreev reflection in the leads 1 and 2 (see, e.g., Fig. 5a). The current in lead 2 consists of two contributions: one, $I_2^{(i)}$, comes from the electron injection from lead 1 due to crossed Andreev and cotunneling processes, the other, $I_2^{(I)}$, from the direct electron tunneling between the lead and the

superconductor. The same applies for lead 1. $I_2(V_1, V_2) = I_2^{(I)}(V_1, V_2) + I_2^{(D)}(V_2)$, where

$$I_2^{(I)}(V_1, V_2) = \Gamma_{EC} - \Gamma_{CA}, \tag{5a}$$

$$\Gamma^{EC} = \Gamma_{EC}^{1\leftarrow 2} + \Gamma_{EC}^{2\leftarrow 1}, \tag{5b}$$

$$\Gamma^{CA} = \Gamma_{CA}^{S\to 1,2} - \Gamma_{CA}^{S\leftarrow 1,2}, \tag{5c}$$

$$I_2^{(D)}(V_2) = \Gamma_{DA}^{S\leftarrow 2} - \Gamma_{DA}^{S\to 2}. \tag{5d}$$

Using the Fermi golden rule, the rates can be found in the second order in the tunneling amplitude $t_{k,p}$. Following the approach described in [2, 14, 15], we finally obtain

$$\Gamma_{CA}^{S\leftarrow 12}(V_1, V_2) = 4\pi^3 \int d\xi \sum_\sigma n_\sigma^{(1)}(\xi - V_1) \times n_{-\sigma}^{(2)}(-\xi - V_2) \frac{\Delta^2}{[\Delta^2 - \xi^2]} \tilde{\Xi}_\sigma^{CA}(2\sqrt{\Delta^2 - \xi^2}), \tag{6}$$

where $n^{(i)}$ is the distribution function in the lead $i = 1, 2$. Hereafter, we take $\hbar = 1$, $e = 1$ [we do not assume $n^{(i)}$ to be only equilibrium Fermi function]. The rate $\Gamma_{CA}^{S\to 12}$ can be obtained from the expression for $\Gamma_{CA}^{S\leftarrow 12}$ by substitution of $(1 - n)$ for n.

The kernel $\tilde{\Xi}_\sigma^{CA}(s) \equiv \int_0^\infty dt \Xi_\sigma^{CA}(t) e^{-st}$ is the Laplace transform of $\Xi_\sigma^{CA}(t)$. It can be expressed through the classical probability, $P(X_1, \hat{p}_1; X_2, \hat{p}_2, t)$ meaning that an electron with the momentum directed along $\hat{p}_1$ initially located at the point X_1 near the NS boundary arrives at time t at some point X_2 near the NS boundary with the momentum directed along $\hat{p}_2$ spreading in the superconducting region as

$$\Xi_\sigma^{CA}(t) = \frac{1}{8\pi^3 e^4 \nu_S} \times \int d\hat{p}_{1,2} \int dX_{1,2} P(X_1, \hat{p}_1; X_2, \hat{p}_2, t) \times \left\{ G^{(1)}(X_1, \hat{p}_1, \sigma) G^{(2)}(X_2, \hat{p}_2, \sigma) \sin^2\left(\frac{\theta(X_1, X_2)}{2}\right) + G^{(1)}(X_1, \hat{p}_1, \sigma) G^{(2)}(X_2, \hat{p}_2, -\sigma) \cos^2\left(\frac{\theta(X_1, X_2)}{2}\right) \right\}.$$

Here, the spatial integration is performed over the N_1S and N_2S surfaces. We choose the spin quantization axis

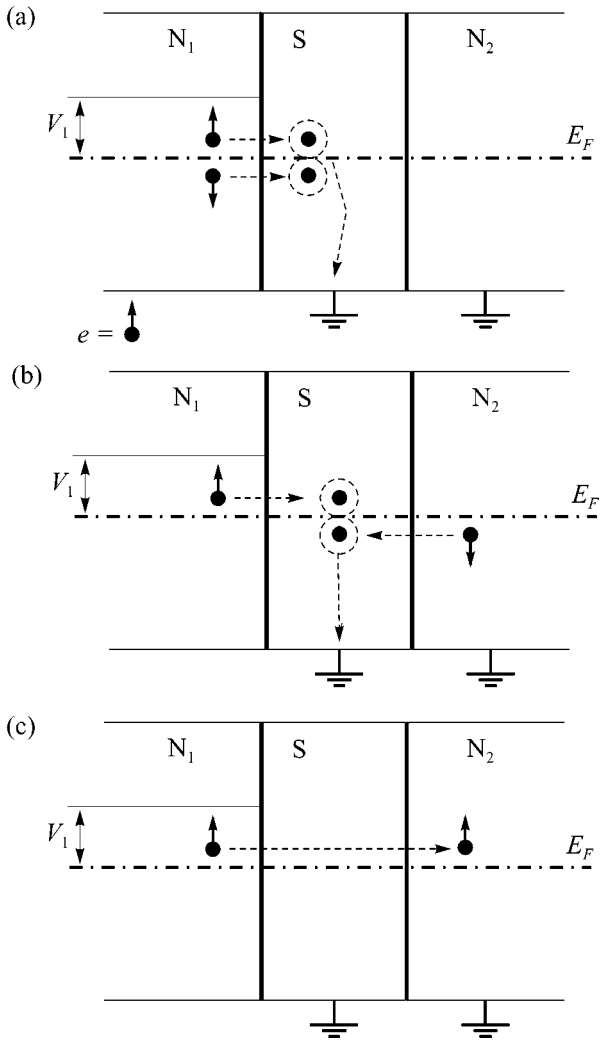


Fig. 5. (a) Direct Andreev tunneling (Andreev reflection), (b) Crossed Andreev tunneling (Andreev transmission), and (c) Elastic cotunneling (normal transmission).

in the direction of the local magnetization in the terminal $N_{1(2)}$. The quasiclassical probabilities $G^{(i)}(X, \hat{p}, \sigma)$, $i = 1, 2$ for the electron with spin polarization σ tunneling from the terminal N_i to the superconductor are normalized such that the junction normal conductance per unit area $g_\sigma^{(i)}(X)$ and the total normal conductance $G_N^{(i)}$ are determined as [14, 16]

$$g_\sigma^{(i)}(X) = \int d\hat{p} G^{(i)}(X, \hat{p}, \sigma), \quad G_N^{(i)} = \int dX \sum_\sigma g_\sigma^{(i)}(X).$$

Then the normal conductance per unit area, discussed above, is defined as $g_N^{(i)} = G_N^{(i)}/\mathcal{A}$, where $\mathcal{A}$ is the surface area of the junction. Symbol $\theta(X_1, X_2)$ is the angle between the magnetizations of the terminals N_1 and N_2 at points X_1 and X_2 near the junction surface. If elec-

trons in N_1 and N_2 are not polarized, then $\theta = 0$. In a similar way:

$$\Gamma_{EC}^{1\to 2} = 4\pi^3 \int d\xi \sum_\sigma n_\sigma^{(1)}(\xi - V_1)(1 - n_\sigma^{(2)}(\xi - V_2)) \times \frac{\Delta^2}{[\Delta^2 - \xi^2]} \tilde{\Xi}_\sigma^{EC}(2\sqrt{\Delta^2 - \xi^2}),$$

where

$$\Xi_\sigma^{EC}(t) = \frac{1}{8\pi^3 e^4 \nu_S} \times \int d\hat{p}_{1,2} \int dX_{1,2} P(X_1, \hat{p}_1; X_2, \hat{p}_2, t) \times \left\{ G^{(1)}(X_1, \hat{p}_1, \sigma) G^{(2)}(X_2, \hat{p}_2, -\sigma) \sin^2\left(\frac{\theta(X_1, X_2)}{2}\right) + G^{(1)}(X_1, \hat{p}_1, \sigma) G^{(2)}(X_2, \hat{p}_2, \sigma) \cos^2\left(\frac{\theta(X_1, X_2)}{2}\right) \right\}. \quad (7)$$

The rate $\Gamma_{EC}^{1\leftarrow 2}$ can be obtained from the expression for $\Gamma_{EC}^{1\to 2}$ by substitution of $(1 - n)$ for n. DA rates are written in [14]. Equations (5a)–(7) derived here allow one to describe the transport properties of many types of junctions.

Consider a FISIN junction with biased ferromagnet with the respect to the superconductor; the normal metal N has same voltage as S. The ferromagnet plays the role of a current "injector"; electrons coming from F are distributed with some distribution function $n^{(1)}$. Electrons in the deep of the terminal N are Fermi-distributed. It follows from Eqs. (5a)–(7) that contributions to the current from EC and CA processes compensate each other for subgap voltages, so $I_N(V_F \neq 0, 0) = 0$. However, spin current is finite:

$$I_N^{(\mathrm{spin})}(V_F, 0) = 4\pi^3 \int d\xi \sum_\sigma \sigma [n_\sigma^{(1)}(\xi - V_F) - n^{(2)}(\xi)] \times \frac{\Delta^2}{[\Delta^2 - \xi^2]} \frac{1}{8\pi^3 e^4 \nu_S} \int d\hat{p}_{1,2} \int dX_{1,2} \times P(X_1, \hat{p}_1; X_2, \hat{p}_2, t) G^{(1)}(X_1, \hat{p}_1, \sigma) G^{(2)}(X_2, \hat{p}_2). \quad (8)$$

Finally, we consider a FISIF junction. It was shown in [2] that in this junction $I_2(V_1, 0) \neq 0$ and $I_2(V_1, 0)$ changes its sign when the ferromagnetic terminals change their orientation from parallel to antiparallel. Naively it can be supposed that, in a FISIN junction where F is a jafrrent injector, S, N are grounded spin accumulation at the interface of the normal metal would lead to spin splitting of the density of states in N and a charge current. However, this is not so; these corrections are of a higher order over tunneling amplitudes than the processes in Fig. 5 and can be neglected, because we assume that tunneling amplitudes are small.

It was also noted in [2, 3] that the cross conductance $G_{12} \equiv \delta_{V_1} I_2(V_1, 0)|_{V_1 = 0}$ is suppressed in a FISIF structure as $1/(k_F r)^2$ when the characteristic distance between the ferromagnets $r < \xi$. In the dirty regime there is no conductance suppression at atomic scales. Consider, for instance, the layout sketched in Fig. 1b; the width d of the superconducting film is supposed to be smaller than ξ. According to Eqs. (5a)–(7), the cross-conductance dependence from the distance r is determined by the Laplace transform $\tilde{P}(s = 2\sqrt{\Delta})$ of the probability $P(r, t) = \exp(-r^2/4D|t|)/4\pi d|t|$, where D is the diffusion constant in the superconductor, $d < \xi$. When $\lambda_F \ll r < \xi$, $G_{12} \sim \tilde{P} \sim \ln(r/\xi)$ and if $r \gg \xi$, $G_{12} \sim \tilde{P} \sim \exp(-r/\xi)$. When the superconductor is bulk ($d > \xi$), similarly we find $G_{12} \sim \xi/r$, $\lambda_F \ll r < \xi$. All considerations above apply also for CA and EC rates. Thus, it is practically more convenient to measure finite effects related to electron subgap tunneling through a superconductor when it is dirty rather than clean. In the dirty case, the terminals are not restricted to being as close as λ_F, as in the clean case, but closer than $\xi \gg \lambda_F$.

We are grateful to M. Mar'enko, Yu.V. Nazarov, V.V. Ryazanov, M.V. Feigelman, A.S. Iosselevich, and Ya.V. Fominov for stimulating discussions. M. Mar'enko directed my attention to suppression of the zero bias cross conductance in dirty NISIN junctions (of certain type) and the long-range decay of the EC and CA rates with the characteristic distance between the normal terminals, which appeared important for reviewing in the general case for spin and charge transport in superconducting junctions with weak coupling to the normal (ferromagnetic) terminals. When the paper was nearly completed, I got the information that spin injection in a normal layer of a FISIN junction was mentioned in one sentence of [17]. We thank D. Feinberg for criticism and for pointing our attention to [17]. We thank the Russian Foundation for Basic Research (project no. 03-02-16677), the Swiss NSF, and the Russian Ministry of Science.

REFERENCES

1. G. Deutscher and D. Feinberg, Appl. Phys. Lett. **76**, 487 (2000).
2. G. Falci, D. Feinberg, and F. W. J. Hekking, Europhys. Lett. **54**, 225 (2001).
3. R. Melin and D. Feinberg, Eur. Phys. J. B **26**, 101 (2002).
4. G. B. Lesovik, T. Martin, and G. Blatter, Eur. Phys. J. B **24**, 287 (2001); P. Recher, E. V. Sukhorukov, and D. Loss, Phys. Rev. B **63**, 165314 (2001); N. M. Chtchelkatchev, G. Blatter, G. B. Lesovik, *et al.*, Phys. Rev. B

66, 161320 (2002); M. S. Choi, C. Bruder, and D. Loss, Phys. Rev. B **62**, 13569 (2000).

5. F. J. Jedema, B. J. van Wees, B. H. Hoving, *et al.*, Phys. Rev. B **60**, 16549 (1999).
6. D. Bouwmeester, A. Ekert, and A. Zeilinger, *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computations* (Springer, Berlin, 2000).
7. F. Giazotto, F. Taddei, R. Fazio, *et al.*, Appl. Phys. Lett. **82**, 2449 (2003).
8. C. J. Lambert, J. Phys. C **3**, 6579 (1991); Y. Takane and H. Ebisawa, J. Phys. Soc. Jpn. **61**, 1685 (1992).
9. M. P. Anantram and S. Datta, Phys. Rev. B **53**, 16390 (1996).
10. Ya. Blanter and M. Büttiker, Phys. Rep. **336**, 1 (2000).
11. A. F. Andreev, Sov. Phys. JETP **19**, 1228 (1964).
12. M. Božović and Z. Radović, Phys. Rev. B **66**, 134524 (2002).
13. Y. Imry, *Introduction to Mesoscopic Physics* (Oxford Univ. Press, Oxford, 1997).
14. F. W. Hekking and Yu. V. Nazarov, Phys. Rev. B **49**, 6847 (1994).
15. N. M. Chtchelkatchev and I. Burmistrov, cond-mat/0303014.
16. D. V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **65**, 2446 (1990).
17. D. Feinberg, G. Deutscher, G. Falci, *et al.*, in *Proceedings of Rencontres de Moriond 2001*, Ed. by T. Martin, G. Montambaux, and J. Tran Thanh Van (EDP Sciences, 2001), p. 535.

JETP Letters, Vol. 78, No. 4, 2003, pp. 236–240. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 271–275.
Original Russian Text Copyright © 2003 by Iskhakov, Komogortsev, Balaev, Okotrub, Kudashov, Kuznetsov, Butenko.

Fe Nanowires in Carbon Nanotubes as an Example of a One-Dimensional System of Exchange-Coupled Ferromagnetic Nanoparticles

R. S. Iskhakov[1], S. V. Komogortsev[1,*], A. D. Balaev[1], A. V. Okotrub[2], A. G. Kudashov[2], V. L. Kuznetsov[3], and Yu. V. Butenko[3]

[1]*Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia*

[2]*Nikolaev Institute of Inorganic Chemistry, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 3, Novosibirsk, 630090 Russia*

[3]*Boreskov Institute of Catalysis, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 5, Novosibirsk, 630090 Russia*

**e-mail: komogor@iph.krasn.ru*

Received July 21, 2003

The cooperative phenomena revealed in the field and temperature dependences of the magnetization in a system of iron nanoparticles in carbon nanotubes were studied experimentally. The character of the temperature dependences of the magnetization indicates that the ferromagnetic Fe particles in carbon nanotubes are exchange-coupled. In the region where the magnetization approaches saturation, the magnetization curves reveal the power dependence $\Delta M \sim H^{-3/2}$ typical for a one-dimensional system of exchange-coupled ferromagnetic nanoparticles. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 75.75.+a; 75.60.Ej; 75.50.Bb

At present, many researchers are interested in the magnetic properties of ferromagnetic nanowires from the viewpoints of both possible applications and fundamental research [1–22]. Such effects as giant (70% [1], 20% [15]) and colossal (up to 10000% [2]) magnetoresistance and giant (as compared to bulk ferromagnets) coercive force (1–5 kOe) [1, 3–7, 17–19] have already been observed in magnetic nanowires.

