JETP Letters, Vol. 77, No. 4, 2003, pp. 151–153. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 181–183.

Cosmic Rays of Extremely High Energies and Pulsars

A. A. Mikhaĭlov

Institute of Cosmophysical Research and Aeronomy, Yakut Research Center, Siberian Division, Russian Academy of Sciences, pr. Lenina 31, Yakutsk, 677891 Russia

e-mail: mikhailov@ikfia.ysn.ru

Received December 31, 2002

Extensive air showers with energy above 4×10^{19} eV that were detected at the Yakutsk array from 1974 to 2001 are analyzed. The directions of their arrival are found to correlate with pulsars located in the direction of the Orion Arm of the Galaxy. The origin of cosmic rays is discussed. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 98.70.Sa; 97.60.Gb; 96.40.Pq

The origin of cosmic rays is one of the fundamental problems in high-energy astrophysics. At present, numerous hypotheses about the origin of cosmic rays with energy $E > 4 \times 10^{19}$ eV exist: from their production due to the decay of superheavy relict particles, which have mass $>10^{21}$ eV and lifetime exceeding the universe's age and belong to the dark matter of the Galactic halo [1], to their arrival from the nuclei of active galaxies [2, 3], radio galaxies [4], etc. These hypotheses must answer the following two questions.

(i) If particles are produced in the Galaxy or nearby galaxies, how do they form an isotropic distribution on the celestial sphere?

(ii) If particles are produced in galaxies more than 40 Mpc away, why is there no bend in the energy spectrum above 6×10^{19} eV [5] (according to calculations made by Greisen [6] and Zatsepin and Kuzmin [7], the spectrum of particles must have a bend at 6×10^{19} eV due to the interaction of particles with the microwave relict radiation with a temperature of 2.7 K)?

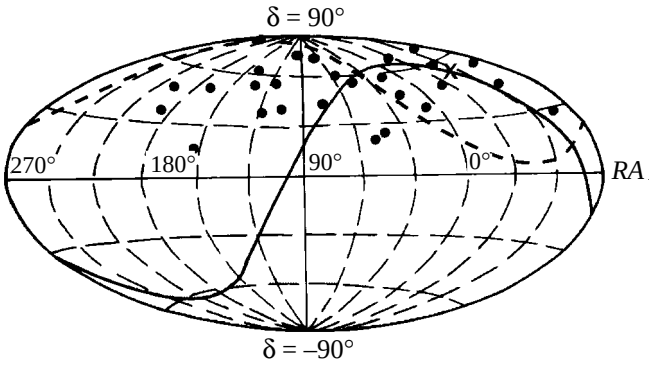
Unfortunately, none of the above hypotheses can answer these questions.

In this work, we analyze the distribution of showers with energy $E > 4 \times 10^{19}$ eV that were detected at the Yakutsk EAS array from 1974 to 2001, when 27 showers, whose arrival axes fell within the array perimeter, were detected. One shower had energy $E > 10^{20}$ eV (1.2×10^{20} eV). The energy of the primary particles and spatial angles were determined with accuracies ~30% and ≤3°, respectively.

The figure shows the distribution of the showers in the map of the celestial sphere in the second system of equatorial coordinates δ (declination) and *RA* (right ascension). As is seen, the distribution is almost isotropic, and a cluster is observed only at δ ~ 27° and *RA* ~ 48°. The probability of random formation of the cluster among 27 uniformly distributed showers is $P \sim 1$.

Further, we determined correlation between the arrival directions of showers and pulsars [8]. To this end, we took the following directions: (i) over the entire celestial sphere region visible by the array and (ii) along the field lines of the large-scale regular magnetic field in the direction of the Orion Arm within a cone with angle <45° (dashed line in figure) from the field-line axis with the galactic coordinates $b = 0°$ and $l = 90°$ (marked by *X* in figure) or δ = 48.1° and *RA* = 317.5°. This direction was chosen because the magnetic field minimally deviates particles moving along field lines, and the probability of correlation between shower arrival directions and pulsars increases.

We calculated the angular distances between the arrival direction of each shower and all pulsars and determined the number $N(\vartheta)$ of showers observed within the angle ϑ from pulsars. A given shower was considered only once and to the minimum angle ϑ. The probability *P* of the randomness of the number $N(\vartheta)$ of showers was calculated by the Monte Carlo method by drawing 27 events distributed isotropically over the



Distribution of 27 showers with $E > 4 \times 10^{19}$ eV in the map of the celestial sphere; δ and *RA* are the declination and right ascension, respectively, *X* is the axis of magnetic-field lines, and the dashed line is the conditional boundary of the Orion Arm.

0021-3640/03/7704-0151$24.00

Showers that arrived from the direction of the Orion Arm and had energies $E > 4 \times 10^{19}$ eV and pulsars correlating with them

n	Shower arrival date	*RA*, deg	δ, deg	Pulsar, PSR	Distance, kpc	Log *T*, yr
1	Jan. 16, 1988	8.7	36.3	0045 + 33	3.6	6.91
2	Jan. 26, 1996	21.3	45.7	0053 + 47	1.0	6.32
3	Oct. 19, 1989	24.0	57.0	0136 + 57	2.8	5.61
4	Dec. 1, 1991	235.4	79.8	1322 + 83	0.7	7.27
5	Feb. 21, 1999	274.2	54.5	1839 + 56	1.6	7.24
6	Dec. 15, 1989	283.5	29.3	1912 + 25	2.0	7.64
7	Nov. 4, 1985	297.3	45.2	1953 + 50	1.7	6.78
8	Jan. 13, 1995	314.8	57.8	2045 + 56	8.5	5.83
9	Oct. 26, 1985	335.2	51.0	2217 + 47	2.4	6.49
10	Feb. 8, 1983	342.9	65.8	2224 + 65	2.0	6.05

celestial sphere, with allowance for the array exposure (for details, see [9]). The number of draws was determined by the accuracy of determining the randomness probability and reached 10^6 in some cases.

Analysis shows that a correlation between the particle arrival directions with pulsars is observed only in the direction of the Orion Arm of the Galaxy. Among eleven showers arrived from the direction of the Orion Arm, ten showers fall within the angle $\theta < 6°$ from pulsars. The table presents the shower arrival date, its coordinate, pulsars, distances to them, and their age *T*. The probability of random observation of ten showers within $\theta < 6°$ from pulsars is $P = 0.05$. Among the remaining 16 showers beyond the Orion Arm, only six are within $\theta < 6°$ from pulsars. The probability of the randomness of this event is $P \sim 0.7$. We emphasize that three of these six showers have maximum detected energies. The shower detected on May 7, 1989, had an energy of 1.2×10^{20} eV and angle $\theta < 1°$ from the pulsar PSR 0458 + 46 spaced by 1.8 kpc. Two other showers had energies 7.4×10^{19} and 6.8×10^{19} eV and angles less than 6° from the pulsars PSR 0940 + 10 and 0917 + 63 spaced by 1.6 and 0.7 kpc, respectively. Correlation between these showers and pulsars is apparently due to the fact that the magnetic field slightly deviates high-energy particles.

The shower with the maximum observed energy $E = 3.2 \times 10^{20}$ eV was detected at the Fly's Eye array (USA) and was interpreted in [10] as being formed by the gold nucleus. At this array, in contrast to the Yakutsk array and AGASA array (Japan), the development of a shower in the atmosphere is directly measured by detecting ionization radiation of air atoms excited by the particles of the shower.

The mean lifetime of gold nuclei in the Galaxy is equal to the mean lifetime of protons with energy $E_p = E_{\mathrm{Au}}/z$ (z is the charge of Au). Calculations made in [11] show that the lifetime of Au nuclei with $E_{\mathrm{Au}} = 4 \times 10^{19}$ eV is longer than the lifetime of protons with the same energy by a factor of 10 and is equal to $\sim 10^5$ yr. This particle lifetime is sufficient for the isotropization of their arrival directions [11, 12] and is much less than the $\sim 10^8$ yr [6, 7] necessary for the formation of the spectrum cutoff caused by their interaction with the relict radiation. Thus, the two requirements presented at the beginning of the paper are satisfied.

According to data reported in [13, 14], the chemical composition of cosmic rays changes gradually to heavier elements with energy; protons prevail (92.5%) for energies $\sim 10^{10}$ eV [13], heavy nuclei dominate for $\geq 3 \times 10^{15}$ eV [13], iron nuclei prevail for $\sim 10^{19}$ eV [14], and gold nuclei dominate for $\sim 10^{20}$ eV. The change in the spectrum of cosmic rays [13] for $\sim 3 \times 10^{15}$, 6×10^{17}, and 10^{19} eV is apparently caused by a change in the chemical composition of particles.

One can conclude that cosmic rays with extremely high energies originate, in all likelihood, from pulsars.

This work was supported by the Russian Foundation for Basic Research, project no. 00-02-16325. The Yakutsk EAS array was supported by the Ministry of Education of the Russian Federation, project no. 01-30.

REFERENCES

1. V. S. Berezinsky, M. Kachelriess, and A. Vilenkin, Phys. Rev. Lett. **79**, 4302 (1997).
2. P. Tinyakov and I. Tkachev, J. Phys. Cosm. Jpn. **70**, 58 (2001).
3. A. V. Uryson, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 71 (1996) [JETP Lett. **64**, 77 (1996)].
4. J. Rachen, T. Stanev, and P. Biermann, Astron. Astrophys. **273**, 377 (1993).
5. N. Sakaki, M. Chikawab, and M. Fukushima, in *Proceedings of 27th International Cosmic Rays Conference* (Hamburg, 2001), Vol. 1, p. 333.
6. K. Greisen, Phys. Rev. Lett. **16**, 748 (1966).
7. G. T. Zatsepin and V. A. Kuzmin, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 144 (1966) [JETP Lett. **4**, 99 (1966)].

8. J. N. Taylor, R. N. Manchester, and A. G. Lyne, Astrophys. J., Suppl. **88**, 529 (1993).
9. N. N. Efimov, A. A. Mikhailov, and M. I. Pravdin, in *Proceedings of 18th International Cosmic Rays Conference* (Bangalore, 1983), Vol. 2, p. 149.
10. L. A. Anchordoqui, M. T. Dova, T. P. McCauley, *et al.*, Nucl. Phys. B (Proc. Suppl.) **97**, 203 (2001).
11. V. S. Berezinsky, S. I. Grigoryeva, A. A. Mikhailov, *et al.*, in *Proceedings of International Symposium on Astrophysical Aspects of the Most Energetic Cosmic Rays* (World Sci., Singapore, 1990), p. 134.
12. V. N. Zirakashvili, D. N. Pochepkin, V. S. Ptuskin, and S. I. Rogovaya, Izv. Ross. Akad. Nauk, Ser. Fiz. **59**, 153 (1995).
13. M. Nagano and A. A. Watson, Rev. Mod. Phys. **72**, 689 (2000).
14. A. D. Erlykin, A. A. Mikhailov, and A. W. Wolfendale, J. Phys. G: Nucl. Part. Phys. **28**, 2225 (2002).

Translated by R. Tyapaev

JETP Letters, Vol. 77, No. 4, 2003, pp. 154–157. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 184–187.

Stabilization of Yang–Mills Chaos in Non-Abelian Born–Infeld Theory¶

D. V. Gal'tsov and V. V. Dyadichev

Faculty of Physics, Moscow State University, Moscow, 119899 Russia

e-mail: galtsov@grg.phys.msu.su, vlad@grg1.phys.msu.su

Received January 10, 2003

We investigate dynamics of the homogeneous time-dependent $SU(2)$ Yang–Mills fields governed by the non-Abelian Born–Infeld Lagrangian, which arises in superstring theory as a result of summation of all orders in the string slope parameter α'. It is shown that, generically, the Born–Infeld dynamics are less chaotic than those in the ordinary Yang–Mills theory, and at a high enough field strength the Yang–Mills chaos is stabilized. More generally, a smothering effect of the string nonlocality on the behavior of classical fields is conjectured. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 11.15.Kc; 11.25.-w; 05.45.Mt

The chaotic nature of the classical Yang–Mills (YM) dynamics was recognized more than twenty years ago [1–5]. It can be revealed even in simple models of the homogeneous fields depending only on time [6]. Further studies have shown the persistence of chaotic features in classical dynamics of spatially varying YM fields, leading to a new kind of "ultraviolet catastrophe" [7].

Here, we would like to investigate whether this tendency is preserved on a deeper level of superstring theory or whether the latter provides a smothering effect. The possibility of exploring this question quantitatively is related to the existence of a closed-form effective action for open strings or the stuck of *D*-branes, which accumulates all orders in the string slope parameter α'. Strictly speaking, an effective action exact in α' is known only for the $U(1)$ gauge field, in which case it is the Born–Infeld (BI) action [8, 9]. In the non-Abelian case, the exact calculation of the effective action is not possible, but a certain non-Abelian version of the Born–Infeld (NBI) action can still be envisioned in a closed form. This problem was extensively discussed recently [10–12]. One reasonable proposal (though perhaps still not giving the precise answer in the sense of superstring theory) involves a symmetrized trace [10] construction of the action. Another, simpler form uses a direct non-Abelian generalization of the $U(1)$ BI action, which we will call the ordinary trace model. Investigation of static solitons in NBI theory [13] has shown qualitative agreement between the results of both the symmetrized trace and the ordinary trace models, and so here, for reasons of simplicity, we adopt the latter one:

$$S_{NBI} = \beta^2\int d^4x\left(1-\sqrt{1+\frac{1}{2\beta^2}F^a_{\mu\nu}F^{\mu\nu}_a-\frac{1}{16\beta^4}(F^a_{\mu\nu}\tilde{F}^{\mu\nu}_a)^2}\right), \quad (1)$$

with β being the critical BI field, which in string theory is $\beta = 1/2\pi\alpha'$.

The simplest non-Abelian configuration for which the ordinary YM theory predicts chaotic behavior [6] is

$$A = u\mathbf{T}_1 dx + v\mathbf{T}_2 dy \quad (2)$$

where u and v are functions only of time and $\mathbf{T}_1$ and $\mathbf{T}_2$ are the gauge group generators. The corresponding field strength

$$F = \dot{u}\mathbf{T}_1 dt\wedge dx + \dot{v}\mathbf{T}_2 dt\wedge dy + uv\mathbf{T}_3 dx\wedge dy,$$

contains both electric and magnetic components, but these are mutually orthogonal, so the pseudoscalar invariant $\mathrm{tr}F\tilde{F}$, generically dominant (1) at high field strength, is equal to zero. The corresponding one-dimensional Lagrangian reads

$$L = \beta^2(1-\sqrt{1-\beta^{-2}(\dot{u}^2+\dot{v}^2-v^2u^2)}). \quad (3)$$

The equations of motion have the form

$$\ddot{u} = -uv^2+\frac{2\dot{u}vu(\dot{u}v+\dot{v}u)}{\beta^2+u^2v^2}, \quad (4)$$

$$\ddot{v} = -vu^2+\frac{2\dot{v}vu(\dot{u}v+\dot{v}u)}{\beta^2+u^2v^2}. \quad (5)$$

In the limit $\beta\longrightarrow\infty$ of the ordinary YM action, the second terms on the right-hand sides of these equations

¶This article was submitted by the authors in English.

vanish, and one obtains the "hyperbolic billiard" [6, 14, 15], known in various problems of mathematical physics, which is a two-dimensional mechanical system with the potential u^2v^2. This potential has two valleys along the lines $u = 0$, $v = 0$, the particle motion being confined by hyperbolas $uv = \text{const}$. The existence of these valleys is crucial for the emergence of chaos.

The trajectories of the system governed by the BI action depend on the energy integral E and the parameter β. Rescaling of the field variables together with time rescaling maps configurations with different E and β onto each other; therefore, it will be enough to consider only the dependence of the dynamics on β assuming for the energy any fixed value, e.g., $E = 1$. One simple (though nonrigorous) method of revealing chaoticity is analysis of the geodesic deviation equation for a suitably defined pseudoriemannian space whose geodesics coincide with the trajectories of the system [16, 17]. The Lagrangian (3) gives rise to the metric tensor

$$ds^2 = \left(1 + \frac{u^2v^2}{\beta^2}\right)dt^2 - \frac{1}{\beta^2}(du^2 + dv^2), \qquad (6)$$

which is regular for all finite $\beta \neq 0$. Consider the deviation equation for two close geodesics

$$\frac{D^2n^i}{ds} = R^i{}_{jkl}u^j u^k n^l, \qquad (7)$$

where D^2n^i is the covariant differential of their transversal separation, R^i_{jkl} is the Riemann–Christoffel tensor, and u^j is the three-speed in the space (6). Locally, the geodesic deviation is described by the matrix $R^i_{jkl}u^j u^k$ depending on points u, v, $\dot{u}$, and $\dot{v}$ of the phase space. If at least one eigenvalue of this matrix is positive, the geodesics will diverge exponentially with time. Negative eigenvalues correspond to oscillations or to slower divergence, e.g., a power-law. In our case, one of the eigenvalues of the matrix $R^i_{jkl}u^j u^k$ vanishes (due to the static nature of the metric); the other two are the roots of a quadratic equation:

$$\lambda_{1,2} = \frac{1}{2}(\mathcal{B} \pm \sqrt{\mathcal{B}^2 + 4\mathcal{C}}), \qquad (8)$$

where

$$\mathcal{B} = \frac{2uv\dot{u}\dot{v} - \beta^2(u^2 + v^2)}{\beta^2 UW} + \frac{(\dot{v}u + \dot{u}v)^2}{\beta^2 U^2 W}, \qquad (9)$$

$$\mathcal{C} = \frac{v^2u^2(3\beta^2 + u^2v^2)}{\beta^2 U^2 W}, \qquad (10)$$

with W and U being positive functions

$$W = 1 + \frac{u^2v^2 - \dot{u}^2 - \dot{v}^2}{\beta^2}, \quad U = 1 + \frac{u^2v^2}{\beta^2}.$$

It is easy to see that both nonzero eigenvalues (8) are real, one positive and the other negative. Therefore, for any finite β, there exist locally divergent geodesics, and it can be expected that for any β the motion will remain chaotic. But in the limit $\beta \longrightarrow 0$, these eigenvalues tend to zero, and the analysis becomes inconclusive. Qualitatively, this phenomenon of decrease of the geodesic divergence with decreasing β can be attributed to the lowering of the degree of chaoticity. It is also worth noting that for $\beta \longrightarrow 0$ the second terms in the equations of motion (4), (5) look like singular friction terms.

Another simple tool in the analysis of chaos is the calculation of the Lyapunov exponents, defined as

$$\chi = \lim_{t\to\infty}\frac{1}{t}\log\frac{|\delta x(t)|}{|\delta x(0)|}, \qquad (11)$$

where $\delta x(t)$ is a solution of the linearized perturbation equation along the chosen trajectory. Here, x stands for the (four-dimensional) phase space coordinates of the original system and the Cartesian metric norm is chosen. The positive value of χ signals that the close trajectories diverge exponentially with time, so the motion is unstable. Stable regular motion corresponds to zero Lyapunov exponents.

The third standard tool is the construction of Poincaré sections. Recall that the Poincare section is some hypersurface in the phase space. The phase space for the system (13) is four-dimensional, but energy conservation restricts the motion to a three-dimensional manifold. Imposing one additional constraint will fix a two-dimensional surface, and we can find a set of points corresponding to the intersection of the chosen trajectory with this surface. For regular motion, one typically gets some smooth curves, while in the chaotic case the intersection points fill finite regions on the surface.