Two methods are currently used for producing nanowire ensembles based on Fe, Co, Ni, and their alloys. The first method is the electrochemical deposition of a metal into cylindrical pores of such porous matrices as aluminum oxide [3–11], silicon [12], and polycarbonate membranes [13–16]. The second method is based on the arrangement of metal nanoparticles inside carbon nanotubes produced by the decomposition of compounds containing a magnetic 3D metal. This method uses chemical vapor deposition (CVD) [17–21] and carbon electric arc decomposition [2, 22].

Unusual magnetic properties of ferromagnetic nanoparticles are caused to a considerable extent by the cooperative effects in the magnetic system of strongly coupled nanoparticles. The cooperative effects in a system of exchange-coupled nanoparticles depend mainly on the strength of exchange coupling and the spatial arrangement of nanoparticles. In nanomaterials with a random distribution of the anisotropy axes of small particles, the main structural characteristics are the size of the nanoparticles and the dimensionality of their arrangement. As shown in [23], the average magnetic anisotropy (and, therefore, the coercive force H_c and the magnetic permeability) of such ferromagnetic nanomaterials is described by the power function $\langle K \rangle \sim K(R_c/\delta)^{2d/(4-d)}$, where K is the energy of the local magnetic anisotropy, R_c is the correlation radius of the random magnetic anisotropy (in nanocrystals R_c is usually taken to be half the grain size), $\delta = (A/K)^{1/2}$ is the exchange correlation length, A is the exchange interaction constant, and d is the dimensionality of the grain arrangement. It was shown in [23, 24] that the approach of the magnetization to saturation in ferromagnetic nanomaterials with $R_c < \delta$ (such materials can be considered as systems of exchange-coupled nanoparticles) is determined by the dimensionality d of the arrangement of nanoparticles. According to [23, 24], the reversible part of the magnetization curve of such ferromagnetic nanomaterials in fields $H < H_{ex} = 2A/MR_c^2$ (so-called exchange field) is described by the following formula:

$$\frac{M(H) - M_s}{M_s} = -\left(\frac{D^{1/2}H_a}{H_{ex}}\right)^2\left(\frac{H_{ex}}{H}\right)^{(4-d)/2}, \qquad (1)$$

where $H_a = 2K/M_s$ is the local anisotropy field. Therefore, studies of the magnetization curves in the region where the magnetization approaches saturation combined with the study of the low-temperature dependence of the magnetization make it possible to obtain information about the parameters A, K, R_c, and d. In principle, these parameters could also be determined by studying the dependence of H_c on R_c, but this method is very laborious and methodologically difficult.

The goal of this work was to study experimentally the cooperative effects that are revealed in the field and temperature dependences of the magnetization of a system of iron nanoparticles in carbon nanotubes.

Experiment. Samples of two types were studied: samples *s*1 synthesized by the electric arc decomposition of $Fe(CO)_5$ [22] and samples *s*2 synthesized by the thermolysis of a mixture of C_{60} fullerene with ferrocene [21]. The samples were obtained in the form of powder consisting of carbon nanotubes filled with iron. Microphotographs of the nanotubes obtained by transmission electron microscopy (TEM) are shown in Fig. 1. It can be seen that the inner cavities of the nanotubes are partially filled with iron (dark regions in the microphotographs of the nanotubes correspond to Fe particles; semitransparent regions, to the carbon walls). The weight fraction of Fe in the nanocomposites under consideration was estimated from the magnetization measurement results: in *s*1 it was ~50–60%; in *s*2, ~15–25%. It can also be seen in Fig. 1 that the nanotubes constituting powder samples *s*1 and *s*2 differ in their morphology: *s*1 consists of distorted nanotubes (Fig. 1a), whereas *s*2 is composed of straight-wall nanotubes (Fig. 1b). The nanotubes of both *s*1 and *s*2 types have a characteristic inner diameter of ~100 Å. Scanning electron microscopy showed that the average length of the nanotubes of both types was ~10 μm [21, 22]. X-ray diffraction studies and Mössbauer spectroscopy showed that the nanowires inside the nanotubes consisted of α-Fe and Fe_3C magnetic particles [21, 22].

Magnetic measurements were performed using a vibrating-sample magnetometer with a superconducting solenoid in fields of up to 60 kOe and in the temperature range 4.2 to 200 K. The contribution of the insert with an empty powder container was measured separately (it was ~1%) and subtracted from the results of measurement.

Results and discussion. The low-temperature magnetization curves for samples *s*1 and *s*2 are shown in Fig. 2. These curves $M(T)$ were measured in the external field $H = 20$ kOe within the temperature range 4.2 to 200 K. It can be seen that the curves do not exhibit singularities typical for superparamagnetic particles, which means that the small Fe particles inside the carbon nanotubes are exchange-coupled.

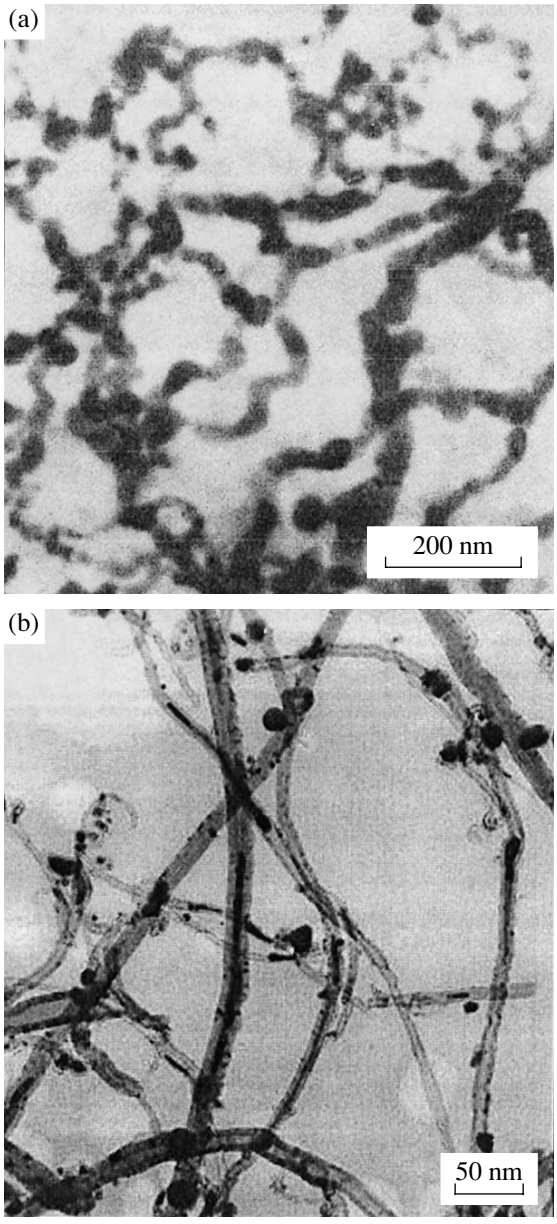


Fig. 1. (a) TEM image of iron-filled carbon nanotubes synthesized by electric arc decomposition of $Fe(CO)_5$ [22]; (b) TEM image of iron-filled carbon nanotubes synthesized by thermolysis of a mixture of C_{60} fullerene with ferrocene [21].

The experimental dependences $M(T)/M(0)$ are well described by the theoretical expression known as the Bloch law (solid lines in Fig. 2):

$$M_s(T) = M_{s0}(1 - BT^{3/2} - CT^{5/2}). \quad (2)$$

The relation of the coefficients B and C in Eq. (2) with the main magnetic constants of the material (exchange interaction constant A and the mean length $\langle r^2 \rangle^{1/2}$ of the atomic exchange coupling) is described

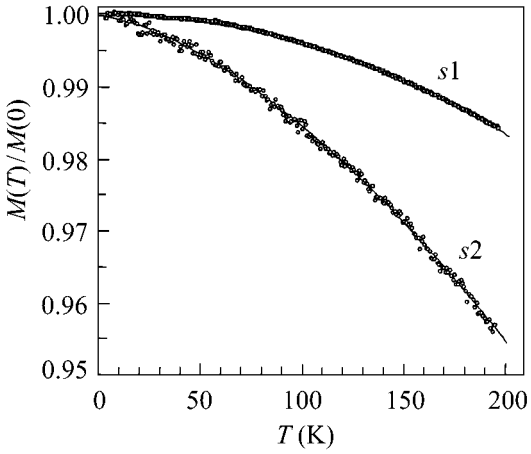


Fig. 2. Normalized magnetization $M(T)/M(0)$ as a function of temperature in an external field H = 20 kOe for Fe nanowires in carbon nanotubes.

by the well-known formulas

$$A = \frac{k}{8\pi}\left(\frac{g\mu_B}{M_{s0}}\right)^{1/3}\left(\frac{2.612}{B}\right)^{2/3}, \quad (3)$$

$$\langle r^2\rangle \equiv \sum_{ij} J_{ij} r_{ij}^4 / \sum_{ij} J_{ij} r_{ij}^2 = \frac{1.57}{\pi}\frac{C}{B^{5/3}}\left(\frac{g\mu_B}{M_{s0}}\right)^{2/3}, \quad (4)$$

where J_{ij} is the exchange integral for a pair of neighboring spins in the atomic lattice r_{ij}.

The calculated constants B, C, A, and $\langle r^2\rangle^{1/2}$ for the Fe nanowires under consideration are given in the table together with the well-known values of the same constants for an α-Fe crystal and a cementite (Fe_3C) crystal.

The measured magnetization curves of the Fe nanowires under consideration are shown in Fig. 3 both as standard isotherms $M(H)$ and as $\Delta M/M_s$ vs. $H^{-3/2}$ curves. The high-field curves of magnetization of the Fe nanowires plotted in coordinates $\Delta M/M_s$ vs. $H^{3/2}$ include rectilinear segments both at liquid helium temperature and at T = 200 K. These segments indicate that $\Delta M/M_s$ is proportional to $H^{-3/2}$ in a field range up to 60 kOe. The curves shown in Fig. 3 comply with Eq. (1). They indicate that the exponent of the power dependence describing the approach of the magnetization to saturation is independent of the temperature and the process of the synthesis of the ferromagnetic material under consideration, but depends only on the dimensionality d of the arrangement of the exchange-coupled grains. In the case under consideration, the dimensionality is 1.

Table

	(*s*1)	(*s*2)	α-Fe	Fe_3C
B, 10^{-5} K$^{-3/2}$	0.35	1.7	0.34	2.9
C, 10^{-8} K$^{-5/2}$	1.7	1.1	0.1	–
$\langle r^2\rangle^{1/2}$, Å	7	1.5	2	–
A, 10^{-6} erg/cm	2.0	0.75	2.1	0.49

The dependence $\Delta M/M_s \sim H^{-1/2}$ (d = 3) describing the approach of the magnetization to saturation in amorphous and nanocrystalline magnetic materials was predicted in [25] and experimentally obtained in [26]. It is well known to magnetologists and widely used for interpreting the experimental data on the approach of the magnetization to saturation in amorphous and nanocrystalline magnetic materials [27–32]. The dependence $\Delta M/M_s \sim H^{-1}$ (d = 2) has been recently observed in experiments with ultrathin nanocrystalline and amorphous Co layers [23]. The experimental curves of the Fe nanowire magnetization obtained in this work reveal the power dependence $\Delta M/M_s \sim H^{-3/2}$ typical for a one-dimensional chain of exchange-coupled ferromagnetic grains. It should be noted that the magnetic properties of such ferromagnetic nanowires are determined mainly by the specific structure of the spin system, which can be described as an ensemble of one-dimensional magnetic units [23, 27] or Imry–Ma domains [33]. Therefore, an increase in the coercive force H_c in nanowires in comparison with ferromagnetic films and bulk materials can be explained using the equation $\langle K\rangle = K/N^{1/2} = K(R_c/R_f)^{d/2}$ [23], where $2R_f$ is the magnetic unit size and N is the number of nanoparticles constituting a single magnetic unit. In nanostructured magnetic materials, the following relation is usually observed: $x = R_c/R_f < 1$ [23, 27]. Raising both sides of the relation to a power $d/2$, we find that, all other factors being the same, the result is the largest for d = 1 than for d = 3 or d = 2. This means that the effective anisotropy (and, therefore, the coercive force) is greater in one-dimensional exchange-coupled systems of ferromagnetic nanoparticles than in similar 2D and 3D systems.

It should be noted in conclusion that a number of papers have been published recently on the theoretical estimation and numerical simulation of the magnetization distribution and the magnetic properties of one-dimensional exchange-coupled nanosystems [34–38]. The results obtained in these papers can be applied to new magnetic systems, such as ferromagnetic nanowires (though with some reservations about the magnetodipole interaction, which is disregarded in these works). In particular, they can be used to interpret the

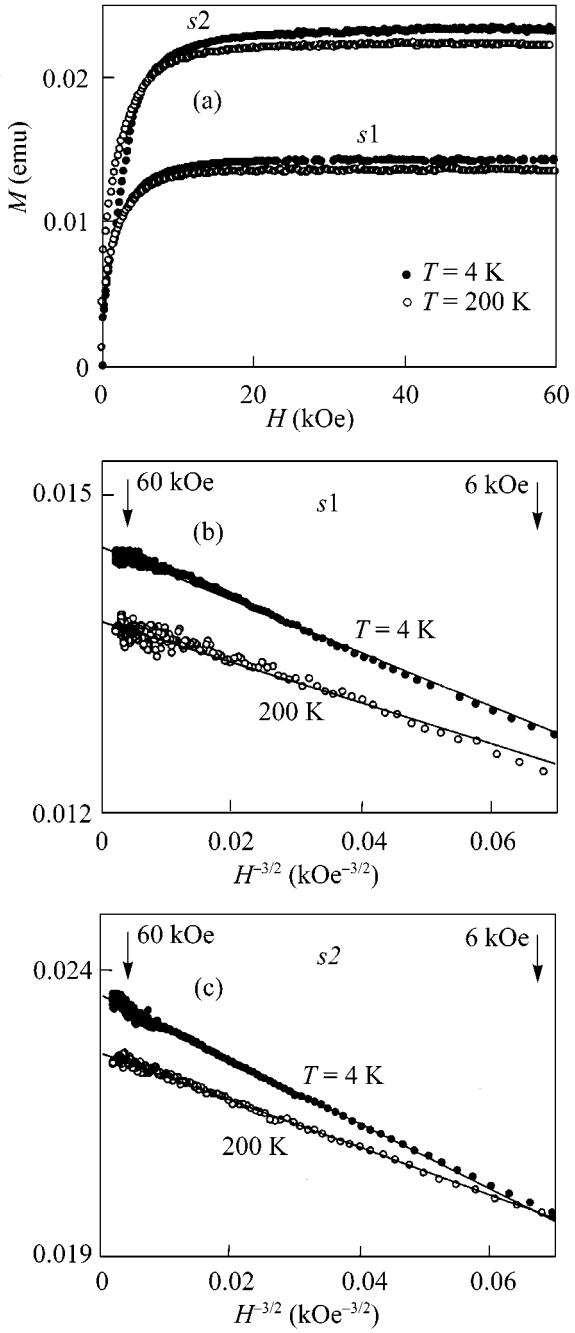


Fig. 3. (a) Curves of magnetization of Fe nanowires in carbon nanotubes; (b) high-field regions of the magnetization curves plotted in coordinates $\Delta M/M_s$ vs. $H^{-3/2}$ for sample *s*1; (c) the same for sample *s*2.

changes in some magnetic properties of nanowires through the changes in the main magnetic constants and structural parameters measured in the experiment.

This work was partially supported by INTAS, project no. 01-254. A.G. Kudashov is grateful to the Haldor Topsoe Company for financial support.

REFERENCES

1. A. Fert and L. Piraux, J. Magn. Magn. Mater. **200**, 388 (1999).
2. J.-E. Wegrowe, T. Wade, X. Hoffer, *et al.*, Phys. Rev. B **67**, 104418 (2003).
3. H. R. Khan and K. Petrikowski, J. Magn. Magn. Mater. **215–216**, 526 (2000).
4. N. Tsuya, T. Tokushima, and M. Shiraki, IEEE Trans. Magn. **23**, 53 (1987).
5. H. Zeng, R. Skomski, L. Menon, *et al.*, Phys. Rev. B **65**, 134426 (2002).
6. R. M. Metzerg, V. V. Konovalov, and M. Sun, IEEE Trans. Magn. **36**, 30 (2000).
7. D. J. Sellmyer, M. Zheng, and R. Skomski, J. Phys.: Condens. Matter **13**, R433 (2001).
8. U. Ebels, J.-L. Duvail, P. E. Wigen, *et al.*, Phys. Rev. B **64**, 144421 (2001).
9. S. G. Yang, H. Zhu, G. Ni, *et al.*, J. Phys. D **33**, 2388 (2000).
10. M. Zheng, R. Skomski, Y. Liu, *et al.*, J. Phys.: Condens. Matter **12**, L497 (2000).
11. J. B. Wang, Q. F. Liu, D. Sh. Xue, *et al.*, J. Phys. D: Appl. Phys. **34**, 3442 (2001).
12. S. A. Gusev, N. A. Korotkova, D. B. Rozenshteĭn, *et al.*, Pis'ma Zh. Tekh. Fiz. **20** (11), 50 (1994) [Tech. Phys. Lett. **20**, 450 (1994)].
13. R. Ferre, K. Ounadjela, J. M. George, *et al.*, Phys. Rev. B **56**, 14066 (1997).
14. J. H. Blythe, V. M. Fedosyuk, O. I. Kasyutich, *et al.*, J. Magn. Magn. Mater. **208**, 251 (2000).
15. G. P. Heydon, S. R. Hoon, A. N. Farley, *et al.*, J. Phys. D **30**, 1083 (1997).
16. C. Schonenberger, B. M. I. van der Zande, L. G. J. Fokkink, *et al.*, J. Phys. Chem. **101**, 5497 (1997).
17. X. X. Zhang, G. H. Wen, Sh. Huang, *et al.*, J. Magn. Magn. Mater. **231**, L9 (2001).
18. G. H. Lee, S. H. Huh, J. W. Jeong, *et al.*, J. Magn. Magn. Mater. **246**, 404 (2002).
19. N. Muhl, D. Elefant, A. Graff, *et al.*, J. Appl. Phys. **93**, 7894 (2003).
20. C. Prados, P. Crespo, J. M. Gonzalez, *et al.*, Phys. Rev. B **65**, 113405 (2002).
21. A. G. Kudashov, A. A. Pruss, O. G. Abrosimov, *et al.*, in *Proceedings of Kharkov Scientific Assembly on Diamond Films and Films of Related Materials* (2002), p. 83.
22. A. I. Okotrub, V. L. Kuznetsov, A. Sharaya, *et al.*, Khim. Interes. Ust. Razvitiya **6**, 781 (2002).
23. R. S. Iskhakov, S. V. Kolmogortsev, A. D. Balaev, and L. A. Chekanova, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 440 (2000) [JETP Lett. **72**, 304 (2000)].
24. V. A. Ignatchenko and R. S. Iskhakov, Fiz. Met. Metalloved., No. 6, 75 (1992).
25. V. A. Ignatchenko and R. S. Iskhakov, Zh. Éksp. Teor. Fiz. **72**, 1005 (1977) [Sov. Phys. JETP **45**, 526 (1977)].
26. V. A. Ignatchenko, R. S. Iskhakov, and G. V. Popov, Zh. Éksp. Teor. Fiz. **82**, 1518 (1982) [Sov. Phys. JETP **55**, 878 (1982)].

27. R. S. Iskhakov, S. V. Kolmogortsev, Zh. M. Moroz, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 872 (2000) [JETP Lett. **72**, 603 (2000)].
28. J. F. Löffler, J. P. Meier, B. Doudin, *et al.*, Phys. Rev. B **57**, 2915 (1998).
29. L. Thomas, J. Tuaillon, J. P. Perez, *et al.*, J. Magn. Magn. Mater. **140**–**144**, 437 (1995).
30. M. W. Grinstaff, M. B. Salamon, and K. S. Suslick, Phys. Rev. B **48**, 269 (1993).
31. J. Tejada, B. Martinez, A. Labarta, *et al.*, Phys. Rev. B **42**, 898 (1990).
32. M. J. O'Shea, K. M. Lee, and A. Fert, J. Appl. Phys. **67**, 5769 (1990).
33. Y. Imry and S.-K. Ma, Phys. Rev. Lett. **35**, 1399 (1975).
34. R. Dickman and E. M. Chudnovsky, Phys. Rev. B **44**, 4397 (1991).
35. A. A. Ivanov, V. A. Orlov, and G. O. Patrushev, Fiz. Met. Metalloved. **84**, 47 (1997).
36. A. A. Ivanov and G. O. Patrushev, Fiz. Met. Metalloved. **86**, 331 (1998).
37. A. A. Ivanov, V. A. Orlov, and G. O. Patrushev, Fiz. Tverd. Tela (St. Petersburg) **41**, 1432 (1999) [Phys. Solid State **41**, 1311 (1999)].
38. R. S. Iskhakov, S. V. Komogortsev, and A. V. Lukyanenko, in *Abstracts of 6th Symposium on Physics and Chemistry of Advanced Materials, Novosibirsk* (2002), p. 95.

Translated by K. Chamorovskiĭ

JETP Letters, Vol. 78, No. 4, 2003, pp. 241–245. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 276–280.
Original Russian Text Copyright © 2003 by Yakimov, Nenashev, Dvurechenskiĭ, Timonova.

Many-Electron Coulomb Correlations in Hopping Transport along Layers of Quantum Dots

A. I. Yakimov*, A. V. Nenashev*, A. V. Dvurechenskiĭ*, and M. N. Timonova**

** Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
e-mail: yakimov@isp.nsc.ru
*** Novosibirsk State University, Novosibirsk, 630090 Russia*

Received July 21, 2003

Experimental data are analyzed on the hopping transport of holes in two-dimensional layers of Ge/Si(001) quantum dots (QDs) under conditions of the long-range Coulomb interaction of charge carriers localized in QDs, when the temperature dependence of the conductivity obeys the Efros–Shklovskii law. It is found that the parameters of hopping conduction significantly deviate from the predictions of the model of one-electron excitations in "Coulomb glasses." Many-particle Coulomb correlations associated with the motion of holes localized in QDs play a decisive role in the processes of hopping charge transfer between QDs. These correlations lead to a substantial decrease in the Coulomb barriers for the tunneling of charge carriers. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 73.63.Kv; 72.20.Ee

Introduction. The problem of the effect of electron–electron interaction on the conductivity of disordered systems has always been actual and has become especially acute after the recent stormy discussion of the existence of a metal–insulator transition (MIT) in two dimensions. The role of Coulomb correlations must be especially significant on the insulator side of the MIT, because localized electronic states are ineffective in screening processes as compared to extended states. The systems in which charge carriers are randomly localized in space with long-range Coulomb interactions existing between them received the name "Coulomb glasses." In the regime of strong electron localization, when the localization length ξ is much smaller than the distance between the localized states, charge transfer is carried out by means of tunnel hops of electrons from one center to another, and the hop length increases with decreasing temperature [1]. Under these conditions, the dependence of conductivity G on temperature T has the form

$$G(T) = G_0 \exp[-(T_0/T)^x], \quad (1)$$

where parameter T_0 is determined by the properties of the material, and exponent $x < 1$ is determined by the energy dependence of the density of states at the Fermi level $g(E_f)$. If the electron–electron interaction in the system is insignificant and $g(E_f) = \text{const}$, $x = 1/3$ (Mott's law for two dimensions), $T_0 = 13.8/k_B g(E_f)\xi^2$, where k_B is the Boltzmann constant [2]. Efros and Shklovskii [3] showed that the interaction of localized electrons in Coulomb glasses lead to the expression

$$G(T) = G_0 \exp[-(T_0/T)^{1/2}], \quad (2)$$

subsequently called the Efros–Shklovskii law. Here,

$$T_0 = C_2 e^2/k_B \kappa \xi, \quad (3)$$

C_2 is a numerical parameter, and κ is the relative permittivity. The quantity $k_B T_0$ serves as a measure of the characteristic scale of the Coulomb interaction in the system.

It is important to note that Eq. (2) with the numerical constant $C_2 = 6.2$ for a two-dimensional system [4] was obtained upon the consideration of only one-electron excitations, which represent the transfer of an electron from one center to another under the condition that all the other electrons are frozen at their positions. The one-electron model did not take into account possible many-particle correlations of electron hops, when, for example, hops of some electrons to small distances, either parallel or sequential in time, facilitate the motion of other electrons to long distances, decreasing the corresponding energy barriers on the way of the current flow by their Coulomb potential. Because the formation of such a many-electron polaron leads to the screening of the Coulomb potential at long hop lengths, the characteristic energy scale of Coulomb correlations for such quasiparticles must be significantly reduced as compared to the case of one-electron excitations.

A number of experiments showed that actually many-particle effects can notably decrease the Coulomb gap width in the spectrum of electronic states in the impurity band of doped semiconductors under conditions of moderate compensation [5, 6].

0021-3640/03/7804-0241$24.00 © 2003 MAIK "Nauka/Interperiodica"

A numerical simulation of hopping conduction in a two-dimensional Coulomb glass was performed in [7, 8]. The authors of [7] extended the percolation approach to the space of many-particle configurations of site occupation numbers on a network of Miller–Abrahams resistances and showed that law (2) is also obeyed in the case of many-electron correlations. However, in this case $C_2 = 0.6 \pm 0.2$; that is, the characteristic interaction parameter T_0 is an order of magnitude smaller than the value obtained in the one-electron approximation. Tsigankov and Efros [8] applied the kinetic Monte Carlo method to modeling the hopping charge transfer in a two-dimensional system of interacting electrons in which all the sequential transitions and the simultaneous (in the quantum-mechanical sense) transitions of two electrons were taken into account. The conclusion was made that many-particle correlations are insignificant and $C_2 = 5.8$. Even though the majority of experimental data obtained with doped semiconductors are in agreement with this conclusion, such a strong difference between the conclusions of the theoretical studies cannot help but stimulate experimental studies with objects in which the dominating role of long-range Coulomb interaction in charge transport is well established.

Layers of self-assembled quantum dots (QDs) of Ge in Si that are formed in the heteroepitaxy of elastically strained systems are among such objects. In experiments with an artificial screen introduced into Ge/Si heterostructures parallel to the array of Ge QDs, it was demonstrated that electron–electron interaction is fully responsible for the temperature dependence of the hopping conductance along the QD layer [9]. In [10], the preexponential factor of the hopping conductance G_0 (see expression (2)) was found to be independent of temperature, and its value was found to be a multiple of the conductance quantum e^2/h. This result provided evidence for the phononless character of hopping conduction in the systems with low disorder [11]. The goal of this work was to analyze experimental data related to the hopping transport of holes in two-dimensional layers of Ge/Si(001) QDs from the viewpoint of various models of electron excitations in a two-dimensional disordered system.

Formation of Ge/Si heterostructures with QDs. Ge/Si heterostructures containing Ge QD arrays were grown by molecular beam epitaxy. An ensemble of pyramidal Ge nanoclusters was grown on the Si(001) surface using the effect of spontaneous morphological transformation of an elastically strained Ge layer in the process of Stranskii–Krastanov growth [12]. The conductance was measured in the planar geometry (along the layers of Ge/Si QDs) with samples of four types.

Samples of series #A were grown on Si substrates with a resistivity of 1000 Ω cm doped with boron up to a concentration of $\sim 10^{13}$ cm^{-3}. A Ge layer eight monolayers thick (≈10 Å) was embedded into the middle of a 90-nm Si layer grown on the substrate. The mean sizes of the base of the forming Ge nanoclusters in the growth plane were 10 nm, the height was ~1 nm, and the layer density of nanoclusters was $\sim 4 \times 10^{11}$ cm^{-2} (the details of sample preparation can be found in [10]).

In the samples of the subsequent series #B, #C, and #D, the effective thickness of a Ge layer was ten monolayers. The lateral sizes of Ge QDs amounted to 15 nm, the mean height was 1.5 nm, and the layer density was $\sim 3 \times 10^{11}$ cm^{-2} [9, 13].

The sequence and structure of the first layers in the samples of series #B and #C were the same as in #A. In #B, the layer of Ge QDs was placed inside the Si layer at a distance of 40 nm from its surface. In samples #C, the array of Ge islands was covered by a 10-nm Si layer and then by a 25-nm surface layer of SiO_2. The presence of the oxide layer changed the relative permittivity of the medium and, hence, the potential of the long-range electron–electron interaction in the sample.

The controlled filling of Ge islands with holes in structures #A, #B, and #C was carried out by embedding a Si layer δ-doped with boron at a distance of 5 nm below the QD layer. Because the ionization energy of boron impurities in silicon is only 45 meV and the depth of the first ten energy levels of holes in germanium pyramids of such sizes measured from the top of the Si valence band is 200–400 meV [14], holes at low temperatures leave the impurities and occupy levels in QDs. The concentration of boron in different samples varied from 2×10^{11} cm^{-2} to 2.25×10^{12} cm^{-2}, which allowed the mean number of holes N_h per one Ge QD to be varied in the range from $N_h = 0.5$ to $N_h = 6.5$. A numerical simulation of the energy spectrum of holes and their wave functions in pyramidal nanoclusters was performed in [14]. The ground state of a hole has an *s*-type symmetry and is doubly degenerate with respect to the spin orientation. The first excited state has a *p*-type symmetry with a degeneracy factor of 4 [14]. Therefore, for example, at $N_h = 6.5$, the first two electron shells turn out to be occupied, and the third one is occupied partially.

The samples of series #D represented silicon metal–oxide–semiconductor (MOS) field-effect transistors (MOSFETs) formed on silicon-on-insulator wafers with an array of Ge QDs embedded at a distance of 40 nm from the Si surface [13]. The filling of QDs with holes was controlled by applying a corresponding voltage across the aluminum gate. The corresponding degree of filling was determined by oscillations of the dependence of the source–drain current on the voltage across the gate.

The ohmic contacts were formed by sputtering Al plates onto the sample surface and subsequently heating the structure at a temperature of 400°C in a nitrogen atmosphere. In all cases, the measurements of the current at various temperatures were performed in the ohmic section of the current–voltage and drain characteristics.

Temperature dependence of the hopping conductance. Previously in [9, 10, 13], we showed that the Efros–Shklovskii law (2) with the preexponential factor $G_0 \simeq e^2/h$ is obeyed to a good accuracy in all four series of samples. To find the hopping conductance parameters G_0, T_0, and x in the analysis of experimental data, we used methods of the differential analysis of the dimensionless activation energy of the conductance $w(T) = \partial \ln G(T)/\partial \ln T$ [15] and nonlinear regression methods [10], which give most accurate results as compared to the method of the rectification of the dependence $G(T)$ constructed in different coordinates.