Both methods reveal that with decreasing β the degree of chaoticity is diminished, although we were not able to perform calculations for very small β to decide definitely whether the chaos is stabilized or not. But in fact ansatz (2) is not generic enough from the point of view of BI dynamics, since there the leading term $(F\tilde{F})^2$ at high field strength in (1) is zero. Thus, we can hope to draw more definitive conclusions by exploring another ansatz for which the invariant $(F\tilde{F})$ is nonzero. Such an example is provided by a simple axially symmetric configuration of the $SU(2)$ gauge field, which is also parameterized by two functions of time u, v [18]:

$$A = \mathbf{T}_1 u dx + \mathbf{T}_2 u dy + \mathbf{T}_3 v dz. \qquad (12)$$

The field strength contains both electric and magnetic components

$$F = \dot{u}(\mathbf{T}_1 dt \wedge dx + \mathbf{T}_2 dt \wedge dy) + \dot{v}\mathbf{T}_3 dt \wedge dz$$
$$+ u^2\mathbf{T}_3 dx \wedge dy + uv(\mathbf{T}_1 dy \wedge dz + \mathbf{T}_2 dz \wedge dx),$$

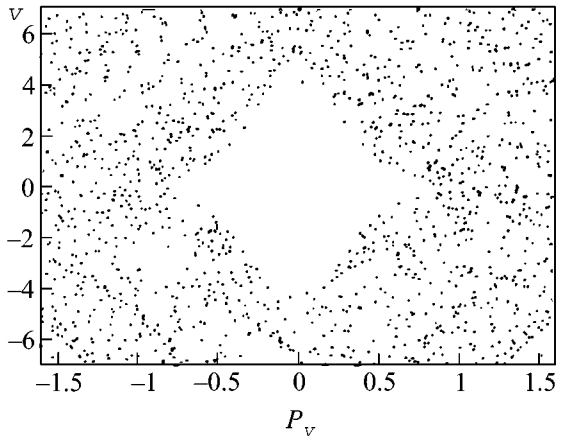


Fig. 1. Poincare plane v, P_v of the system (13), where P_v is the canonical momentum conjugate to v, for $\beta = 0.32$.

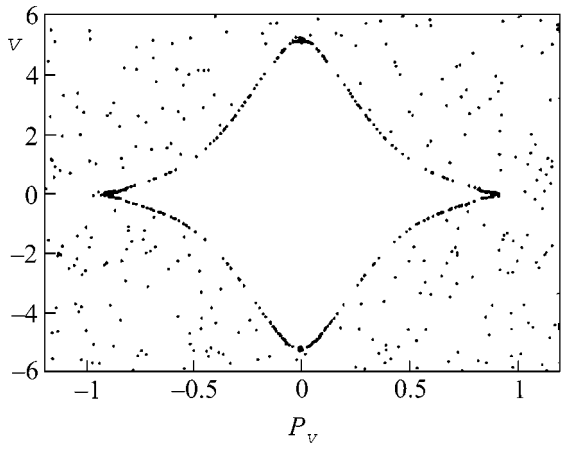


Fig. 2. $\beta = 0.317$.

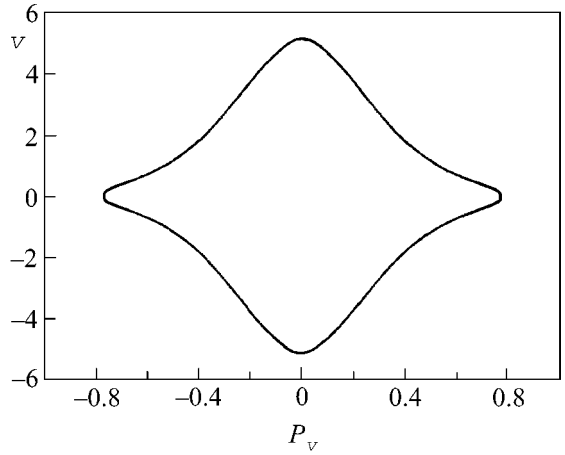


Fig. 3. $\beta = 0.31$.

and now the pseudoscalar term $(F\tilde{F})^2$ is nonzero. Substituting this ansatz into the action (1), we obtain the following one-dimensional Lagrangian:

$$L_1 = \beta^2(1-\mathcal{R}), \tag{13}$$

$$\mathcal{R} = \sqrt{1-\frac{2\dot{u}^2+\dot{v}^2-u^2(u^2+2v^2)}{\beta^2}-\frac{u^2(2\dot{u}v+\dot{v}u)^2}{\beta^4}}.$$

In the limit of the usual YM theory $\beta \longrightarrow \infty$, we recover the system that was considered in [18]:

$$L_{YM} = \frac{1}{2}(2\dot{u}^2+\dot{v}^2-u^4-2u^2v^2). \tag{14}$$

The corresponding potential has valleys along $v = 0$. As in the case of the hyperbolic billiard with potential v^2u^2, the dynamics governed by the usual quadratic YM Lagrangian motion exhibit a chaotic character.

However, for the NBI action with finite β, the situation is different. One can again use the geodesic deviation method, the corresponding metric being

$$ds^2 = \left(1+\frac{u^4+2u^2v^2}{\beta^2}\right)dt^2 - \frac{du^2+2dv^2}{\beta^2}-\frac{u^2(2vdu+udv)^2}{\beta^4}. \tag{15}$$

The energy integral is given by the expression

$$E = \frac{\beta^2+u^4+2v^2u^2}{\mathcal{R}}-\beta^2. \tag{16}$$

One can calculate the eigenvalues of the geodesic deviation operator as before. In general, one finds that the larger eigenvalue is smaller than one in the case of the previous ansatz (2). Now regions in the configuration space in which both nonzero eigenvalues are negative appear; this corresponds to locally stable motion. With decreasing β, the relative volume of the local stability regions increases, although for every β there still exist regions of local instability as well. To obtain more definitive conclusions, we performed numerical experiments using Lyapunov exponents and Poincare sections. The results look as follows. For large β, the dynamics of the system (13) are qualitatively similar to those in the ordinary YM theory (14) and exhibit typical chaotic features. The trajectories enter deep into the valleys along $u = 0$. The Poincare sections consist of clouds of points filling the finite regions of the plane, while the Lyapunov exponents remain essentially positive. The situation changes drastically with decreasing β. For some value of this parameter, depending on a particular choice of the initial conditions, one observes after a series of bifurcations a clear transition to regular motion. The points on the Poincare sections $u = 0$ line up along the smooth curves (sections of a torus), while the Lyapunov exponent goes to zero. The motion becomes quasiperiodic. A typical feature of such regu-

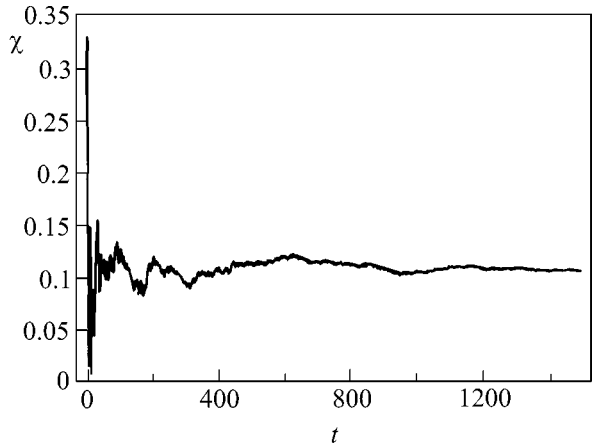


Fig. 4. Lyapunov exponent corresponding to Fig. 1.

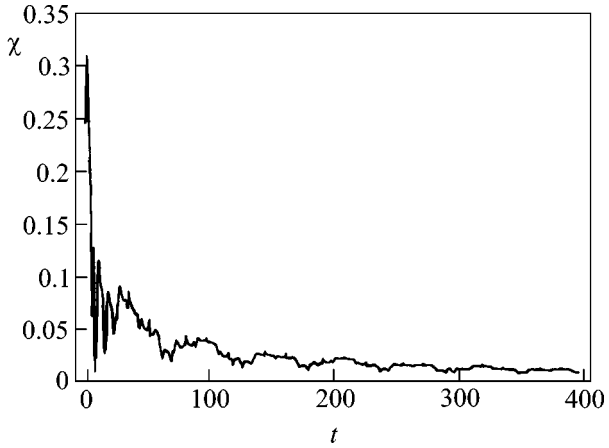


Fig. 5. Lyapunov exponent corresponding to Fig. 3.

lar motion is that the trajectory no longer enters the potential valleys. Some Poincare sections illustrating this chaos–order transition are presented in Figs. 1–3. The Lyapunov exponents corresponding to the first and the last picture of this series are shown in Figs. 4 and 5, respectively.

Comparing with our first system (2), we can guess that a substantial role in the chaos–order transition in the NBI theory is played by the pseudoscalar term in the action (1). Since generically this term does not vanish, it is tempting to say that the transition observed is a generic phenomenon in the non-Abelian BI theory. The effect is essentially nonperturbative in α', and we conjecture that it reflects the typical smothering effect of the string nonlocality on the usual stiff field-theoretical behavior. A related question is whether the cosmological singularity, which was recently shown to be chaotic at the supergravity level of string models [19], will also be regularized when higher α' corrections are included. Unfortunately, there is no closed form effective action for the closed strings analogous to the BI action for the open strings. But combining gravity in the lowest order in α' with a matter action exact in α', we can probe the nature of the cosmological singularity as well. This will be reported in a separate publication.

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16306).