As an example, Fig. 1 demonstrates the experimental temperature dependences of the conductance in the units of e^2/h constructed on the $\log G - T^{-1/2}$ coordinates for the samples of series #A. The symbols correspond to experimental points, and the solid lines are the results of fitting the experimental data by the equation $G(T) = \gamma T^m \exp[-(T_0/T)^{1/2}]$, where T_0, γ, and m are variable parameters.[1] It was shown that $m \approx 0$, and the values of T_0 were found [9, 10]. The experimental data for samples #C and #D are displayed in Fig. 2.

In order to determine the character of Coulomb correlations experimentally, it is necessary to compare the experimental values of T_0 with the values predicted by different models. However, because, according to expression (3), T_0 depends on the permittivity and the localization length and, hence, on the sizes of the QDs, their density, etc., it is more convenient to compare the values of the universal parameter C_2, which describes the hopping transport observed experimentally, with the values that give various models of Coulomb correlations in Coulomb glasses. Parameter C_2 can be found from expression (3) if the localization length ξ and T_0 are known. Usually, ξ is determined from measurements of the positive magnetoresistance in the hopping conduction mode caused by the contraction of the wave functions of localized carriers in the plane perpendicular to the magnetic field [5]. However, it can turn out that this procedure does not give the correct result, because the expressions for the magnetic-field dependence of the hopping conductance themselves were already obtained within the framework of a priori assumptions of the occurrence of correlation effects. In the next section, we will present the results of modeling the wave functions of holes in arrays of Ge/Si QDs, which allowed us to determine the localization lengths and, then, parameter C_2.

Calculation of the localization radius of holes in arrays of Ge/Si QDs. In modeling the localization length of the wave functions of holes in the ground and first excited states, we considered an array of Ge/Si QDs located at sites of a square lattice 15×15 in size.

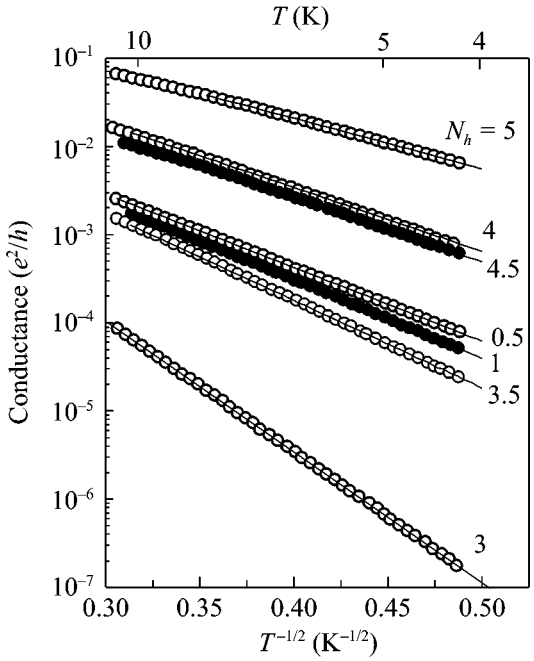


Fig. 1. Temperature dependences of the conductance for the samples of series #A constructed on the $\log G - T^{-1/2}$ coordinates. Symbols are experimental points, and solid lines are the results of fitting the experimental data by the equation $G(T) = \gamma T^m \exp[-(T_0/T)^{1/2}]$. T_0, γ, and m are variable parameters.

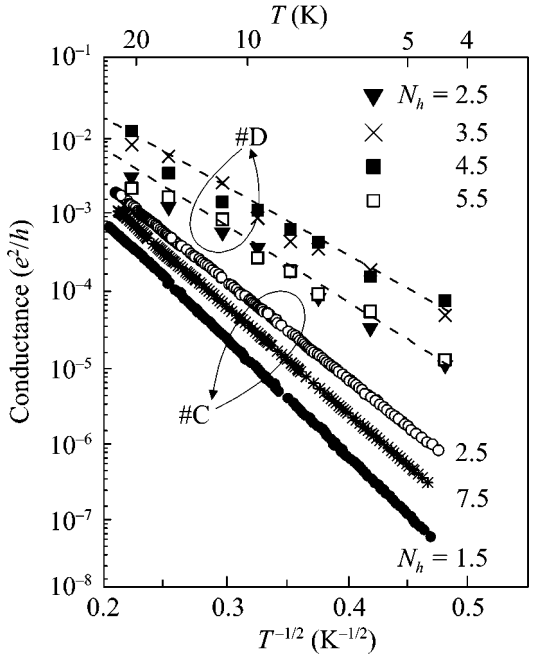


Fig. 2. Temperature dependences of the conductance for the samples of series #C and #D.

Only the overlaps of the states that belong to neighboring QDs were taken into account, because the overlap integrals rapidly drop with increasing distance. The dis-

[1] In [10], the dependences $G(T)$ are also given for N_h = 1.5, 2, 2.5, 5.5, and 6. However, because, at these occupation numbers, $T_0 \sim T$ these data cannot be used for correctly revealing the hopping conductance parameters.

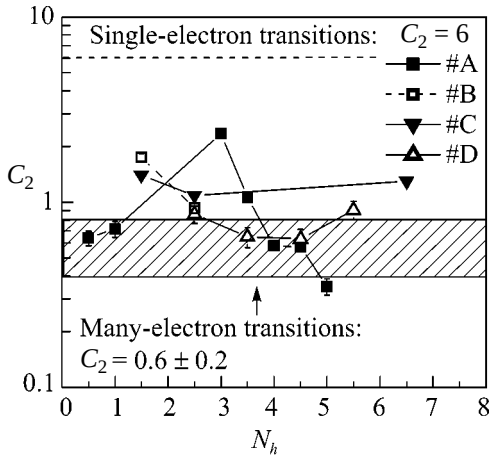


Fig. 3. The values of parameter C_2 in the samples of various series and with various occupations of quantum dots with holes (symbols). The value $C_2 = 6.2$ obtained within the one-electron model [4] is shown with a dashed line. The hatched area corresponds to the range of values $C_2 = 0.6 \pm 0.2$, which describe many-particle excitations in Coulomb glasses [8]. The experimental value of C_2 averaged over all samples equals 0.98 ± 0.47.

tances between the centers of neighboring points were taken equal to $1/\sqrt{n_{QD}}$, where n_{QD} is the two-dimensional density of QDs plus a random component that has a Gaussian distribution. The Hamiltonian of the system was written in the form

$$\hat{H} = \sum_{i,\alpha} E_{i,\alpha}\hat{a}^{+}_{i,\alpha}\hat{a}_{1,\alpha} + \sum_{i,j,\alpha,\beta} J_{i,j,\alpha,\beta}\hat{a}^{+}_{j,\beta}\hat{a}_{i,\alpha}, \qquad (4)$$

where index i numbers the QDs, index α ranges over the hole bound states in the QDs (only the first nine states were taken into account); $\hat{a}^{+}_{i,\alpha}$ and $\hat{a}_{1,\alpha}$ are operators of hole creation and annihilation in the αth state of the ith QD, $E_{i,\alpha}$ is the hole energy in this state, and $J_{i,j,\alpha,\beta}$ is the overlap integral between the αth state in the ith QD and the βth state in the jth QD. Because Ge nanoclusters in real structures have a dispersion of sizes (~20% [16]), a random energy $E_{i,\alpha}$ was assigned to each site. This energy was determined as the size-quantization energy in QDs with the particular sizes and was calculated in the tight-binding approximation with the sp^3 basis set with regard to the spin–orbit interaction and deformation effects [14].

The overlap integrals were found by the following procedure. The energies of the hole states were calculated for model structures containing Ge QDs inside the silicon region shaped as a parallelepiped. At the parallelepiped boundaries, the following periodic boundary conditions were introduced: $\psi(-d/2, y, z) = \psi(d/2, y, z)$ or $\psi(-d/2, y, z) = -\psi(d/2, y, z)$, where d is the parallelepiped size in the direction x, and $\psi(x, y, z)$ is the wave function. Analogous boundary conditions were written for the directions y and z. Such structures are similar to an infinite crystal at whose sites QDs are located with period d. The value of the overlap integral for the distance d between the centers of QDs was found in the direction of the x axis lying in the growth plane and was determined as $J(d) = |E_+ - E_-|/4$, where E_+ and E_- are the values of the hole energy corresponding to the above boundary conditions. The obtained dependence $J(d)$ can be presented in the form

$$J(d) = A_\alpha \exp(-B_\alpha d), \qquad (5)$$

where coefficients A_α and B_α depend on the energy level number α. Coefficient B_α represents the inverse of the localization radius of the αth hole state in the isolated QD. The overlap integrals $J_{i,j,\alpha,\beta}$ between states with different numbers α,β were determined as the geometric mean of the integrals between the states with number α and number β

$$J_{i,j,\alpha,\beta} = \sqrt{A_\alpha A_\beta}\exp\left(-\frac{B_\alpha + B_\beta}{2}d_{ij}\right), \qquad (6)$$

where d_{ij} is the distance between the ith and jth QDs.

The Schrödinger equation $\hat{H}|\varphi\rangle = E|\varphi\rangle$ was solved numerically using the ARPACK program package in the MatLab system. The calculations were performed with 5000 random realizations of the QD array. For each realization, a state was determined whose energy was most close to the hole energy in the ground and excited states averaged over the array. Next, the probabilities p_i that a hole can be found in each QD were calculated. After that, the section of the QD array passing in the direction parallel to the sides of the lattice squares through the point with the maximum probability of hole finding was considered. In this section, the values of the probabilities p_i were fitted using a function of the form $p_i = a\exp(-2bd_i)$, where d_i is the distance between the ith QD and the point with the maximum probability of hole finding. The localization radius ξ was found as the inverse of the parameter b averaged over all the realizations of the QD array. The following results were obtained: for the samples of series #A, ξ = 2.41 nm for the hole ground state and 2.78 nm for the first excited state; for samples of series #B, #C, and #D, ξ = 2.06 nm for the hole ground state, 2.23 nm for the first excited state, and 2.30 nm for the second excited state.

Finding parameter C_2. The numerical value of parameter C_2 was determined from expression (3). We took the experimental values of T_0 and the calculated values of ξ. Because, in the samples of series #A, #B, and #D, the layer of Ge QD is embedded into Si, the permittivity κ for these samples was taken as $\kappa \equiv \kappa_{Si}$ = 12. For sample #C, $\kappa_{Si} = 9$ [9], because a layer of SiO_2 is located near the QD, which decreases the effective permittivity. The results are given in Fig. 3.

The value of C_2 obtained in experiments with various samples (averaging over the samples gives $C_2 = 0.98 \pm 0.47$) was found to be much less than the "one-electron" value. This means that the processes of hopping charge transfer in two-dimensional Ge/Si QD arrays are determined to a large extent by many-electron Coulomb correlations.

This work was supported by the Russian Foundation for Basic Research, project no. 03-02-16526; the program of the President of Russian Federation for the support of young Doctors of science, project no. MD-28-2003-02; and the program "Universities of Russia," project no. UR.01.01.019).

REFERENCES

1. N. Mott, J. Non-Cryst. Solids **1**, 1 (1968).
2. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer, New York, 1984).
3. A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).
4. V. L. Nguen, Fiz. Tekh. Poluprovodn. (Leningrad) **18**, 335 (1984) [Sov. Phys. Semicond. **18**, 207 (1984)].
5. A. G. Zabrodskiĭ, Usp. Fiz. Nauk **168**, 804 (1998) [Phys. Usp. **41**, 722 (1998)]; A. G. Zabrodskii, A. G. Andreev, and S. V. Egorov, Phys. Status Solidi B **205**, 61 (1998).
6. M. Lee and J. G. Massey, Phys. Status Solidi B **205**, 25 (1998).
7. A. Pérez-Garrido, M. Ortuno, E. Cuevas, and J. Ruiz, Phys. Rev. B **55**, R8630 (1997).
8. D. N. Tsigankov and A. L. Efros, Phys. Rev. Lett. **88**, 176602 (2002).
9. A. I. Yakimov, A. V. Dvurechenskii, V. V. Kirienko, *et al.*, Phys. Rev. B **61**, 10 868 (2000).
10. A. I. Yakimov, A. V. Dvurechenskiĭ, A. I. Nikiforov, and A. A. Bloshkin, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 445 (2003) [JETP Lett. **77**, 376 (2003)].
11. V. I. Kozub, S. D. Baranovskii, and I. S. Shlimak, Solid State Commun. **113**, 587 (2000).
12. O. P. Pchelyakov, Yu. B. Bolhovityanov, A. V. Dvurechenskii, *et al.*, Thin Solid Films **367**, 75 (2000).
13. A. I. Yakimov, A. V. Dvurechenskii, V. V. Kirienko, *et al.*, J. Phys.: Condens. Matter **11**, 9715 (1999).
14. A. V. Dvurechenskii, A. V. Nenashev, and A. I. Yakimov, Nanotechnology **13**, 75 (2002).
15. A. G. Zabrodskiĭ and K. N. Zinov'eva, Zh. Éksp. Teor. Fiz. **86**, 727 (1984) [Sov. Phys. JETP **59**, 425 (1984)].
16. A. I. Yakimov, A. V. Dvurechenskii, Yu. Yu. Proskuryakov, *et al.*, Appl. Phys. Lett. **75**, 1413 (1999).

Translated by A. Bagatur'yants

JETP Letters, Vol. 78, No. 4, 2003, pp. 246–249. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 281–285.
Original Russian Text Copyright © 2003 by Konorov, Sidorov-Biryukov, Bugar, Chorvat, Jr., Chorvat, Zheltikov.

Quantum Control of Two-Photon Photochromism in the Solid Phase

S. O. Konorov[1,2], D. A. Sidorov-Biryukov[2], I. Bugar[3], D. Chorvat, Jr.[3], D. Chorvat[3], and A. M. Zheltikov[1,2,*]

[1] *Physics Department, Moscow State University, Vorob'evy gory, 119899 Moscow, Russia*
[2] *International Laser Center, Moscow State University, Vorob'evy gory, 119899 Moscow, Russia*
[3] *International Laser Center, Ilkovićova 3, 81219 Bratislava, Slovak Republic*
**e-mail: zheltikov@top.phys.msu.su*

Received July 11, 2003

Phase and polarization control of photochromism in the solid phase is experimentally demonstrated. Photochromism initiated by a two-photon absorption of femtosecond laser pulses in a three-dimensional solid-phase polymer sample of spiropyran is controlled by varying the polarization state and the chirp of laser pulses.
© *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 82.50.Nd; 42.70.Gi

Photochromism—a light-induced reversible transformation in chemical species between two forms having different absorption spectra—has been known since the 19th century. In 1867, Fritzsche [1] published the first observation of this phenomenon in tetracene exposed to the daylight. The term photochromism, combining the Greek words phos (light) and chroma (color), was introduced in the middle of the 20th century [2], when the photochromism of spiropyrans, one of the most widespread photochromic compounds, was also discovered [3, 4]. Traditional applications of photochromic materials include [5–7] sunglasses, optical filters, switches, and photography. Materials of this class are currently intensely studied in the context of three-dimensional optical memory [8, 9], creation of components for integrated optics and photonics [10, 11], reversible micromachining of polymers [12], and development of optically switchable biomaterials and biophotonic devices [13, 14].

The photochromic reaction in spiropyran compounds proceeds [5–7] through C–O bond cleavage in an initially nonexcited molecule (form A, see the inset in Fig. 1) and relaxation from a transient state X to a metastable merocyanine-form state (form B). Time-resolved studies of the photochromic reaction performed with the use of femtosecond pulses [15–17] have revealed the existence of the initial, subpicosecond-scale phase of the photochromic process. The key idea of this work is to use femtosecond pulses not only to probe, but also to control the ultrafast dynamics involved in photochromism. Quantum-control strategies are now widely used to direct ultrafast physical and chemical processes in the gas and liquid phases [18, 19]. Recent experiments [20] on the quantum control of wave-packet dynamics in polymers extend the technique of quantum control to the solid phase. The main goal of this work is to explore the possibility of applying this technique to photochromism in the solid phase. We have studied the photochromic reaction initiated by a two-photon absorption of femtosecond laser pulses in a three-dimensional polymer sample of spiropyran. The results of our experiments suggest the possibility of controlling photochromic transformations in the solid phase by varying the polarization state and the chirp of laser pulses.

The femtosecond laser system employed in our experiments (Fig. 2) consisted of a Ti: sapphire master oscillator, a stretcher, an amplifier, and a pulse compressor. The Ti: sapphire master oscillator was pumped by 4-W cw radiation of a diode-laser-pumped Nd: YVO_4 Verdi laser. The master oscillator generated laser

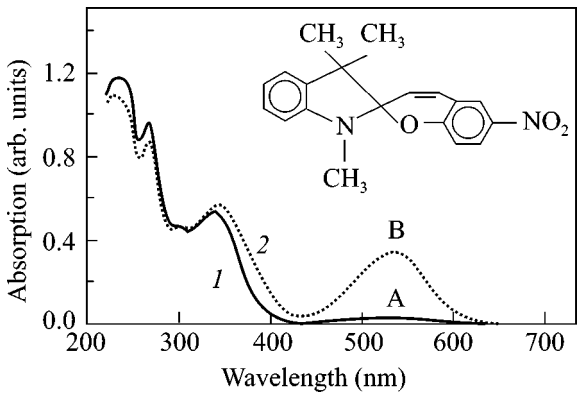


Fig. 1. Absorption spectra of spiropyran molecules in form A (*1*) and form B (*2*). The inset shows the structure formula of spiropyran molecules.

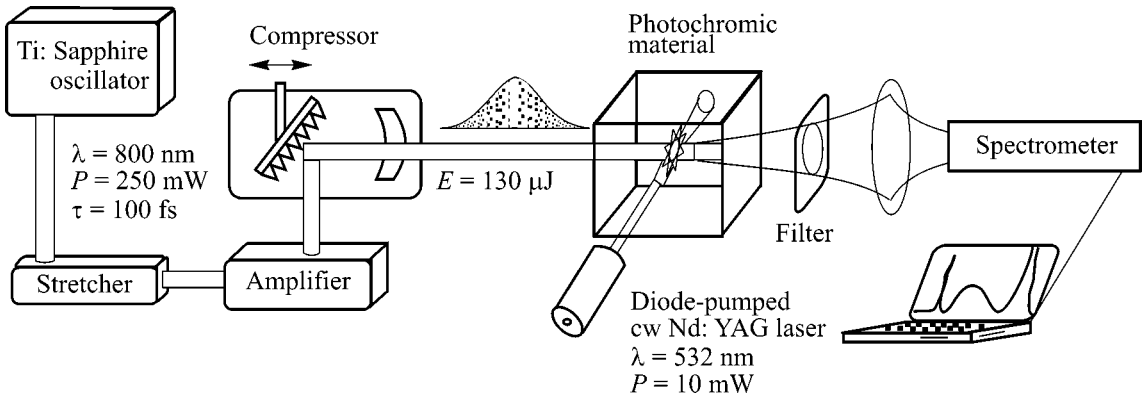


Fig. 2. Diagram of the experimental setup based on a femtosecond Ti: sapphire laser with an amplification stage.

pulses with a duration of 50–100 fs, a typical average output power on the order of 250 mW, and a pulse repetition rate of 100 MHz. Femtosecond pulses produced by the master oscillator were stretched up to 800 ps and launched into a multipass Ti: sapphire amplifier pumped with a nanosecond Nd: YAG laser with intracavity second-harmonic generation. Amplified 1-kHz picosecond pulses with an energy up to 300 μJ were then compressed to a duration of 100–130 fs in a single-grating pulse compressor. Approximately 50% of laser energy was lost at this pulse-compression stage.

Our experiments were performed with PMMA samples doped with spiropyran molecules (shown in the inset to Fig. 1). The concentration of spiropyran in a $9 \times 10 \times 10$ mm^3 sample was 1.6×10^{-2} mol/l [21, 22]. Absorption spectra of the uncolored and colored forms (forms A and B) of spiropyran molecules in the PMMA host are shown by lines *1* and *2* in Fig. 1, respectively.

Second-harmonic radiation (with a wavelength of 532 nm), produced by a diode-pumped continuous-wave Nd: YAG laser, was employed to excite the photoluminescence signal in the area where the photochromic reaction was initiated through the two-photon absorption of Ti: sapphire-laser radiation (Fig. 3). The photoluminescence signal served in our experiments for a quantitative characterization of the efficiency of the photochromic transformation. Continuous-wave 532-nm radiation also stimulated the reverse photochromic transformation, partially recovering form-A spiropyran in the sample. With Ti: sapphire laser pulses in two arms of the optical scheme being mismatched in time or space, the photoluminescence signal remained constant in time, indicating a dynamic equilibrium of the forward and backward photochromic reactions, thus showing no systematic pulse-to-pulse accumulation of form-B spiropyran generated by each of the femtosecond pulses in the laser-irradiated area of the photochromic sample.

When Ti: sapphire-laser pulses coming from different arms of the optical scheme were matched in time and space in the photochromic sample, the photoluminescence signal increased (Fig. 4), indicating the growth in the concentration of form-B spiropyran, i.e., the enhancement of the photochromic reaction. The results of photoluminescence measurements are presented by the solid line in Fig. 4 against the cross-correlation trace of the same pair of pulses measured in a 1-mm-thick BBO crystal (the dashed line), showing the negligibility of dispersion and pulse propagation effects in our experiments.

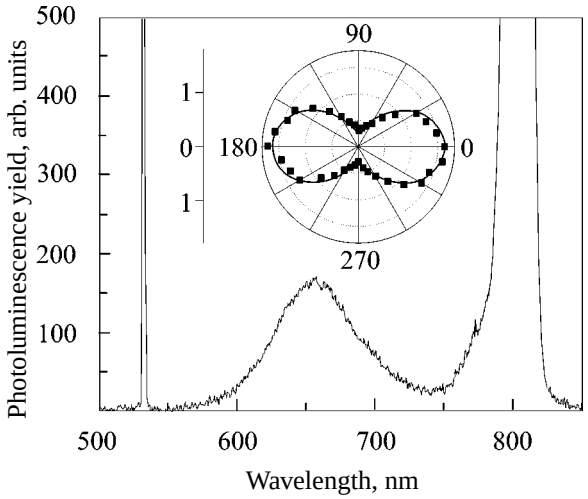


Fig. 3. The spectrum of the photoluminescence (PL) signal from form-B spiropyran molecules excited by 532-nm second-harmonic radiation of the continuous-wave Nd: YAG laser in the area of a photochromic reaction induced by two-photon absorption of Ti: sapphire-laser radiation. Spectral lines at 532 and 800 nm are related to second-harmonic radiation, employed to excite photoluminescence (532 nm), and radiation initiating the photochromic reaction (800 nm). The dots in the inset show the dependence of the photoluminescence signal on the angle φ between the polarization vectors of linearly polarized Ti: sapphire laser pulses initiating the photochromic reaction. The polar coordinates represent the photoluminescence yield and the angle φ. The solid line in the inset displays the approximation function $I(\varphi) \propto \eta \sin^2\varphi + \cos^2\varphi$ with $\eta = 1.6$.

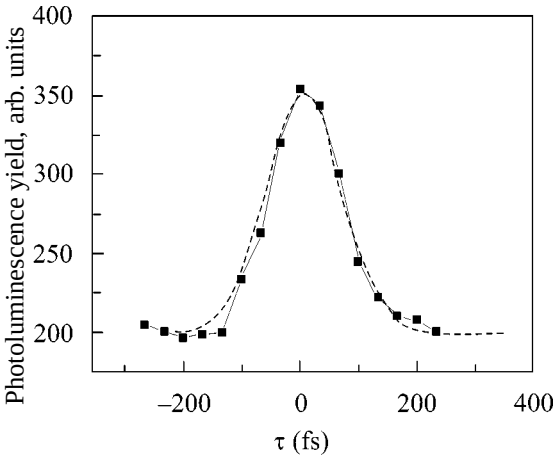


Fig. 4. The solid line displays the yield of photoluminescence excited in the spiropyran/PMMA sample by 532-nm second-harmonic radiation of a diode-pumped continuous-wave Nd: YAG laser as a function of the delay time τ between Ti: sapphire-laser pulses initiating the photochromic reaction. The dashed line shows the cross-correlation trace of the same pair of pulses measured with a 1-mm-thick BBO crystal.

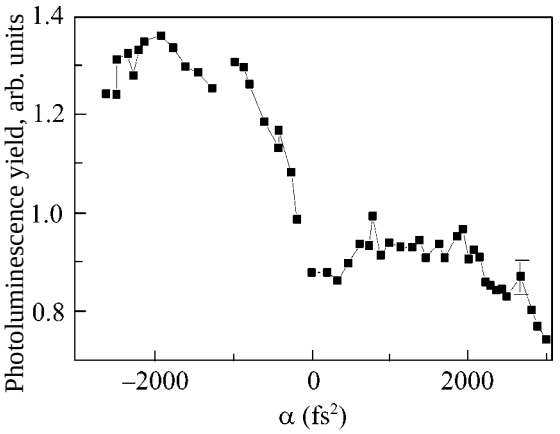


Fig. 5. The yield of the photoluminescence signal excited with continuous-wave 532-nm laser radiation in the Ti: sapphire-laser-irradiated area of a spiropyran/PMMA sample as a function of the initial chirp α of Ti: sapphire-laser pulses initiating the photochromic reaction.

Ti: sapphire laser pulses initiate the photochromic reaction in our experiments through two-photon absorption. The cross section of two-photon absorption is thus the key parameter, which controls the efficiency of photochromic processes. The cross section of two-photon absorption can be controlled, in turn, by varying the polarization state of the fields initiating the photochromic reaction [23]. The polarization dependence of the two-photon absorption cross section, determined by the tensor properties of this cross section, suggests the way to control two-photon photochromism by changing polarizations of the electromagnetic fields in laser pulses initiating the photochromic reaction.

To investigate the influence of the polarization of Ti: sapphire laser radiation on the efficiency of photochromic processes, we measured the yield of photoluminescence excited by second-harmonic radiation of the continuous-wave Nd: YAG laser as a function of the angle φ between the polarization vectors of linearly polarized fields in 130-fs Ti: sapphire-laser pulses initiating the photochromic reaction. This experimental polarization dependence (shown by dots in the inset to Fig. 3) was approximated with the function $I(\varphi) \propto \eta\sin^2\varphi + \cos^2\varphi$. The physical meaning of the parameter η in this approximation of the function $I(\varphi)$ can be understood by considering the relation between the two-photon absorption cross section and the imaginary part of the cubic nonlinear-optical susceptibility tensor $\chi^{(3)}_{ijkl}$, which gives $\eta = [\mathrm{Im}(\chi^{(3)}_{1122})]^2/[\mathrm{Im}(\chi^{(3)}_{1111})]^2$. The results of experiments presented in the inset to Fig. 3 show that linearly polarized femtosecond pulses with parallel polarization vectors produce form-B spiropyran in a spiropyran/PMMA sample approximately six times more efficiently than linearly polarized pulses with perpendicular polarization vectors. In the presence of a low-power 532-nm second-harmonic radiation, stimulating the reverse photochromic process, we were able to shift the dynamic equilibrium between the forward and backward photochromic reactions toward either form-A or form-B spiropyran generation by changing the angle between the polarization vectors of the fields in Ti: sapphire-laser pulses.

Figure 5 displays the yield of photoluminescence excited with second-harmonic radiation of the continuous-wave Nd: YAG laser as a function of the initial chirp of Ti: sapphire-laser pulses, initiating the photochromic reaction. As can be seen from these results, the efficiency of the photochromic process is highly sensitive to the initial chirp of femtosecond pulses. Negatively chirped laser pulses produce form-B spiropyran much more efficiently than transform-limited or positively chirped laser pulses do. This finding can be interpreted in terms of the dynamics of vibrational wave packets produced by femtosecond laser pulses in spiropyran molecules. The C–O bond cleavage is then represented as an evolution of a wave packet consisting of vibrations of an electronically excited spiropyran molecule [16]. Femtosecond pulses are considered as wave packets in the frequency domain. Being absorbed by a ground-state photochromic molecule through a two-photon process, such pulses excite a group of vibrational levels, leading to the formation of a vibrational wave packet in the electronically excited state. This two-photon-absorption-excited vibrational wave packet then evolves, as indicated by extensive experimental data [16, 17], toward the transient state X on the femtosecond time scale, eventually resulting in the formation of the metastable colored form (form B) of spiropyran molecules. Since the vibrational potential sensed by the wave packet in the electronically excited

state is generally anharmonic, individual vibrations forming the excited-state wave packet reach the target transient state X at different moments of time. Initially chirped laser pulses, according to the general idea of quantum control with chirped pulses [24, 25], help to synchronize individual oscillations in the target state. The lack of the information on the potential surface for spiropyran molecules in a solid-phase host does not allow us to quantify the details of this wave-packet scenario at the present stage of research.

Experimental studies of photochromic processes induced by a two-photon absorption of femtosecond laser pulses in a three-dimensional polymer sample of spiropyran demonstrate the possibility to control photochromism in the solid phase by changing the polarization state and the chirp of laser pulses. Linearly polarized femtosecond pulses with parallel polarization vectors have been shown to enhance the two-photon-absorption-induced photochromic reaction by a factor of about 6 as compared with linearly polarized femtosecond pulses having perpendicular polarization vectors. Negatively chirped laser pulses have been demonstrated to initiate the photochromic process much more efficiently than transform-limited or positively chirped laser pulses. Quantum-controlled photochromism, demonstrated in this work, offers new strategies for the creation of coherence- and polarization-controlled components for photonics and optical telecommunications, three-dimensional optical data storage, reversible microfabrication, and the development of photoswitchable biomaterials and biophotonic devices.

We are grateful to N.T. Sokolyuk, A.B. Fedotov, and S.A. Magnitskii for illuminating discussions. This study was supported in part by the President of the Russian Federation Grant MD-42.2003.02, the Russian Foundation for Basic Research (projects nos. 03-02-16929 and 02-02-17098), and the Volkswagen Foundation (project no. I/76 869).