REFERENCES

1. G. Z. Baseian, S. G. Matinian, and G. K. Savvidy, Pis'ma Zh. Éksp. Teor. Fiz. **29**, 641 (1979) [JETP Lett. **29**, 587 (1979)].
2. B. V. Chirikov and D. L. Shepelyansky, Pis'ma Zh. Éksp. Teor. Fiz. **34**, 171 (1981) [JETP Lett. **34**, 163 (1981)].
3. B. V. Chirikov and D. L. Shepelyansky, Yad. Fiz. **36**, 1563 (1982) [Sov. J. Nucl. Phys. **36**, 908 (1982)].
4. A. A. Birò, S. G. Matinyan, and B. Müller, *Chaos and Gauge Theory* (World Sci., Singapore, 1994).
5. S. G. Matinian, gr-qc/0010054.
6. S. G. Matinian, G. K. Savvidy, and N. G. Ter-Arutunian Savvidy, Zh. Éksp. Teor. Fiz. **80**, 830 (1981) [Sov. Phys. JETP **53**, 421 (1981)].
7. H. B. Nielsen, H. H. Rugh, and S. E. Rugh, hep-th/9605013.
8. E. S. Fradkin and A. A. Tseytlin, Phys. Lett. B **163B**, 123 (1985).
9. R. G. Leigh, Mod. Phys. Lett. A **4**, 2767 (1989).
10. A. Tseytlin, Nucl. Phys. B **501**, 41 (1997); hep-th/9701125.
11. A. Tseytlin, hep-th/9908105.
12. S. Gonorazky, F. A. Schaposnik, and G. A. Silva, Phys. Lett. B **449**, 187 (1999); hep-th/9812094.
13. V. V. Dyadichev and D. V. Gal'tsov, Nucl. Phys. B **590**, 504 (2000); hep-th/0006242.
14. G. M. Asatrian and G. K. Savvidy, Phys. Lett. A **99A**, 290 (1983).
15. P. Dahlqvist and G. Russberg, Phys. Rev. Lett. **65**, 2837 (1990).
16. M. Szydlowski, Phys. Lett. A **176**, 22 (1993).
17. Y. Sota, S. Suzuki, and K. Maeda, Class. Quantum Grav. **13**, 1241 (1996); gr-qc/9505036.
18. B. K. Darian and H. P. Kunzle, Class. Quantum Grav. **13**, 2651 (1996); gr-qc/9608024.
19. T. Damour and M. Henneaux, Phys. Rev. Lett. **85**, 920 (2000); hep-th/0003139.

JETP Letters, Vol. 77, No. 4, 2003, pp. 158–161. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 188–191.

Hyper-Rayleigh Light Scattering in the Third Harmonic Generation in Silver Island Films

E. M. Kim, S. S. Elovikov, and O. A. Aktsipetrov

Physics Department, Moscow State University, Moscow, 119992 Russia

e-mail: JaneKim@shg.ru

Received January 14, 2003

Hyper-Rayleigh scattering in the third harmonic generation (incoherent optical third harmonic) was experimentally observed in spatially inhomogeneous, quasi-two-dimensional structures formed by silver island films. A comprehensive analysis of the linear and nonlinear light scattering in combination with the results of atomic-force microscopy revealed the fractal nature of the island films studied in the experiment. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 78.35.+c; 42.65.Ky; 78.66.Bz

Hyper-Rayleigh scattering (HRS) in the optical second harmonic (SH) generation, which is, on the one hand, an incoherent analogue of the conventional SH generation and, on the other hand, a nonlinear analogue of the Rayleigh scattering (RS), has been intensively studied in bulk molecular systems starting with the 1970s [1–3]. Later, HRS in the SH generation in molecular systems, i.e., liquids and gases, in which the source of incoherence was represented by the thermal fluctuations of optical and nonlinear optical parameters of the medium, was extended to the case of spatially inhomogeneous solid systems with random inhomogeneities, such as inhomogeneous thin films and rough surfaces. Despite the low efficiency of the incoherent SH generation in inhomogeneous thin films and, hence, the low intensity of the diffuse HRS signal at the SH frequency, these studies made it possible to develop correlation methods for analyzing the random inhomogeneities of the optical, structural, and morphological properties of micron-sized and nanosized objects, such as the microcrystal and polydomain structures of ferroelectric ceramic thin films [4], the structural inhomogeneities of magnetic [5] and ferroelectric [6] Langmuir monolayers, layer-by-layer assembled ferromagnetic films [7], Langmuir fullerene films [8], etc. However, the low intensity of the HRS signal corresponding to the diffuse SH component resulted in no attempts being made to experimentally observe the HRS for higher harmonics.

In this paper, we describe an experimental observation of the HRS in the optical third harmonic (TH) generation in silver island films, which serve as a convenient model of a two-dimensional random ensemble of nonlinear scatterers. We perform a combined analysis of the linear and nonlinear (at the SH and TH frequencies) scattering indicatrices in combination with the analysis of the results of atomic-force microscopy of the films to study the fine morphological features of the samples under investigation.

The silver island films were deposited on a silicon substrate with the (001) orientation by the thermal evaporation technique. The substrate surface was preliminarily cleaned by chemical etching in sulfuric acid. The deposition was performed with a residual gas pressure of $p < 10^{-5}$ torr in the vacuum conditions. The structural properties of the films were studied by an atomic-force microscope in the constant-force regime with a lateral resolution of 30 Å and a resolution of 4 Å along the normal to this plane. The AFM tip radius was 20 nm. Since the size of the inhomogeneities under study was comparable with the tip radius, the instrumental function of the measurement caused distortions of the true inhomogeneity profile and led to overestimated values of its dimension in the plane perpendicular to the tip: $W = \sqrt{8bR}$, where b is the average transverse dimension of a particle and R is the tip radius.

The experimental study of the HRS indicatrices in the TH and SH generation was performed using a YAG:Nd^{3+} laser with a wavelength of 1064 nm, a pulse power density of 10 MW/cm^2, a pulse width of 15 ns, and a pulse repetition rate of 25 Hz. Part of the pumping radiation was diverted into a reference channel, and the SH and TH responses of the silver island films were normalized to the corresponding signal from the reference quartz crystal. This technique made it possible to reduce the effect of the fluctuations of laser radiation power on the accuracy with which the polar angular indicatrices of the weak HRS signal were measured.

Figure 1a shows the HRS indicatrix in the TH generation (the scattered light intensity at the TH wavelength $\lambda = 355$ nm) as a function of the polar scattering angle. The plot demonstrates the appearance of a diffuse incoherent TH component due to the reflection of

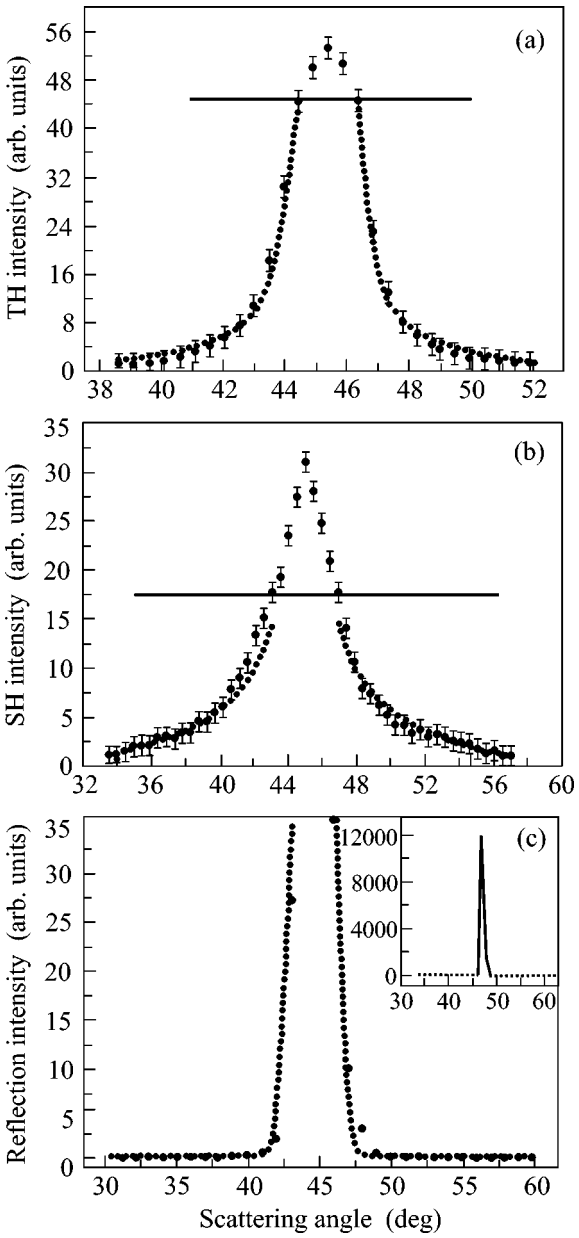


Fig. 1. (a) HRS indicatrix for the TH generation, (b) HRS indicatrix for the SH generation, and (c) linear RS indicatrix at the SH wavelength in an silver island film. The inset shows the RS indicatrix in the whole intensity range.

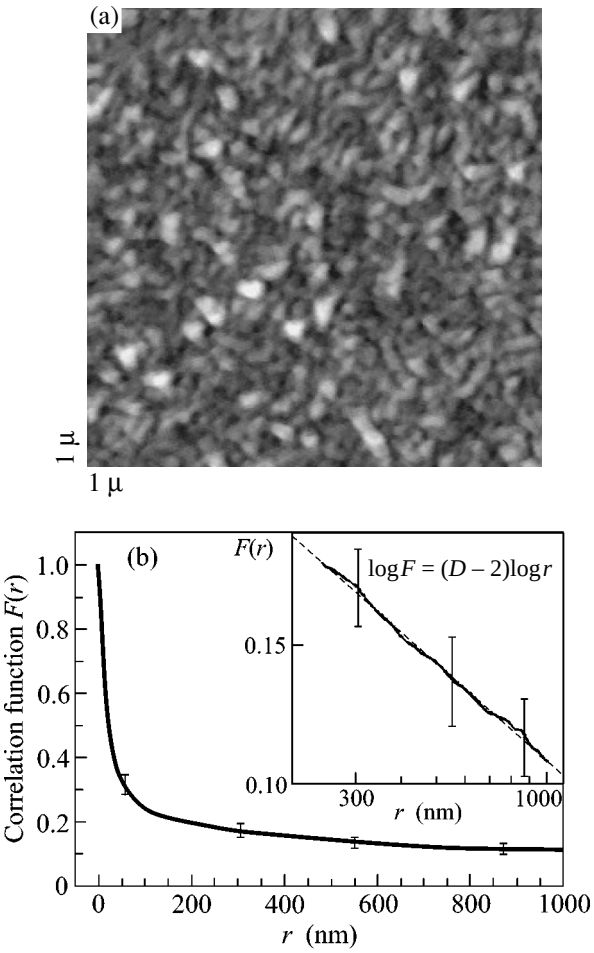


Fig. 2. (a) Atomic-force microscope image of a silver island film; the frame size is 1000 × 1000 nm. (b) The density–density correlation function $F(r)$; the inset shows the density–density correlation function on the logarithmic scale for distances greater than the maximal monomer size (250 nm).

the laser radiation from the two-dimensional random ensemble of nonlinear scatterers, i.e., from the island film formed by silver nanoparticles. Figures 1b and 1c show the HRS and linear RS indicatrices, respectively, at the SH wavelength $\lambda = 532$ nm for the same silver island film. All indicatrices are measured in the *s*–*s* geometry, i.e., for the *s* polarization of the incident and scattered light. The aperture of the detecting system was 2° for the HRS indicatrix at the SH and 1° for the HRS indicatrix at the TH. Since the Si(001) substrate of the island film was the source of a coherent specular SH and TH signal, the desired SH and TH signal from the ensemble of silver clusters, I_{Ag}, was set off by subtracting the specular SH and TH signal generated by the substrate, I_{Si}, from the total signal received from the substrate–silver system I_{Si+Ag}: $I_{Ag} = I_{Si+Ag} - I_{Si}$. The resulting angular indicatrices demonstrate the appearance of a broadened scattered component localized in the angular interval 32°–58° with a maximum in the specular direction and with a monotonic decrease down to zero outside this interval. The horizontal line separates the regular signal from the diffuse radiation according to the instrumental function of the detecting system. The linear RS indicatrix exhibits a narrow intense maximum in the specular direction (inset in Fig. 1c), which is associated with the specular reflection of light from the surface of the Si(001) substrate.

Figure 2a shows a typical atomic-force microscope image of the surface of an silver island film. One can see that the silver islands have the form of clusters that are stuck together and have an average size of about

150 nm. The average monomer height is 6 nm, while the inhomogeneities of the silicon surface are smaller, about 1 nm. The system of clusters is a two-dimensional ensemble, because the longitudinal dimension of the particles is much greater than their transverse dimension.

To determine the correlations between the morphological and nonlinear optical properties of the silver island films from the analysis of the atomic-force microscope images, we constructed the statistical function of the structure, namely, the density–density correlation function, which characterizes the correlation of the substance distribution over the sample plane:

$$F(\mathbf{r}) = \frac{1}{S}\int n(\mathbf{r}+\mathbf{r}')n(\mathbf{r}')d\mathbf{r}',$$
$$n(\mathbf{r}') = \begin{cases}1, & r \in S_{\mathrm{Ag}}\\ 0, & r \notin S_{\mathrm{Ag}},\end{cases} \quad (1)$$

where S_{Ag} is the space occupied by silver; $n(\mathbf{r})$ is the surface density function, which is equal to unity when the end of the radius vector falls within a metal particle (the same or neighboring) and to zero when the end of radius vector falls within a metal-free region of the substrate surface; and S is the film area over which the averaging is performed. Since the samples are isotropic, we assume that $F(\mathbf{r}) = F(r)$. Figure 2b represents the density–density correlation function for the silver island film for which the light scattering indicatrices shown in Fig. 1 were obtained. At distances greater than the maximal monomer size (250 nm), the function $F(r)$ smoothly decreases and tends to zero. The inset in Fig. 2b shows the function $F(r)$ on a logarithmic scale for distances greater than 250 nm. This dependence was approximated by the power function $1/r^{2-D}$ with the parameter $D = 1.62$, which is represented by the dashed line in the inset in Fig. 2b. Hence, the silver cluster structure of the silver island film has the properties of a fractal with the dimension $D = 1.62$.

To interpret the indicatrices that were obtained for the linear and nonlinear scattering, we use the assumption that the correlations of the local optical-field factors and the corresponding optical susceptibilities in the sample plane are statistically independent. Then, the expression for the two-point correlation function of the nonlinear polarization can be represented in the form

$$\Pi_\alpha(r) = A\Pi_\chi(r) + B\Pi_L(r) + AB\Pi_\chi(r)\Pi_L(r), \quad (2)$$

where $\Pi_\chi(r)$ is the correlation function of the nonlinear susceptibility, $\Pi_L(r)$ is the correlation function of the local optical-field factor, A and B are the adjustable parameters, and the subscript $\alpha = 1, 2, 3$ indicates the RS and the HRS in the SH and TH generation. The contributions of χ_ω, $\chi_{2\omega}$, and $\chi_{3\omega}$ can be determined through the silver density function as $\Pi_\chi(r) \sim F(r) - \bar{n}^2$. The correlation function of the local optical-field factor can be related to the Fourier transform of the linear indicatrix of the RS: $\Pi_L(r) \sim \int I_\omega(\theta) J_0[-ikr]kdk$, where $k = \omega(\sin\theta - \sin\theta_0)/c$ and J_0 is the zero-order Bessel function. Then, the expression for the correlation function of the nonlinear polarization takes the form

$$\Pi_\alpha(r) \sim AF(r) + Be^{-r^2/w^2} + ABF(r)e^{-r^2/w^2}, \quad \alpha = 1, 2, 3, \quad (3)$$

where $\Pi_L(r) \sim e^{-r^2/w^2}$ is the analytic expression for the correlation function of the local optical-field factor, which is obtained from the approximation of the numerical calculation of the integral $\int I_\omega(\theta)J_0[-ikr]kdk$. From expression (3), it follows that the correlation function of the nonlinear polarization is the sum of three terms, the first of which is characterized by the density–density correlation function, the second, by the Fourier transform of the linear RS indicatrix, and the third is the product of the two aforementioned functions.

The general phenomenological expression for the incoherent diffuse RS signal and the HRS signal in the SH and TH generation is as follows:

$$I_{\alpha\omega}(\theta) \sim \int \Pi_\alpha(\mathbf{r}) \exp\{-i\alpha(\sin\theta - \sin\theta_0)\omega\mathbf{e}_\parallel\mathbf{r}/c\}d\mathbf{r}, \quad \alpha = 1, 2, 3, \quad (4)$$

where θ_0 and θ are the angle of incidence and the scattering angle, respectively, and $\mathbf{e}_\parallel$ is the unit vector of the wave-vector projections of the incident and scattered waves. If we represent the silver island film as a two-dimensional isotropic ensemble of scatterers, we can replace the vector integration in expression (4) with the scalar integration:

$$I_{\alpha\omega}(\theta) \sim \int \Pi_\alpha(r) J_0[-i\alpha(\sin\theta - \sin\theta_0)\omega r/c]rdr. \quad (5)$$

For the approximation of the indicatrices of the RS and the HRS in the SH and TH generation, it is necessary to substitute expression (3) obtained for the correlation function of the nonlinear polarization into expression (5) for the diffuse component of the reflected linear signal and the SH and TH signals:

$$I_{\alpha\omega}(\theta) \sim A\frac{1}{k^D} + Be^{-M_\alpha k^2 w^2} + ABe^{-N_\alpha k^2 a^2}, \quad \alpha = 1, 2, 3. \quad (6)$$

Here, $D = 1.62$ is the fractal dimension, ω is the half-width of the Fourier transform of the linear RS indicatrix, $a = 0.27$ is a parameter, M_1 and N_1 are parameters corresponding to the RS indicatrix, M_2 and N_2 are parameters corresponding to the HRS indicatrix in the SH generation, M_3 and N_3 are parameters corresponding to the HRS indicatrix in the TH generation, and A and B are adjustable parameters. The approximation by formula (6) describes the indicatrices of the RS and the

HRS for the SH and TH most closely in terms of the total square deviation. To find out which of the functions, the power function or the Gaussian one, makes the greater contribution to the angular indicatrices, we approximate the latter by the two functions separately. The power function with the exponent determined from the density–density correlation function $D = 1.62$ provides a closer approximation of the HRS indicatrix in the SH generation (dotted line in Fig. 1b) than the Gaussian function. This means that, at large distances, the fractal property is retained with the same exponent $D = 1.62 < 2$ as in the nanometer range, for which the density–density correlation function $F(r)$ was calculated. At the same time, the Gaussian function more adequately describes the RS indicatrix in terms of the total square deviation (dotted line in Fig. 1c) than does the power function. As for the HRS indicatrix in the TH generation, it is equally well described by the Gaussian function and the power function in terms of the total square deviation. In Fig. 1a, the dotted line shows the approximation of the HRS indicatrix in the TH generation by formula (6). The comparison shows that the assumption about the statistical independence of the susceptibility and the local optical field factor is justified for describing the HRS in the SH generation and is less suitable for describing the linear and nonlinear HRS indicatrices in the TH generation.

In closing, we note that the HRS was experimentally observed in the incoherent diffuse third optical harmonic generation, and, despite the weakness of this cubic nonlinear optical effect, the corresponding angular scattering indicatrices of the HRS from silver island films were thoroughly investigated. A comprehensive analysis of the indicatrices of the linear RS and the quadratic and cubic HRS from randomly inhomogeneous silver island films in combination with the analysis of the statistical properties of the silver island films by atomic-force microscopy revealed the fractal nature of the island film structure on both nanometer and micron scales. It was shown that the fractal property of the silver clusters is retained in the micron range with the same dimension $D = 1.62 < 2$ as in the nanometer range, for which the density–density correlation function was calculated.

We are grateful to N.V. Didenko and D.A. Muzychenko for the assistance in the experiment, and to A.A. Nikulin for numerous consultations on the theory of the HRS.

REFERENCES

1. S. Kielich, *Molecular Nonlinear Optics* (Wydawn. Naukowe Uniw., Poznan', 1972; Mir, Moscow, 1974).
2. P. K. Schmidt and G. W. Rayfield, Appl. Opt. **33**, 4286 (1994).
3. K. Clays and A. Persoons, Phys. Rev. Lett. **66**, 2980 (1991).
4. A. A. Fedyanin, N. V. Didenko, N. E. Sherstyuk, *et al.*, Opt. Lett. **24**, 1260 (1999).
5. T. V. Murzina, G. B. Khomutov, A. A. Nikulin, *et al.*, J. Opt. Soc. Am. B **17**, 63 (2000).
6. O. A. Aktsipetrov, T. V. Misuryaev, T. V. Murzina, *et al.*, Opt. Lett. **25**, 411 (2000).
7. T. V. Murzina, A. A. Nikulin, O. A. Aktsipetrov, *et al.*, Appl. Phys. Lett. **79**, 1309 (2001).
8. E. D. Mishina, T. V. Misuryaev, A. A. Nikulin, *et al.*, J. Opt. Soc. Am. B **16**, 1692 (1999).