REFERENCES

1. J. Fritzsche, C. R. Acad. Sci. **69**, 1035 (1867).
2. Y. Hirshberg, C. R. Acad. Sci. **231**, 903 (1950).
3. E. Fischer and Y. Hirshberg, J. Chem. Soc. 4522 (1952).
4. Y. Hirshberg, J. Am. Chem. Soc. **78**, 2304 (1956).
5. *Photochromism: Molecules and Systems*, Ed. by H. Dürr and H. Bouas-Laurent (Elsevier, Amsterdam, 1990), Studies in Organic Chemistry, Vol. 40.
6. *Photochromism, Techniques of Chemistry III*, Ed. by G. H. Brown (Wiley, New York, 1971).
7. V. A. Barachevskiĭ, G. I. Lashkov, and V. A. Tsekhomskiĭ, *Photochromism and Its Applications* (Khimiya, Moscow, 1977).
8. D. A. Parthenopoulos and P. M. Rentzepis, Science **245**, 843 (1989).
9. D. A. Akimov, A. M. Zheltikov, N. I. Koroteev, *et al.*, Quantum Electron. **26**, 848 (1996).
10. S. Lecomte, U. Gubler, M. Jäger, *et al.*, Appl. Phys. Lett. **77**, 921 (2000).
11. S. Cattaneo, S. Lecomte, Ch. Bosshard, *et al.*, J. Opt. Soc. Am. B **19**, 2032 (2002).
12. S. O. Konorov, A. B. Fedotov, and A. M. Zheltikov, Appl. Phys. B **76**, 707 (2003).
13. G. Berkovic, V. Krongauz, and V. Weiss, Chem. Rev. **100**, 1741 (2000).
14. H. Asanuma, K. Shirasuka, T. Yoshida, *et al.*, Chem. Lett. **2**, 108 (2001).
15. N. P. Ernstig, B. Dick, and Th. Arthen-Engeland, J. Phys. Chem. **95**, 5502 (1991).
16. J. Z. Zhang, B. J. Schwartz, J. C. King, and C. B. Harris, J. Am. Chem. Soc. **114**, 10921 (1992).
17. S. A. Antipin, A. N. Petrukhin, F. E. Gostev, *et al.*, Chem. Phys. Lett. **331**, 378 (2000).
18. *Femtosecond Coherent Raman Spectroscopy*, Ed. by W. Keifer, Special Issue of J. Raman Spectrosc. **31** (1/2) (2000).
19. D. Zeidler, S. Frey, K.-L. Kompa, and M. Motzkus, Phys. Rev. A **64**, 023420 (2001).
20. D. Zeidler, S. Frey, W. Wohlleben, *et al.*, J. Chem. Phys. **116**, 5231 (2002).
21. D. A. Akimov, A. M. Zheltikov, N. I. Koroteev, *et al.*, Laser Phys. **7**, 1242 (1997).
22. D. A. Akimov, A. B. Fedotov, N. I. Koroteev, *et al.*, Opt. Mem. Neural Netw. **6**, 31 (1997).
23. D. A. Akimov, N. I. Koroteev, S. A. Magnitskii, *et al.*, Jpn. J. Appl. Phys. **36**, 426 (1997).
24. B. Kohler, V. V. Yakovlev, J. Che, *et al.*, Phys. Rev. Lett. **74**, 3360 (1995).
25. C. J. Bardeen, Q. Wang, and C. V. Shank, Phys. Rev. Lett. **75**, 3410 (1995).

Translated by A. Zheltikov

JETP Letters, Vol. 78, No. 4, 2003, pp. 250–254. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 286–290.
Original English Text Copyright © 2003 by Fyodorov.

Induced vs. Spontaneous Breakdown of *S*-Matrix Unitarity: Probability of No Return in Quantum Chaotic and Disordered Systems¶

Y. V. Fyodorov

Department of Mathematical Sciences, Brunel University, Uxbridge UB83PH, United Kingdom
Petersburg Nuclear Physics Institute, Russian Academy of Sciences, Gatchina, Leningrad region, 188350 Russia

Received July 22, 2003

We investigate systematically sample-to-sample fluctuations of the probability τ of no return into a given entrance channel for wave scattering from disordered systems. For zero-dimensional ("quantum chaotic") and quasi-one-dimensional systems with broken time-reversal invariance, we derive explicit formulas for the distribution of τ and investigate particular cases. Finally, relating τ to violation of *S*-matrix unitarity induced by internal dissipation, we use the same quantity to identify the Anderson delocalization transition as the phenomenon of spontaneous breakdown of *S*-matrix unitarity. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 03.65.Nk; 05.45.Mt; 11.55.-m

Various aspects of chaotic wave scattering in the presence of absorption or internal losses have attracted considerable attention in recent years [1–10]. In the general case, a convenient framework for extracting the universal properties of the corresponding *S*-matrix is provided by the method of effective non-Hermitian Hamiltonian

$$\mathcal{H} = H - i\pi \sum_{a=1}^{M} W_a \otimes W_a^{\dagger},$$

in terms of which the energy-dependent element S_{ab} of the $M \times M$ scattering matrix $\hat{S}$ is expressed as

$$S_{ab} = \delta_{ab} - 2i\pi W_a^{\dagger} \frac{1}{\mathcal{E} - \mathcal{H}} W_b, \qquad (1)$$

see [11, 12] and references therein. Here, H stands for a self-adjoint Hamiltonian describing the closed counterpart of the disordered or chaotic system under consideration, $\mathcal{E}$ stands for the energy of incoming waves and the energy-independent vectors W_a, $a = 1, 2, \ldots, M$, contain matrix elements coupling the internal motion to one of open M channels. As is easy to verify, such a construction ensures the unitarity of the scattering matrix $S^{\dagger}S = \mathbf{1}_M$ provided the energy $\mathcal{E}$ takes only real values. When one allows the energy parameter to have a nonzero imaginary part $\epsilon = \mathrm{Im}\,\mathcal{E} > 0$, the *S*-matrix unitarity is immediately lost: $S^{\dagger}S < \mathbf{1}_M$. Physically, the parameter ϵ stands for uniform damping inside the system and is responsible for the losses of the outgoing flux of the particles as compared to the incoming flux. The balance between the two fluxes is precisely the physical mechanism behind the *S*-matrix unitarity.

In the present paper, we concentrate on the "probability of no return" (PNR), which is defined as the quantum-mechanical probability for a particle entering the system via a given channel a to never exit through the same channel. This quantity is well-defined for a given realization of disorder and will show sample-to-sample fluctuations, whose statistics we are going to study. In the case of no internal dissipation, PNR is the same as the probability to exit via any of the remaining channels, known as the transmission probability $\sum_{b \neq a} |S_{ab}|^2 \equiv 1 - |S_{aa}|^2$. For a system with absorption, the last equality is violated, and we keep the notation $\tau_a = 1 - |S_{aa}|^2$ for the PNR related to the reflection probability R_a in the same channel as $\tau_a = 1 - R_a$. In particular, if only a single open channel is attached to our disordered system and the boundaries are purely reflecting, then neglecting dissipation trivially results in $\tau \equiv 0$. The nontrivial statistics of τ then arise solely due to absorption, and for small absorption the PNR value $\tau \simeq 2\epsilon\tau_W$ [2, 3, 7], where τ_W is the so-called Wigner delay time intensively studied in recent years; see [3, 4, 11, 13–15] and references therein.

It is convenient to write explicitly the normalization of the channel vectors as $W_a^{\dagger}W_a = \gamma_a/\pi$ and to assume that different channel vectors are orthogonal: $W_a^{\dagger}W_b = 0$ for $a \neq b$. In fact, one should remember that the effective strength of every open channel is more appropriately characterized by the so-called "transmission coeffi-

¶ This article was submitted by the author in English.

cients" [12] (also known as the "sticking probabilities") $T_a = 1 - |\langle S_{aa}\rangle|^2$, related to bare couplings γ_a by

$$\frac{1}{T_a} = \frac{1}{2}(1 + g_a), \quad g_a = \frac{1}{2\pi\nu}(\gamma_a + \gamma_a^{-1}), \tag{2}$$

where ν is the mean spectral density. Here and below, the angular brackets mean the disorder-averaged value of the quantities. The two limiting cases $T_a = 1$ and $T_a = 0$ correspond to situations of perfectly coupled and decoupled (closed) channel a, respectively.

The starting point of our analysis is based on the following convenient representation for the diagonal elements of the scattering matrix:

$$S_{aa} = \frac{1 - iG_a}{1 + iG_a}, \quad G_a = \pi W_a^\dagger \frac{1}{\mathcal{E} - \mathcal{H}} W_a, \tag{3}$$

where

$$\mathcal{H}_a = H - i\pi \sum_{b \neq a}^{M} W_b \otimes W_b^\dagger.$$

In this way, we reduce our problem to investigating the statistics of the diagonal entries of the resolvent G_a of the "reduced-rank" non-Hermitian operator $\mathcal{H}_a$ independent of the vector W_a. In particular, for the single-channel case $M = 1$, the operator $\mathcal{H}_a$ will not contain the channel vector W at all and will therefore be self-adjoint: $\mathcal{H} = H$.

The statistics of the diagonal entries of the resolvent of a random self-adjoint Hamiltonian H describing the motion of a quantum particle in a static disorder were discussed in much detail by Mirlin and Fyodorov [16] in the framework of the supermatrix nonlinear σ model [17]. In particular, for the case of systems with broken time reversal invariance, they were able to find a very compact representation for the joint probability density $\mathcal{P}(u, v)$ of the real $u = \mathrm{Re}\, G_a$ and imaginary $v = \mathrm{Im}\, G_a$ parts of the quantity G_a, assuming normalization $W_a^\dagger W_a = 1/\pi$. Physically, the variable v is the most important and, as it is the local density of states, has enjoyed thorough investigations; see [18, 19] and references therein.

In fact, it is quite straightforward to incorporate the non-Hermitian part of the Hamiltonian $\mathcal{H}_a$ into the method, as was already partly done in [20], where the statistics of $\mathrm{Im}\, G_a$ were addressed as describing fluctuations of the photodissociation cross-section.

According to [16], the function $\mathcal{P}(u, v)$ is given, for the center of the spectrum $\mathrm{Re}\,\mathcal{E} = 0$, by

$$\mathcal{P}(u, v) = \frac{1}{4\pi v^2} P(x),$$

where x is the combination $x = (u^2 + v^2 + 1)/2v$ and the function $P(x)$ is given by

$$P(x) = \hat{\mathbf{L}} F(x), \quad F(x) = \int_{-1}^{1} \frac{d\lambda}{x - \lambda} \mathcal{F}(x, \lambda), \tag{4}$$

where we have introduced the (Legendre) operator

$$\hat{\mathbf{L}} = \frac{d}{dx}(x^2 - 1)\frac{d}{dx}.$$

The particular form of the function $\mathcal{F}(x, \lambda)$ depends crucially on the effective spatial dimension of the underlying disordered system and is, for example, quite different for zero-dimensional systems ("quantum chaos") and for "diffusive" extended quasi-one-dimensional or higher dimensional systems. We will give an explicit analysis of several physical possibilities later on in the paper.

Having at our disposal the expression for $\mathcal{P}(u, v)$, we can relate the PNR distribution $\mathcal{P}(\tau_a)$ to the function $P(x)$. After a set of algebraic transformations, we find the following attractively simple formula:

$$\mathcal{P}(\tau_a) = \frac{1}{\pi\tau_a^2}\int_0^\pi P[x(\tau_a, \theta)]\, d\theta,$$
$$x(\tau_a, \theta) = \left(\frac{2}{\tau_a} - 1\right)\left(\frac{2}{T_a} - 1\right) \tag{5}$$
$$+ 4\cos\theta \frac{\sqrt{(1 - \tau_a)(1 - T_a)}}{\tau_a T_a}.$$

Now we proceed with a separate analysis of a few physical situations possible in disordered systems. In all cases, we assume time reversal symmetry to be broken.

1. "Zero-dimensional" quantum chaotic system. We assume that the disordered region is coupled to M scattering channels characterized by effective coupling constants $g_1, \ldots, g_M$ (see Eq. (2)), with g_1 corresponding to the chosen entrance channel. The strength of uniform damping will be characterized by the parameter $\eta = 2\pi\epsilon/\Delta$, where Δ the mean level spacing generated by the Hermitian Hamiltonian H. According to the standard argumentation, H can be effectively replaced by a large $N \times N$ random Hermitian matrix taken from the Gaussian Unitary Ensemble; see, e.g., [11, 12]. Then, in the limit of large enough $N \gg M$, the function $\mathcal{F}(x, \lambda)$ depends on the remaining $M - 1$ coupling constants, as well as on the effective damping η, as [20]

$$\mathcal{F}(x, \lambda) = \prod_{a=2}^{M} \frac{g_a + \lambda}{g_a + x} e^{-\eta(x - \lambda)}. \tag{6}$$

The function $F(x)$ in Eq. (4) can be found in closed form for any M, since one gets, in fact, a simple recursion relating $F_M(x)$ to $F_{M-1}(x)$. Here we restrict our-

selves mainly to the cases of one and two open channels $M = 1, 2$:

$$F_1(x) = \int_{x-1}^{x+1} \frac{du}{u} e^{-\eta u},$$
$$F_2(x) = F_1(x) - 2\frac{e^{-\eta x}}{g_2 + x}\frac{\sinh\eta}{\eta}. \quad (7)$$

The distribution $\mathcal{P}(\tau)$ for $M = 1$ is then equal to (cf. [4])

$$\mathcal{P}(\tau) = \frac{2}{\tau^2} e^{-\eta A}\{I_0(\eta B)[\sinh\eta(\eta A - 1)\eta\cosh\eta] - \eta B\sinh\eta I_1(\eta B)\}, \quad (8)$$
$$A = \left(\frac{2}{\tau} - 1\right)\left(\frac{2}{T_1} - 1\right), \quad B = 4\frac{\sqrt{(1-\tau)(1-T_1)}}{\tau T_1},$$

where $I_0(z)$, $I_1(z)$ are the modified Bessel functions of the respective order. For particular case of perfectly coupled channel $T_1 = 1$, Eq. (8) reduces to the formula

$$\mathcal{P}_1(\tau) = \frac{1}{\tau^3} e^{-2\eta/\tau}[\tau(1 + 2\eta - e^{2\eta}) + 2\eta(e^{2\eta} - 1)] \quad (9)$$

derived earlier [2] with a very different method.

The function $\mathcal{P}(\tau)$ for $M = 2$ can be obtained straightforwardly, but the general formula is too long, and we restrict our discussion to a few particular cases. First of all, when dissipation is absent ($\eta = 0$), we recover the exact distribution of the transmission probability found earlier in [21, 22] by rather different methods. The next case to be considered is that of a lossy system coupled to two perfectly open channels $g_1 = g_2 = 1$:

$$\mathcal{P}_2(\tau) = \mathcal{P}_1(\tau) + \frac{1 - e^{-2\eta/\tau}}{\eta}\left[\frac{1}{2} + \frac{\eta}{\tau} + \frac{2\eta^2}{\tau^2} - \frac{2\eta^2}{\tau^3}\right], \quad (10)$$

where $\mathcal{P}_1(\tau)$ is given in Eq. (9). In fact, it is not difficult to find a similar recursive formula relating $\mathcal{P}(\tau)$ for M perfectly coupled channels to the same function for $M - 1$ perfect channels. We do not give that formula, apart from the simplest case of no dissipation:

$$\mathcal{P}_M(\tau) = \mathcal{P}_{M-1}(\tau) + [(M-1)\tau^{M-2} - (M-2)\tau^{M-3}],$$

which immediately yields $\mathcal{P}_M(\tau) = (M-1)\tau^{M-2}$, $M \geq 1$. This formula (as well as its counterpart $\mathcal{P}(\tau) \propto \tau^{\frac{M-3}{2}}$ for preserved time reversal invariance) in fact follows from the known distribution of 1×1 subunitary block of random unitary scattering matrices; see [23].