Translated by E. Golyamina

JETP Letters, Vol. 77, No. 4, 2003, pp. 162–166. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 192–196.

Enrichment of the Nitrogen Atomic Component with the ^{15}N Isotope in a Post-Discharge Zone

N. M. Gorshunov* and S. V. Gudenko

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

** e-mail: gorshunov@imp.kiae.ru*

Received December 5, 2002; in final form, January 22, 2003

Isotope atomic composition in the post-discharge zone of a pulsed discharge in a nitrogen flow was studied by electron paramagnetic resonance. It is shown that, while the atomic concentration in the post-discharge zone decreases, the relative content of the ^{15}N isotope increases more than 30 times compared to its natural abundance. Such a high isotope enrichment exceeds more than fourfold the corresponding maximal value that is predicted for the nitrogen atoms by the existing theoretical model. Analysis of the experimental results shows that the isotope-selective dissociation of nitrogen molecules proceeds in the post-discharge zone after the most part of atoms created in the discharge zone had recombined at the tube surface. A mechanism that explains that a nitrogen isotope enrichment as high as that is quite possible is proposed. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 28.60.+s

INTRODUCTION

The isotope-selective population of upper vibrational levels v of nitrogen molecules in the nonequilibrium vibrational exchange between the $^{14}N_{2(v)}$ and $^{14}N^{15}N_{(v)}$ molecules, as well as the possibility of enriching the dissociation products or the products of $N_{2(v>12)} + O = NO + N$ reaction with the ^{15}N isotope, was reported in [1]. However, the experimental results on this effect are quite contradictory. In [2], it was reported that the pulsed discharge in a nitrogen–oxygen mixture gave NO molecules that are threefold enriched with the ^{15}N isotope at room temperature of walls and 130-fold upon cooling walls in liquid nitrogen. However, in the subsequent works [3, 4], the enrichment of NO molecules obtained under the same conditions and using a more thorough mass-spectrometric analysis was found to be no higher than 1.5. Recent EPR studies of the influence of impurities on the plasma composition in the post-discharge zone of a microwave discharge in nitrogen at room and higher temperatures have shown that the nitrogen atomic component is enriched with ^{15}N from 10 to 12 times [5]. However, it was shown theoretically in [6] that the highest isotope enrichment of the products of nonequilibrium vibrational exchange between the isotopic modifications of nitrogen molecules in the ground electronic state cannot exceed 8 at temperatures higher than room temperature. In spite of the fact that much work was devoted in the 1990s to studying the nitrogen excited states in electric discharge, the contradiction between different experiments and between the calculations and the experiments on isotope enrichment of the excited-molecule reaction products still remains open. For this reason, the enrichment of reaction products with the ^{15}N isotope, as expected according to the theoretical estimates and some experiments, has motivated us to further study this problem.

The purpose of this work was to study the physical effects leading to isotope enrichment of the nitrogen atoms obtained upon pumping pure nitrogen through the zone of pulsed discharge.

EXPERIMENTAL

In the experiment, nitrogen flow was passed through the zone of pulsed discharge, where electron impact induced population of various vibrational and electronic levels of a nitrogen molecule and dissociation of some of them. The scheme of the setup is shown in Fig. 1. Special purity grade nitrogen was used, which admitted from cylinder *1* through flow-rate-controlling cock *K*1 into discharge quartz tube *2* with an i.d. of

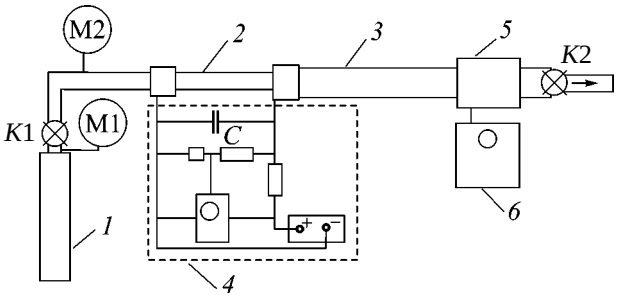


Fig. 1. Scheme of the setup: (*1*) cylinder with nitrogen, (*2*) discharge tube, (*3*) measuring tube, (*4*) pulsed voltage generator, (*5*) EPR cavity, and (*6*) EPR spectrometer.

0021-3640/03/7704-0162$24.00

0.7 cm and a length of 21 cm. The electric discharge was initiated by pulsed voltage generator *4* upon self-breakdown of capacitor *C* into gas in tube *2*. The discharge duration was about 2 μs, the pulse separation was from 50 to 60 ms, and the breakdown voltage was equal to several kilovolts. The specific energy contribution ε to the discharge was determined from a change in the energy accumulated in the capacitor during the discharge, from the nitrogen amount in the discharge tube, and from the number of discharge pulses during the gas flow through the tube. During the course of experiment, ε changed from 0.01 to 0.1 eV per molecule, and the pressure in the tubes between cocks *K*1 and *K*2 changed from 2.5 to 5.8 torr. From tube *2*, the flow admitted to measuring quartz tube *3* with an i.d. of 1.5 cm running along the EPR spectrometer cavity and through pressure- and gas-rate-controlling cock *K*2. The distance from the end of the discharge tube to the center of EPR cavity was $L \cong 20$ cm. The flow velocity V at room temperature was calculated from the measured gas flow rate and pressure. The concentrations of the ^{14}N and ^{15}N atoms in the flow of partially dissociated nitrogen were measured by an EPR spectrometer. Nitrogen molecules do not contain unpaired electron spins and, hence, do not contribute to the EPR signal. Since the spin J of nitrogen nucleus is nonzero, the EPR line of the ^{14}N atoms ($J = 1$) is split into three hyperfine components and ^{15}N ($J = 1/2$) into two components. Measurements were performed on a Bruker ESP-300 spectrometer with a cylindrical microwave cavity ER 4115 OD operating in the TE_{011} mode. The first derivative dI/dH of absorption signal was recorded. The concentration ratio of the ^{14}N and ^{15}N atoms in the nitrogen flow passing through the cavity was determined from the comparison of the results of double integration of the recorded lines. To determine the absolute concentrations of the ^{14}N and ^{15}N atoms, the spectrometer was calibrated against the EPR signal of molecular oxygen [7].

RESULTS AND DISCUSSION

In the experiments, the concentrations of the ^{15}N and ^{14}N isotopes in the nitrogen flow passing through the EPR cavity were measured at various flow rates V. The positions of EPR lines were in excellent agreement with the g factor ($g = 2.002$) and the hfs constants $A_{14} = 10.5$ MHz and $A_{15} = 14.7$ MHz, which coincided with the literature data (see, e.g., [8]) with a good accuracy. The observation of the typical pattern of hfs splitting for different isotopes, together with the coincidence of the quantitative parameters with their expected values, practically excludes the erroneous interpretation of the experimental results. Figure 2 presents a fragment of the EPR spectrum with the downfield components of the ^{14}N triplet and the ^{15}N doublet, which are offset from the signal center by $\delta H_{15} = 2.64$ G and $\delta H_{14} = 3.76$ G. The lines are asymmetrical because of a slight

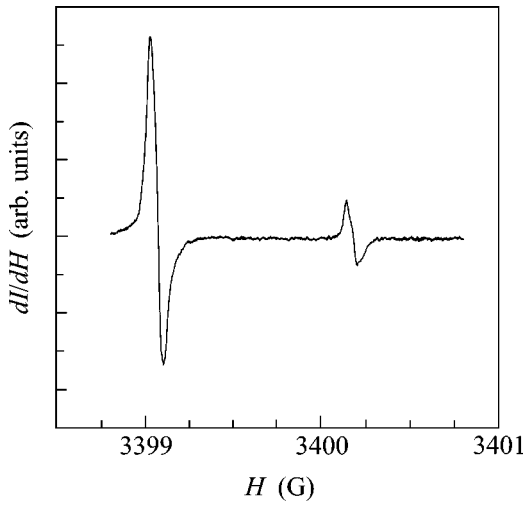


Fig. 2. ^{14}N and ^{15}N EPR spectrum components for the flow rate $V = 11.5$ cm/s, pressure $P = 4$ torr, and specific energy contribution $\varepsilon = 0.1$ eV per molecule. Concentration $^{14}N = 2.5 \times 10^{12}$, enrichment parameter $\beta = 34$, center of the spectrum is at $H_0 = 3402.8$ G, and frequency 9.55 GHz.

inhomogeneity of the field produced by the permanent magnet across the cavity height.

The measurement results were used to determine the $1/V$ dependence of the ^{15}N-isotope-enrichment parameter $\beta = (^{15}N/^{14}N)(^{15}N_0/^{14}N_0)$ for the atomic component, where $^{15}N_0/^{14}N_0 = 0.00366$ is the natural concentration ratio of the nitrogen isotopes. The variable $1/V$ was chosen as an argument because it is proportional to the gas-flow time along any section of the discharge and measuring tubes. The dependence of β on $1/V$ is shown in Fig. 3. One can see from this figure that the enrichment of nitrogen atomic component with the ^{15}N isotope at high flow rates is close to unity and increases drastically with a decrease in the flow rate. The experimentally obtained β values are satisfactorily approximated by the function $\beta = 1.0\exp(39/V)$. Hereafter, V is in cm/s. In contrast to β, one has failed to satisfactorily approximate all experimentally measured ^{14}N concentrations by the exponential function of $1/V$. However, four experimental points obtained for close values $\varepsilon = (0.07–0.08)$ eV per a molecule and for different V are satisfactorily approximated by the exponential curve $^{14}N = 4.5 \times 10^{14}\exp(-62/V)$, as is seen from Fig. 3. Therefore, the energy contribution ε strongly influences the atomic concentration in the flow and has only little effect on the ^{15}N-isotope enrichment. The exponential decrease of the atomic concentration and the increase in the enrichment parameter β with increasing $1/V$ and, hence, with an increase in the gas-flow time through the measuring tube allows the assumption to be made that the ^{15}N-isotope enrichment is due to atomic recombination at the tube walls in the post-discharge zone. For pure nitrogen and at a pressure of a few torr in the tubes with i.d. less than 2 cm, atomic recombination in the

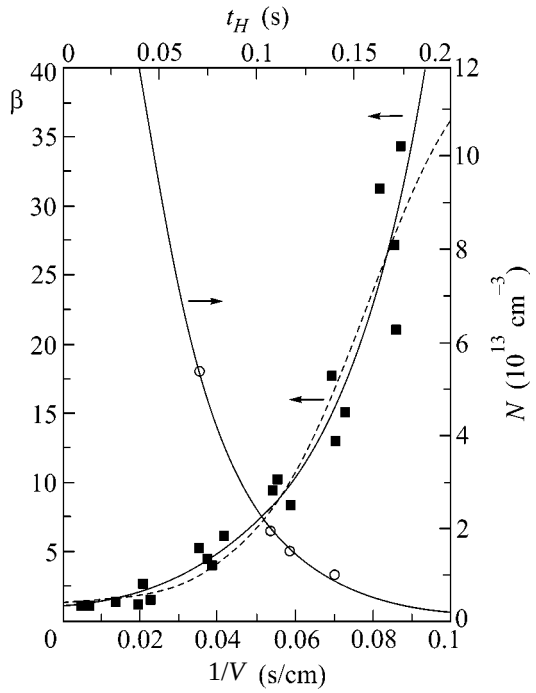


Fig. 3. Dependence of the enrichment parameter β (■) and the ^{14}N concentration (○) on 1/V and on the flow passage time t_H through the hot section of measuring tube. Approximation by exponential functions is shown by solid lines, and approximation by the function $\beta = [1 + 0.005 \times 45\exp(t_H/0.03)]/[1 + 0.005\exp(t_H/0.03)]$ is shown by the dashed line.

post-discharge zone mainly proceeds at the tube surface in the chemisorption centers occupied by the N atoms, after the other physisorbed N atoms or N atoms from the gas phase [9] encounter them. In this model, the surface recombination probability γ is independent of atomic mass. According to [9], if the radial diffusion time in gas is much shorter than the recombination time at the surface, the equation for recombination kinetics of nitrogen atoms into molecules at the surface of a tube with radius R has the form $dN/dt = -(\gamma u N)/2R$, where u is the mean thermal atomic velocity in gas. The integration of this equation gives the dependence of the ^{14}N and ^{15}N concentrations on the time t during which the atoms efficiently recombine at the tube surface:

$$^{15}\mathrm{N} = {}^{15}\mathrm{N}_0\exp[-(t\,{}^{15}\gamma\,{}^{15}u)/2R], \quad (1)$$

$$^{14}\mathrm{N} = {}^{14}\mathrm{N}_0\exp[-(t\,{}^{14}\gamma\,{}^{14}u)/2R]. \quad (2)$$

Setting $^{15}\gamma = {}^{14}\gamma = \gamma$ and dividing the first equation by the second, one obtains

$$\beta = ({}^{15}\mathrm{N}/{}^{14}\mathrm{N})({}^{15}\mathrm{N}_0/{}^{14}\mathrm{N}_0) = \exp[t\gamma\,{}^{14}u(1 - {}^{15}u/{}^{14}u)/2R]. \quad (3)$$

In our experiment, the region of intense high-temperature recombination of atoms escaping the discharge zone was at the opening section of measuring tube, which was covered with a brass connection electrode of length $L_H \cong 2$ cm (Fig. 1) and strongly heated upon the discharge. It is known that the recombination probability at the Pyrex surface increases by more than two orders of magnitude with increasing temperature from 400 to 1000 K [10]. If one takes the same temperature dependence for γ of the quartz surface, the exponent in Eqs. (1) and (2) at the hot section of the measuring tube will be an order of magnitude larger than the exponent for the room-temperature tube section. Then, neglecting the atomic recombination at room temperature, one can obtain from Eq. (2) the following expressions for the atomic concentration and enrichment parameter in the EPR cavity:

$$^{14}\mathrm{N} = {}^{14}\mathrm{N}_0\exp(-t_H/\tau_H), \quad (4)$$

$$\beta = \exp(t_H/{}^{14}\tau_H - t_H/{}^{15}\tau_H), \quad (5)$$

where $t_H = L_H/V_H$; $^{14}\tau_H = 2R/(\gamma\,{}^{14}u)_H$ and $^{15}\tau_H = 2R/(\gamma\,{}^{15}u)_H$ are, respectively, the flow time and the characteristic recombination rate in the hot section of measuring tube and V_H is the flow rate in the hot section. A numerical estimate obtained using Eq. (5) for the enrichment parameter gives $\beta = 1.4$. This maximal enrichment can be achieved in our experiments if atoms recombine at the surface, where the isotope-selective recombination is determined by the difference ($^{14}u - {}^{15}u$) between thermal velocities.

The high isotope enrichment obtained in our experiment can be explained by the nonequilibrium vibrational exchange [1], because the upper vibrational levels are mainly populated by the molecules containing ^{15}N atoms. In this case, the dissociation reactions of the vibrationally excited molecules or the atom–molecule isotope exchange reactions involving the nitrogen molecules mainly at the upper vibrational levels [11] should afford isotope-enriched nitrogen atoms. The enrichment cannot be higher than the isotopic selectivity in the population of high vibrational levels of the nitrogen molecule where these processes occur. The calculations of the enrichment for the products of reaction proceeding with the participation of the upper vibrational levels in the ground electronic state of the nitrogen molecule at a gas temperature higher than 300 K suggest that β does not exceed eight [6], which is four times smaller than the maximal value obtained for β in our experiment. However, if the dissociation or atom–molecule exchange proceeds mainly from the upper vibrational levels of electronic excited states, the maximal ^{15}N-isotope enrichment of the reaction products can be higher than eight. For example, at the start of the post-discharge zone in the region of short-lived afterglow, high populations of vibrational levels (up to $v = 10$) were observed for the $N_2(B^3\Pi_g, v)$ electronic exited state [12–14]. This electronic state, according to [12], appears in the near-afterglow zone upon the inter-

action of vibrationally excited molecules in the ground electronic state $N_2(X^1\Sigma_g^+, v_2)$ and vibrationally excited molecules in the excited electronic state $N_2(A^3\Sigma_u^+, v_1)$:

$$(N_2(A^3\Sigma_u^+, v_1) + N_2(X^1\Sigma_g^+, v_2)) = N_2(B^3\Pi_g, v) + N_2(X^1\Sigma_g^+, v_3), \quad (6)$$

where $v_2 > v_3$. In this case, the ^{15}N isotopes dominate the formation of $N_2(B^3\Pi_g, v)$ state, because the upper vibrational levels v_1 and v_2 of the reactants are enriched with this isotope as a result of the nonequilibrium vibration–vibration exchange. Then the upper vibrational levels of the $B^3\Pi_g$ state are additionally enriched with the ^{15}N isotope in the course of nonequilibrium vibration–vibration exchange, as compared to the lower-lying levels. As a result, the dissociation reaction $N_2(B^3\Pi_g, v > 12) = 2N(^4S)$ with a rate of $((4 \times 10^3–3 \times 10^5)T$ s^{-1} (T is the gas temperature in K) [14] in the post-discharge zone should yield atoms that are heavily enriched with the ^{15}N isotope. Since the number of atoms formed in the post-discharge zone is much smaller than in the discharge, this mechanism can have an appreciable effect on the enrichment of the nitrogen atomic component only after the majority of atoms produced in the discharge zone have recombined. This is favored by the position of the high-temperature recombination zone at the start of the measuring tube. The experimentally measured atomic concentration in a flow passing through the EPR cavity is contributed by the atoms produced in the discharge and not recombined in the hot section of the tube, as well as by the atoms produced in the post-discharge zone. The ratio α of the concentration of atoms formed in the post-discharge zone to the maximal atomic concentration in the discharge depends on the passage time through the hot tube section in two ways. On the one hand, α must decrease with increasing t_H because the electronic and vibrational excitations of reactants (6) relax at the hot wall. On the other hand, it can increase with t_H because of an additional vibrational excitation of nitrogen by electrons escaping the discharge zone and because of a decrease in the concentration of atoms acting as vibrational relaxants. To estimate how the atomic concentration and the enrichment parameter β depend on t_H and, hence, on the flow rate, we assumed that the coefficient α, as the ^{15}N-enrichment parameter β_0 for the atoms produced in the post-discharge zone, vary with the flow rate much slower than $\exp(-t_H/\tau_H)$. Then, by adding to Eq. (4) the term accounting for the atom formation in the post-discharge zone, we obtain

$$^{14}N = {}^{14}N_0[\exp(-t_H/\tau_H) + \alpha]. \quad (7)$$

Similarly, we have

$$^{15}N = {}^{15}N_0[\exp(-t_H/\tau_H) + \alpha\beta_0]. \quad (8)$$

Here, τ_H was assumed to be the same for ^{15}N and ^{14}N. After dividing Eq. (8) by Eq. (7), we obtain $\beta = [1 + \alpha\beta_0\exp(t_H/\tau_H)]/[1 + \alpha\exp(t_H/\tau_H)]$. This dependence satisfactorily approximates the experimental points given in Fig. 3 for $\alpha = 0.005$, $\beta_0 = 45$, and $\tau_H = 0.03$ s. This signifies that the nitrogen atoms should form in the post-discharge zone with a concentration of 0.5% of the atomic concentration in the discharge and with the ^{15}N-isotope content 45 times exceeding the natural abundance. Such a high enrichment parameter is not limiting, because the isotope enrichment of the reaction products of vibrationally excited nitrogen molecules in the vibration–vibration exchange may increase with lowering gas temperature below room temperature [2, 6].