2. Quasi-1D systems. Consider a single channel attached to one edge of a piece of quasi-one-dimensional disordered metal of length L, with the opposite edge being in contact with a perfectly conducting lead of very many channels. When the internal dissipation is absent, the function $F(x)$ was found by Mirlin [24]:

$$F(x) = \ln\frac{x+1}{x-1} - 2\int_0^\infty \frac{dkk}{1+k^2}\tanh\left(\frac{\pi k}{2}\right)P_{-\frac{1}{2}+\frac{ik}{2}}(x)e^{-t(k^2+1)/4}, \quad (11)$$

where the dimensionless parameter $t = L/\xi$ is the sample length L measured in units of the localization lengths ξ. The (real) functions $P_\nu(x)$, $\nu = -1/2 + ik/2$ are known as conical functions and represent a special case of Legendre functions. As such they satisfy $\hat{\mathbf{L}}P_\nu(x) = \nu(\nu+1)P_\nu(x)$. This observation immediately yields the following expression for the PNR distribution:

$$\mathcal{P}(\tau) = \frac{1}{2\tau^2} I\left(t; \frac{2}{\tau} - 1\right), \quad (12)$$

$$I(t; x) = \int_0^\infty dkk\tanh\left(\frac{\pi k}{2}\right)P_{-\frac{1}{2}+\frac{ik}{2}}(x)e^{-t(k^2+1)/4}, \quad (13)$$

where we assumed, for simplicity, that the selected single channel is perfectly coupled to the scattering medium. Surprisingly, this distribution is practically the same as the distribution of the reflection coefficient from a piece of strictly one-dimensional medium obtained long ago in the framework of the Berezinskii technique [25]. In particular, for any value of the parameter t, the distribution displays a log-normal far tail corresponding to very small PNR values $\tau \longrightarrow 0$. To find it for $t \ll 1$, one needs an asymptotic of the conical functions for large arguments $x \gg 1$, which we borrow from Eq. (50) of [24]. A calculation very similar to that presented in [24] yields

$$\mathcal{P}(\tau \ll 1) \simeq \frac{1}{2\sqrt{2}t}\frac{\sqrt{-\ln\tau}}{\tau}\exp\left\{-\frac{1}{4t}\ln^2\tau\right\}, \quad t \ll |\ln\tau|. \quad (14)$$

This log-normal tail is related to the presence of the anomalously localized states [26]. In the opposite case of very long samples $t \gg 1$, the PNR values are exponentially small due to the localization phenomenon, and the distribution is purely log-normal:

$$\mathcal{P}(\tau) \simeq \frac{1}{2\tau}\frac{1}{\sqrt{\pi t}}\exp\left\{-\frac{1}{4t}(t + \ln\tau)^2\right\}, \quad t \sim |\ln\tau| \gg 1. \quad (15)$$

Let us turn our attention now to the case of a quasi-1D disordered sample with a nonvanishing internal dissipation $\epsilon > 0$, assuming the second edge of the sample

to be impenetrable for waves. The scaling physical parameter controlling the role of dissipation is then given by [18] $\delta = \pi\rho\epsilon\xi$, with ρ standing for the mean spectral density. This is just the dissipation ε measured in units of the mean level spacing for a sample whose length is ξ. The most interesting regime is that of small $\delta \ll 1$. In that limit, the function $\mathcal{F}(x, \lambda)$ turned out to be independent of λ, whereas the x dependence persists in a form of the scaling combination $y = 2\delta x$, i.e., $\mathcal{F}(x, \lambda) \longrightarrow \tilde{F}(2\delta x)$. This implies that the relevant values of parameters are $2/\tau \simeq x \sim \delta^{-1} \gg 1$. The latter condition immediately results in the formula $F(x) \longrightarrow 4\delta\tilde{F}(y)/y$ and also converts the Legendre operator $\hat{\mathbf{L}}$ to

$$\hat{\mathbf{L}}_y = \frac{d}{dy} y^2 \frac{d}{dy}.$$

Let us note that the emerging PNR distribution yields, in fact, the distribution of the Wigner delay time via the relation $\tau_W = 2\pi\rho\xi/y$.

The expression for the function $\tilde{F}(y)$ is known explicitly [27]: $\tilde{F}(y) = \tilde{F}_\infty(y) + \tilde{F}_t(y)$, where

$$\tilde{F}_t(y) = \frac{4}{\pi} y^{\frac{1}{2}} \int_0^\infty \frac{dk\,k}{1+k^2} \sinh\frac{\pi k}{2} K_{ik}\left(2y^{\frac{1}{2}}\right) e^{-t(k^2+1)/4} \quad (16)$$

and $\tilde{F}_\infty(y) = 2y^{\frac{1}{2}} K_1\left(2y^{\frac{1}{2}}\right)$, with $K_\nu(u)$ standing for the Macdonald function.

For the case of a very short ($t \ll 1$) sample, the function $\tilde{F}(y)$ is known to be approximated by $\exp(-ty)$ [27]. A simple calculation then yields the distribution $\mathcal{P}(\tau) = [(4t\delta)^2/\tau^3]\exp\{-4t\delta/\tau\}$. Realizing that $2t\delta \equiv \eta$, we see that the distribution coincides with the weak absorption limit of Eq. (9). As expected, the same distribution follows from that of the Wigner delay time [11].

In the opposite limit of very long samples, $t \longrightarrow \infty$, only the first term survives, and by noticing that $\hat{\mathbf{L}}_y[\tilde{F}_\infty(y)/y] = \tilde{F}_\infty(y)$ we find the corresponding PNR distribution:

$$\mathcal{P}(\tau) = \frac{16}{\delta}\left(\frac{\delta}{\tau}\right)^{5/2} K_1(4\sqrt{\delta/\tau}). \quad (17)$$

Although the typical value of τ is on the order of δ, the moments ($\langle\tau^m\rangle$) do not exist for $m \geq 1$ because of the power law tail $\mathcal{P}(\tau \gg \delta) \propto \tau^{-2}$. A similar tail was found in the distribution of the total reflection coefficient from a multichannel long disordered 1D sample in [9] and is also typical for the Wigner delay time distribution in purely 1D system [15]. Negative moments of τ are equal to $\langle\tau^{-k}\rangle = (4\delta)^{-k}k!(k+1)!$. Note that they differ from the corresponding moments in a purely 1D system [15] by an extra factorial factor $(k+1)!$, reminiscent of similar relations between other quantities in 1D and quasi-1D [18].

Finally, in the case of strong absorption $\delta \gg 1$ in a long wire $t \longrightarrow \infty$, the function $\mathcal{F}(x, \lambda) = \exp\{-\sqrt{\delta}(x-\lambda)\}$ [17] and the resulting distribution $\mathcal{P}(\tau)$ coincide with that given by Eq. (9), with η replaced by $\sqrt{\delta}$.

3. Behavior at the Anderson transition. Let us briefly discuss the possible qualitative behavior of the PNR τ in a scattering system formed by a single perfect channel attached to a d-dimensional disordered sample in the vicinity of the point of the Anderson delocalization transition α_c (the mobility edge). Here, we denote by α an effective parameter that controls the transition in the infinite sample, with states being localized (extended) for $\alpha > \alpha_c$ (respectively, $\alpha < \alpha_c$).

Our arguments are based on a picture of the transition as described in terms of a functional order parameter developed in detail in [16]; see also earlier results in [17] and [28]. For a sample of finite size L, the PNR is a function of three parameters: ϵ, L, α. According to the suggested scenario, the behavior of the function $\mathcal{F}(x, \lambda)$ in the insulating phase in the weak absorption limit $\delta \propto \epsilon\xi^d \longrightarrow 0$ is expected to be reminiscent of that described above for the one-dimensional case, i.e., $\mathcal{F}(x, \lambda) \longrightarrow \tilde{F}(2\delta x)$, and the function $\tilde{F}(y)$ decays to zero for $y \gg 1$. Then it is natural to expect that all negative PNR moments in the infinite volume limit $L \longrightarrow \infty$ are to scale as $\langle\tau^{-k}\rangle \sim \epsilon^{-k}\xi^{-dk}$, where ξ is the localization length diverging in the vicinity of the mobility edge.

In contrast, in the delocalized phase, the function $\mathcal{F}(x, \lambda)$ is expected to remain a nontrivial function of both x and λ even when $\epsilon \longrightarrow 0$, provided the latter limit is taken after the infinite volume limit $L \longrightarrow \infty$. This should immediately result in a finite-width distribution $\mathcal{P}(\tau)$ of the PNR. From this point of view, the Anderson transition acquires a natural interpretation as the phenomenon of spontaneous breakdown of *S*-matrix unitarity. As long as $\alpha \longrightarrow \alpha_c$, the widths of the distribution and properly defined (negative) PNR moments should vanish, with some set of critical exponents.

If, however, we take the limit $\epsilon \longrightarrow 0$ first, then for $\alpha < \alpha_c$ PNR in a large but finite sample should scale with the system size L as $\langle\tau^{-k}\rangle \sim C(\alpha)\epsilon^{-k}L^{-dk}$, where $C(\alpha)$ is expected to diverge when $\alpha \longrightarrow \alpha_c$. In some sense, the behavior of the negative moments of the Wigner delay time defined as $\tau_W = \lim_{\epsilon\to 0}\tau(\epsilon, L, \alpha)/2\epsilon$ is reminiscent of that for the inverse participation ratio [26]. This analogy suggests a possibility for anomalous scaling $\langle\tau_W^{-k}\rangle \sim L^{-dr_k}$ with $r_k \neq k$ at the mobility edge

$\alpha = \alpha_c$, which would then reflect the underlying multifractality of the wavefunctions.

It will be very interesting to perform a detailed numerical analysis of PNR and Wigner delay times for realistic and well-controlled models of scattering from disordered systems, e.g., quantum graphs [29] or models used in [13], and to verify the suggested picture qualitatively and quantitatively in various regimes. The statistics of PNR should be also of experimental accessibility in microwave resonator experiments; see, e.g., [6, 7] and references therein.

The author is obliged to Y. Imry and U. Kuhl for useful discussions and to T. Kottos, A. Mirlin, and D. Savin for valuable comments. Financial support by the Brunel University VC grant is gratefully acknowledged.

REFERENCES

1. E. Kogan, P. A. Mello, and H. Liqun, Phys. Rev. E **61**, R17 (2000).
2. C. W. J. Beenakker and P. W. Brouwer, Physica E (Amsterdam) **9**, 463 (2001).
3. S. A. Ramakrishna and N. Kumar, Phys. Rev. B **61**, 3163 (2000).
4. D. V. Savin and H.-J. Sommers, cond-mat/0303083.
5. I. Rozhkov, Y. V. Fyodorov, and R. L. Weaver, Phys. Rev. E **68**, 016204 (2003).
6. R. Schäfer, T. Gorin, T. H. Seligman, and H. J. Stockmann, J. Phys. A: Math. Gen. **36**, 3289 (2003); R. A. Mendez-Sanchez *et al.*, cond-mat/0305090.
7. E. Doron, U. Smilansky, and A. Frenkel, Phys. Rev. Lett. **65**, 3072 (1990).
8. O. Entin-Wohlman, A. Aharony, Y. Imry, *et al.*, Phys. Rev. Lett. **88**, 166801 (2002).
9. T. Sh. Misirpashaev and C. W. J. Beenakker, JETP Lett. **64**, 319 (1996); N. A. Bruce and J. T. Chalker, J. Phys. A: Math. Gen. **29**, 3761 (1996).
10. V. Freilikher, M. Pustilnik, and I. Yurkevich, Phys. Rev. B **50**, 6017 (1994); P. Pradhan and N. Kumar, Phys. Rev. B **50**, 9644 (1994); M. Titov and C. W. J. Beenakker, Phys. Rev. Lett. **85**, 3388 (2000).
11. Y. V. Fyodorov and H.-J. Sommers, J. Math. Phys. **38**, 1918 (1997).
12. J. J. M. Verbaarschot, H. A. Weidenmueller, and M. R. Zimbauer, Phys. Rep. **129**, 367 (1985).
13. T. Kottos and M. Weiss, Phys. Rev. Lett. **89**, 056401 (2002); A. Ossipov, Tsampikos Kottos, and T. Geisel, Phys. Rev. B **61**, 11411 (2000); Europhys. Lett. **62**, 719 (2003).
14. P. W. Brouwer, K. M. Frahm, and C. W. J. Beenakker, Phys. Rev. Lett. **78**, 4737 (1997).
15. C. Texier and A. Comtet, Phys. Rev. Lett. **82**, 4220 (1999); S. K. Joshi, A. K. Gupta, and A. M. Jayannavar, Phys. Rev. B **58**, 1092 (1998); J. Heinrichs, Phys. Rev. B **65**, 075112 (2002).
16. A. D. Mirlin and Y. V. Fyodorov, Phys. Rev. Lett. **72**, 526 (1994); J. Phys. I **4**, 655 (1994).
17. K. B. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge Univ. Press, Cambridge, 1997).
18. A. D. Mirlin and Y. V. Fyodorov, Europhys. Lett. **25**, 669 (1994); C. W. J. Beenakker, Phys. Rev. B **50**, 15170 (1994).
19. H. Schomerus, M. Titov, P. W. Brouwer, and C. W. J. Beenakker, Phys. Rev. B **65**, 121101 (2002); M. Titov and H. Schomerus, Phys. Rev. B **67**, 024410 (2003).
20. Y. V. Fyodorov and Y. Alhassid, Phys. Rev. A **58**, R3375 (1998).
21. V. N. Prigodin, K. B. Efetov, and S. Iida, Phys. Rev. B **51**, 17223 (1995).
22. P. W. Brouwer and C. W. J. Beenakker, Phys. Rev. B **50**, 11263 (1994).
23. Y. V. Fyodorov and H.-J. Sommers, J. Phys. A: Math. Gen. **36**, 3303 (2003); see Eq. (69) of the paper; W. A. Friedman and P. A. Mello, J. Phys. A: Math. Gen. **18**, 425 (1985).
24. A. D. Mirlin, Phys. Rev. B **53**, 1186 (1996).
25. A. A. Abrikosov, Solid State Commun. **37**, 997 (1981).
26. A. D. Mirlin, Phys. Rep. **326**, 259 (2000).
27. Y. V. Fyodorov and A. D. Mirlin, Int. J. Mod. Phys. B **8**, 3795 (1994).
28. M. R. Zirnbauer, Phys. Rev. B **34**, 6394 (1986); A. D. Mirlin and Y. V. Fyodorov, Nucl. Phys. B **366**, 507 (1991).
29. T. Kottos and U. Smilansky, Phys. Rev. Lett. **85**, 968 (2000); J. Phys. A: Math. Gen. **36**, 3501 (2003).

JETP Letters, Vol. 78, No. 4, 2003, pp. 255–258. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 78, No. 4, 2003, pp. 291–294.
Original Russian Text Copyright © 2003 by Freĭnkman, Eletskiĭ, Zaĭtsev.

Laser Ion Beam Formation for Nanotechnologies

B. G. Freĭnkman[1], A. V. Eletskiĭ[2], and S. I. Zaĭtsev[1]

[1]*Institute for Problems of Microelectronic Technologies, Russian Academy of Sciences, Chernogolovka, Moscow region, Russia*

[2]*Kurchatov Institute, Russian State Scientific Center, Moscow, 123182 Russia*

Received June 19, 2003; in final form, July 14, 2003

It is suggested to generate cold ion beams by laser collimation and subsequent laser ionization of a primary atomic beam. The primary beam, formed by a standard method, is collimated through transverse cooling by resonance laser radiation. Laser radiation is also used for the multistep ionization of atoms in the collimated beam. Advantages of the proposed method are a low scatter of the initial ion energy (below 10^{-1} eV) and a high emittance in the region of the virtual source (~10^{-6} cm rad at a beam current on the level of microamperes). The high monochromaticity of the obtained ion beam allows the chromatic aberration effect to be significantly suppressed, which implies good prospects for using such sources in ion beam lithography. The proposed method also allows the spectrum of elements used in ion beam sources to be expanded, which is an independent technological advantage. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 32.80.Fb; 41.75.Ak; 39.10.+j; 07.77.Gx

One of the main directions in the development of rapidly expanding nanotechnologies is related to the need for increasing the resolution of lithographic processes to a characteristic scale below 100 nm [1]. For the ion beam lithography the progress in this direction is limited primarily by the chromatic aberration effect related to the scatter of initial ion energies (within ~5 eV for liquid metal sources and ~2 eV for microwave discharge sources) [2].

In connection with recent developments in the field of laser collimated and focused atomic beams (see, e.g., [3–5]), there is a growing interest in the commercial use of such beams. In this case, principal limitations of the spatial resolution are related only to purely quantum effects: the resolution threshold as determined by the de Broglie wavelength is on the order of 0.1 nm. However, in practical applications, the possibilities of atomic beam lithography are still limited by relatively low energies of atomic beams, which is insufficient for conducting the processes of resist exposure, implantation, etching, etc., on the required resolution level.

The natural way of solving this problem would be to form an ion beam from a laser collimated atomic beam through laser ionization, as is done in the laser ionization spectroscopy and laser isotope separation technology [6–10]. Using this approach, it is possible to provide for a significant increase in the ion beam energy, while retaining a high degree of collimation, by using electrostatic collimating lenses. This study is aimed at an analysis of the possibility of practical implementation of such a combined method for obtaining ion beams for nanotechnologies.

Atomic beam formation. We suggest using a standard source for creating a beam of neutral atoms, an example being offered by an oven producing a saturated vapor of a substance. The flux of neutrals in the beam is determined by the saturated vapor pressure at a given oven temperature and by the transmission of the output channel [11]. By using metals possessing sufficiently high volatility, it is easy to obtain output atomic fluxes in the beams on the order of 10^{13}–10^{15} cm^{-2} s^{-1}. For example, a liquid metal source with gallium provides for this atomic flux density at a temperature of about 1100–1200 K.

By using a diaphragm, the angular divergence of a beam can be reduced to a sufficiently small level at the expense of some loss in the beam intensity. However, even in the case of incomplete ionization, the atomic beam intensity can still provide for an ion beam current of 10^{-8}–10^{-6} A, which significantly exceeds the output current of a liquid-metal ion source (~10^{-9} A). An example of successful solution is offered by [5], where a beam of Cs atoms with a diameter of $d_a \sim 1.4$ mm, an intensity of $j_a \sim 10^{12}$ cm^{-2} s^{-1}, and an angular divergence of $r'_a \sim 0.4$ mrad was obtained using a metal vapor at $T \sim 410$ K.

Laser collimation of atomic beams. The laser collimation of atomic beams is based on the resonance interaction of such a beam with an orthogonal laser beam [3–6]. In the classical description, the force acting upon a two-level atom absorbing a resonance laser radiation is proportional to the differential momentum transfer as a result of the laser photon absorption and the spontaneous and stimulated photon emission. Owing to the coherence of laser radiation, the momentum transfer during the absorption is equal in magnitude and opposite in direction to that during the stimu-

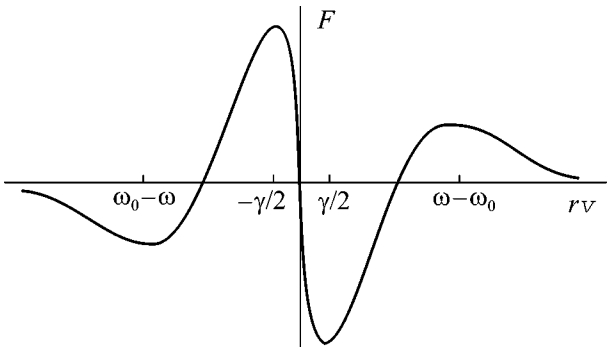


Fig. 1. A plot of the radiative friction force versus atomic velocity according to formula (2).

lated emission, while the direction of spontaneous emission is arbitrary. Therefore, the average force acting upon an atom is directed along the laser beam and proportional to the difference between the photon and atomic momenta and to the probability of spontaneous emission from the upper excited level. The momentum transfer depends on the laser frequency detuning relative to the center of the atomic absorption line with allowance of the Doppler effect, which is determined by the thermal motion of atoms.

The general expression for the radiative friction force with which a resonance laser radiation acts upon an atom is as follows [12]:

$$\mathbf{F} = \eta\mathbf{k}\gamma\left[\frac{V_0^2}{(\omega-\omega_0-k\mathbf{v})^2+\gamma^2/4+2V_0^2}\right], \qquad (1)$$

where ω and $\mathbf{k}$ are the frequency and the wave vector of the laser radiation; ω_0 and γ are the frequency and the radiative lifetime of the atomic transition, respectively; $\mathbf{v}$ is the atomic velocity; and V_0 is the Rabi frequency ($2V_0^2$ is the dipole potential of the atomic transition, proportional to the laser radiation intensity).

For collimating an atomic beam, it is necessary to use at least two opposite laser beams orthogonal to the longitudinal velocity of the atomic beam. By selecting a proper laser radiation frequency, it is possible to provide for the effective interaction of atoms only with the laser beam opposite to the transverse atomic velocity. In this case, an expression for the radiative friction force with an arbitrary V_0 is as follows [12]:

$$\mathbf{F} = \eta\mathbf{k}\gamma\left[\frac{V_0^2}{(\omega-\omega_0-k\mathbf{v})^2+\gamma^2/4+2V_0^2}\right.$$
$$-\frac{V_0^2}{(\omega-\omega_0+k\mathbf{v})^2+\gamma^2/4+2V_0^2} \qquad (2)$$
$$\left.-\left(\frac{V_0}{\omega-\omega_0}\right)^6\frac{k\mathbf{v}(\omega-\omega_0)}{4(k\mathbf{v})^2+\gamma^2}\right].$$

In a sufficiently weak field, the main role (irrespective of the atomic velocity) is played by the first two terms. For $\omega-\omega_0<0$, this force slows down the atoms, and for $\omega-\omega_0>0$, it accelerates the atoms with the transverse velocity components opposite to the laser beam direction. When the field is not too weak ($V_0^2 > (\omega-\omega_0)\gamma$), the last term begins to predominate in the region of moderate velocities ($v \le \gamma/k$). The radiative friction force is zero when $|v| = (\gamma/2k)\sqrt{[V_0/(\omega-\omega_0)\gamma]^4-1}$ and exhibits two extrema in the region of large velocities, at $kv \sim \pm(\omega-\omega_0)$, and two others in the region of small velocities, at $kv \sim \pm\gamma/2$. The plot of this force as a function of the atomic velocity is depicted in Fig. 1.

Using expressions (1) and (2), it is possible to evaluate the possibilities of radiative collimation of atomic beams with respect to the resonance value of the friction force and the resonance width. For the visible radiation ($\gamma \sim 10^7$ Hz) and thermal atomic velocities ($v \sim 10^4$–10^5 cm/s), a change in the velocity in the region of effective deceleration (cooling) amounts to $\Delta v \sim 10^2$–10^3 cm/s during a time of $\tau \sim 10^{-5}$–10^{-4} s (for a laser radiation intensity below 0.1 W/cm^2), which corresponds to a longitudinal size of the irradiated region about 0.1–1 cm. Thus, in order to ensure effective radiative collimation of thermal atomic beams, it is necessary to reduce the initial beam divergence to within 10^{-1}–10^{-2} rad.

It should be noted that even relatively low laser radiation intensity ($\gg$1 W/cm^2) provides for saturation of the resonance excited state in the course of collimation. This circumstance considerably simplifies the problem of subsequent multistep laser ionization of atoms in the collimated beam.

Laser ionization of atomic beams. The choice of a particular scheme for the laser ionization of an atomic beam is determined by the atom type and the beam intensity. In selecting the scheme, it would be expedient using the experience gained from the development of atomic vapor laser isotope separation (AVLIS) technology [7–10, 13, 14]. The spectrum of elements for which the isotope selective laser ionization was successfully implemented is sufficiently broad, including elements such as Pt [7], U [8, 9], Ti [10], Li [13], Ca [14], etc.

The ionization of atoms in a collimated beam is performed according to a multistep scheme employing frequency tunable lasers. In the first step, it is possible to use the resonance radiation of the same laser that was used for the laser beam collimation. Under the conditions of saturation, the populations of the ground (N_0) and resonancely excited (N_1) states are related as

$$N_0/g_0 = N_1/g_1, \qquad (3)$$

where g_0 and g_1 are the statistical weights of the corresponding states.

Subsequent ionization of the resonancely excited atoms can be performed in one or two steps, the limiting stage being determined by the process of photoionization of an excited atom. For estimating the required laser radiation intensity, let us assume that ionization has to take place within a sufficiently narrow spatial region, where variation of the electric field potential would be relatively small. Let the size l of this region be such that

$$I \geq \eta \omega_{pi} v/\sigma_{pi} l, \tag{4}$$

where I is the laser radiation intensity, $\hbar\omega_{pi}$ is the laser quantum energy, v is the transverse atomic velocity in the beam, and σ_{pi} is the photoionizaton cross section of the excited atom. The latter parameter can be estimated using the Kramers formula [15] describing the photoionizaton cross section of a highly excited hydrogen atom with an effective principal quantum number $n \gg 1$:

$$\sigma_{pi} = \frac{16\pi}{3\sqrt{3}} \frac{me^{10}}{cn^5 \eta^6 \omega^3 l}. \tag{5}$$

Substituting this estimate and characteristic values of $\hbar\omega \approx 2$ eV, $l \sim 0.01$ cm, and $v \sim 10^5$ cm/s into condition (4), we conclude that the effective photoionizaton of the excited atoms in the region of preset size l is ensured provided that $I \gg 100n^5$ W/cm^2. This condition is readily satisfied for not too large quantum numbers ($n \sim$ 3–5) by using focused laser beams.

Estimating the main parameters of the ion source. Figure 2 shows a schematic diagram of the setup implementing the proposed method of laser collimation and ionization of atomic beams. Let us estimate the main parameters of this system, proceeding from the atomic beam parameters achieved in [5]. Using the laser ionization of an atomic beam, followed by focusing the ion beam with an immersion lens of spherical capacitor type, it is possible to obtain a virtual ion source with a current of ~10^7 A in the focal plane of this lens. Taking into the parameters of this lens [16], we obtain

$$M = \frac{R_a}{R_c - 2R_a}; \quad R_i = R_c M, \tag{6}$$

where M is the linear magnification scale, R_c is the radius of the external sphere, R_a is the radius of the inner sphere, and R_i is the radius of the image surface. According to the Helmholtz–Lagrange relation [17],

$$MG = \sqrt{U_0/U_i} \tag{7}$$

(where G is the angular magnification scale and U_0, U_i are the potentials in the space of object and image, respectively), the source radius is given by the formula

$$r_i = r'_a R_c \sqrt{U_0/U_i}, \tag{8}$$

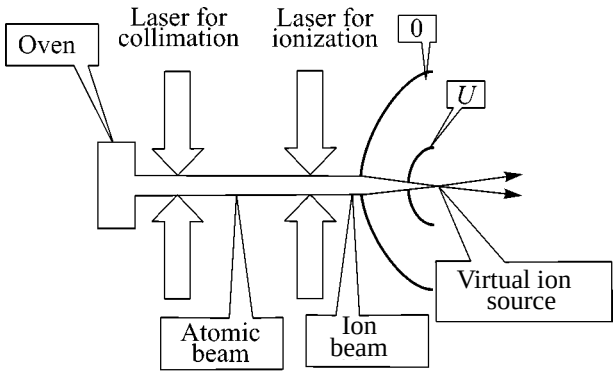


Fig. 2. A schematic diagram of the proposed laser ion beam source.

and the source emittance can be determined as

$$r_i G = r'_a \frac{R_c}{M} \frac{U_0}{U_i}. \tag{9}$$

For $M = 0.1$, $R_c = 50$ mm, $U_0 = 0.1$ eV, and $U_i = 1$ kV, these formulas yield $r_i \sim 1.2 \times 10^{-5}$ cm and $r_i G \sim 7 \times 10^{-7}$ cm rad. The ion energy scatter, determined only by the oven temperature, amounts to ~0.035 eV.

As a particular example, let us consider the formation of a beam of gallium ions (frequently employed in the ion beam etching technology). Gallium atom is characterized by an ionization potential of 6 eV and has a strong resonance absorption line corresponding to the transition $4p^2P_{1/2} \longrightarrow 5s^2S_{1/2}$ with $\lambda = 403$ nm. The probability of spontaneous emission for this transition is 5×10^7 s^{-1} [18]. This yields the following estimate of the laser intensity required for saturation under the condition of dominating Doppler broadening of the resonance transition: $I_s = 8\pi\langle v\rangle\hbar\omega/\lambda^3 \sim 15$ W/cm^2 (here, $\langle v\rangle$ is the average thermal velocity of atoms). As can be seen, the existing dye lasers readily provide for saturation of the transition under consideration. For example, an effective source of coherent radiation (with an efficiency of up to 11%) tunable in a wavelength range from 299 to 416 nm is provided by a laser using a solution of 4,4-diphenylstilbene in dioxane, pumped by a pulsed XeCl excimer laser with $\lambda = 308$ nm [19].

The length L_c of the atomic beam collimation zone is determined by the requirement that the transverse momentum of atoms related to their thermal velocity scatter has to be suppressed by the resonance laser radiation employed. Under the condition of saturation of the resonance transition, this requirement is independent of the laser radiation intensity and is expressed by the relation

$$L_c \sim M\langle v\rangle \alpha/\hbar k\tau, \tag{10}$$

where M is the atomic mass and α is the initial angular divergence of the atomic beam collimated by a dia-

phragm. For $\alpha = 0.01$, relation (10) yields an estimate of $L_c \sim 1.5$ cm.

Ionization of the resonance-excited Ga atoms can be also performed by the laser used in the collimation stage. However, in this case the laser intensity must be not lower than $\sim 10^6$ W/cm^2. This level can be provided by focusing the laser beam in the ionization zone. Since the intensity of excitation is much lower as compared to that required for the ionization, no ionization takes place in the excitation zone.

Thus, the proposed ion beam source is comparable with liquid metal sources and microwave discharge sources [2] with respect to the ion beam current and emittance, while exceeding both alternative sources with respect to the initial ion energy scatter parameter. The proposed method virtually completely suppresses chromatic aberration of the ion-optical system, which is one the main factors determining limiting resolution of the ion beam lithography process. Another important advantage of the new source is the possibility of virtually inertialess control of the ion beam current achieved by varying the intensity of the ionizing laser radiation, although this factor cannot be changed arbitrarily. However, in the case of metal atomic beams, the existing frequency tunable lasers can ensure the effective multistep ionization for atoms of any type. We believe that the above features make the proposed approach to the ion beam formation a promising solution for nanotechnologies.

REFERENCES

1. K. A. Valiev, *Physics of Submicron Lithography* (Nauka, Moscow, 1990).
2. *The Physics and Technology of Ion Sources*, Ed. by I. G. Brown (Wiley, New York, 1989; Mir, Moscow, 1998).
3. S. Chu, C. N. Cohen-Tannoundji, and W. D. Fillips, Usp. Fiz. Nauk **169**, 274 (1999).
4. G. Timp, R. E. Behringer, D. M. Tennant, *et al.*, Phys. Rev. Lett. **69**, 1636 (1992).
5. F. Lison, H.-J. Adams, D. Haubrich, *et al.*, Appl. Phys. B **65**, 419 (1997).
6. N. B. Delone, *Interaction of Laser Radiation with Matter* (Nauka, Moscow, 1989).
7. P. Peuser, F. Ames, J. Bohn, *et al.*, Appl. Phys. B **38**, 249 (1985).
8. V. K. Mago, B. Lal, A. K. Ray, *et al.*, J. Phys. B **20**, 6021 (1987).
9. H. D. V. Böhm, W. Michaelis, and C. Weitkamp, Opt. Commun. **26**, 177 (1978).
10. Y. Mariyama, Y. Suzuki, T. Arisawa, and K. Shiba, Appl. Phys. B **44**, 163 (1987).
11. T. A. Voronchev and V. D. Sobolev, *Physical Fundamentals of Electrovacuum Technology* (Vysshaya Shkola, Moscow, 1967).
12. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991).
13. T. Arisawa, Y. Mariyama, Y. Suzuki, and K. Shiba, Appl. Phys. B **28**, 73 (1982).
14. U. Brinkmann, W. Hartig, H. Telle, and H. Walther, Appl. Phys. **5**, 109 (1974).
15. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Relativistic Quantum Theory* (Nauka, Moscow, 1968), Vol. 1.
16. P. Schange, H. Bruining, and J. C. Francken, Philips Res. Rep. **7**, 119 (1952).
17. L. A. Baranova and S. Ya. Yavor, *Electrostatic Electron Lenses* (Nauka, Moscow, 1986).
18. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions*, 2nd ed. (Énergoatomizdat, Moscow, 1986; Springer, Berlin, 1985).
19. A. A. Malyutin, in *Handbook of Physical Quantities*, Ed. by I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, N.Y., 1997), Chap. 34.

Translated by P. Pozdeev