This mechanism can also be used to explain the observed [5] strong decrease in the isotope enrichment of the atomic component in the nitrogen flow upon the addition of a minor oxygen impurity, which, as reported in [5], strongly changed the recombination rate of nitrogen atoms at the tube surface near the discharge zone and played the same role as a change in the flow rate in our experiments.

CONCLUSIONS

It is shown that the processes occurring in the post-discharge zone must be taken into account when studying the isotope enrichment of the products of reactions initiated by electric discharge in a gas flow.

More than 30-fold enrichment with the ^{15}N isotope was experimentally observed for the nitrogen atoms formed in the post-discharge zone of a pulsed discharge. This effect was explained using the model of two-step ^{15}N-isotope enrichment of the high-lying vibrational levels of the excited electronic state $N_2(B^3\Pi_g, v)$ followed by the dissociation of this state at $v > 12$.

The observation of a high ^{15}N-isotope enrichment of the nitrogen atomic component, together with the simplicity of the apparatus used, allows one to believe that a new promising method of separating nitrogen isotopes can be designed on the basis of this effect.

REFERENCES

1. É. M. Belenov, E. P. Markin, A. N. Oraevskiĭ, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 196 (1973) [JETP Lett. **18**, 116 (1973)].
2. N. G. Basov, É. M. Belenov, V. A. Isakov, *et al.*, Zh. Éksp. Teor. Fiz. **68**, 2032 (1975) [Sov. Phys. JETP **41**, 1017 (1975)].
3. T. J. Manuccia and M. D. Clark, Appl. Phys. Lett. **28**, 372 (1976).
4. V. M. Akulintsev, N. M. Gorshunov, S. E. Kupriyanov, *et al.*, Khim. Vys. Énerg. **12**, 554 (1978).

5. V. Kudrle, A. Talsky, and J. Janca, in *Proceedings of XXIV International Conference on Phenomena in Ionized Gases, ICPIG XXIV*, Warsaw (1999), Vol. 1, p. 167.
6. Sh. S. Mamedov, Tr. Fiz. Inst. im. P.N. Lebedeva, Akad. Nauk SSSR **107**, 3 (1979).
7. S. Krongelb and H. W. P. Standberg, J. Chem. Phys. **31**, 1196 (1959).
8. M. A. Heald and R. Beringer, Phys. Rev. **96**, 645 (1954).
9. B. Gordiets, C. M. Ferreira, and J. Nahorny, J. Phys. D: Appl. Phys. **29**, 1021 (1996).
10. P. A. Sa and J. Loureiro, J. Phys. D: Appl. Phys. **30**, 2323 (1997).
11. A. Lagana, E. Carcia, and L. Ciccarelli, J. Phys. Chem. **91**, 312 (1987).
12. P. Supiot, D. Blois, and S. De Benedicts, in *Proceedings of XXIV International Conference on Phenomena in Ionized Gases, ICPIG XXIV*, Warsaw (1999), Vol. 1, p. 247.
13. M. Capitelli, C. Gorse, and A. Ricard, in *Nonequilibrium Vibrational Kinetics*, Ed. by M. Capitelli (Springer, Berlin, 1986; Mir, Moscow, 1989).
14. L. S. Polak, D. I. Slovetskiĭ, A. D. Urbas, *et al.*, in *Plasma Chemistry*, Ed. by B. M. Smirnov (Atomizdat, Moscow, 1978), No. 5, p. 258.

Translated by V. Sakun

JETP Letters, Vol. 77, No. 4, 2003, pp. 167–171. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 197–201.

Refraction of Thermalized Electrons Emitted Ballistically into Vacuum from p^+-GaAs-(Cs,O)

V. V. Bakin[1], A. A. Pakhnevich[1], S. N. Kosolobov[2], H. E. Scheibler[2], A. S. Jaroshevich[2], and A. S. Terekhov[1, 2]

[1]*Novosibirsk State University, ul. Pirogova 2, Novosibirsk, 630090 Russia*

[2]*Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 13, Novosibirsk, 630090 Russia*

Received January 21, 2003

The refraction of thermalized photoelectrons emitted into vacuum from p^+-GaAs-(Cs,O) with a negative electron affinity was observed. The refraction is caused by a jump in electron effective mass at the semiconductor–vacuum interface. It was observed only for a small group of electrons that were emitted ballistically into vacuum without scattering of the tangential momentum component at the interface. *© 2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 79.60.Bm

The surfaces of p^+-GaAs and some other semiconductors covered with a cesium monolayer with oxygen addition undergo transition to the effective negative electron-affinity (NEA) state [1]. In this state, the surface energy barrier for electrons decreases to an extent that the vacuum level goes below the conduction-band bottom in the semiconductor bulk, so that even slow photoelectrons thermalized near the conduction-band bottom have a high probability of escaping into vacuum. Semiconductors with NEA have long and successfully been used as photocathodes in photodetectors and as sources of spin-polarized electrons. Despite the long history of studies, the physical phenomena accompanying the electron transition from a semiconductor with NEA to vacuum still remain to be understood. The simplest emission model infers the ballistic electron escape into vacuum without scattering of the tangential momentum component and the energy at the interface. In this model, the electron energy distribution width in vacuum must be on the order of kT, and the electron trajectories must be concentrated in a narrow range of angles $\theta \le \theta_0$ to the surface normal. The fact that θ_0 is much smaller than $\pi/2$ is caused by the electron-wave refraction by the interface because of the jump in effective mass [1]. The experimental studies of energy and angular distributions of emitted electrons showed that the predictions of the simplest model were not realized. The measured distributions showed up as broad bands that fully fill the allowed energy intervals from the conduction-band bottom in the semiconductor bulk to the vacuum level and the allowed intervals of angles $\theta \le \pi/2$.

The large width of energy distributions was explained by the emission of phonons [2], while the absence of phonon structure in the distributions was explained by their smoothing out due to the patchy character of the work function for the p^+-GaAs-(Cs,O) surface and by the insufficient resolution of electron spectrometers. The broadening of angular distributions was explained in [3] by the suppression of reflection as a result of the intense scattering of the tangential component of photoelectron momentum at the interface. The macroscopic surface roughness might be one more reason [4]. Moreover, the measured angular distributions could broaden under the influence of weak uncontrolled electric and magnetic fields caused by the patchy character of the work function and by the remanent magnetization of the spectrometer parts and vacuum assembly.

Although these arguments could explain any discrepancies between the theoretical and experimental results, they unfortunately have proved nothing. For this reason, the questions of the photoemission model, of the main energy- and momentum-scattering mechanisms for the thermalized photoelectrons at the p^+-GaAs-(Cs,O)–vacuum interface and of the limiting attainable characteristics of emitters with NEA, are still open. It was only recently that we observed, while studying the photoelectron energy distributions at 4.2 K using an electron spectrometer with a uniform retarding field, a fine structure consisting of an electron elastic-scattering peak and its phonon replicas [5]. This has given proof that the thermalized electrons were emitted from the size-quantization band bottom in the near-surface region of the p^+-GaAs space charge, while the electron interaction with optical phonons was the main mechanism of electron energy loss upon the transition from a semiconductor into vacuum. The collisions with free holes, although dominating the electron energy scattering in the bulk of heavily doped p-type GaAs [6],

0021-3640/03/7704-0167$24.00

proved to be less important in the near-surface region. It has also become clear that the uniformity of the work function of GaAs with NEA and the resolution of a spectrometer with a uniform retarding electric field are sufficient for separating a relatively small group of electrons that are emitted from semiconductor into vacuum without energy loss. Some of these electrons may escape the semiconductor without elastic momentum scattering, i.e., ballistically. The angular distribution of the ballistically emitted electrons should be due to refraction. Consequently, the study of the photoelectron angular distribution can be used to observe this phenomenon. Since the electron refraction due to the jump in effective mass at the interface is a universal phenomenon and influences the charge-carrier transfer through any heterojunctions [7], its study is of interest not only in photoemission physics.

In this work, we have implemented a new high-resolution method for measuring the photoelectron energy and angular distributions, which allows the separation of a group of monoenergetic electrons with the angular distribution halfwidth of few degrees. We believe that we have observed ballistically emitted thermalized photoelectrons, whose angular distribution is formed due to the refraction effect.

A GaAs emitter on a glass substrate was used in our experiments [8]. The photoemission from the gallium arsenide layer was excited by its illumination through the glass substrate. The GaAs surface roughness, which broadens the electron energy distributions, was reduced by chemicomechanical polishing and by lowering the maximal heating temperature to 800 K. The other details of preparing GaAs emitters were the same as described in [8]. As in [5], the spectrometer was based on a hand-made planar vacuum photodiode made from nonmagnetic materials and composed of the GaAs emitter on glass and a copper-covered glass anode, which were hermetically packed, parallel to one another, at the opposite ends of a cylinder made from alumina-based ceramics. The working diameters of the emitter and anode were 18 mm, and the distance between them was 1 mm. The degassing of anode and photodiode frame, the preparation of the surface of GaAs emitter, and their subsequent hermetic assembly using deformable indium gaskets was carried out in a three-chamber ultrahigh-vacuum setup with a base vacuum of $\leq 10^{-8}$ Pa.

The angular distributions of slow photoelectrons were measured by a new technique, which was recently proposed in [9] and based on measuring the photocurrent J_{ph} in a uniform retarding electric field E and the perpendicular magnetic field B. It was shown in [9] that the measured $J_{ph}(E, B)$ dependences can be used to reproduce the electron distribution function Φ in the total energy ε and polar angle θ. The advantage of this method is that, because of the short distance between the emitter and anode and the small size of the light spot on the emitter, the photoelectron trajectories are concentrated within a volume on the order of one cubic millimeter, where the uniformity of electric and magnetic fields can be achieved with a desired accuracy, while the uncontrolled distortions of electron trajectories are insignificant. The "opticomechanical" element of the photoelectron spectrometer included the photodiode, a pair of Helmholtz coils for producing a uniform magnetic field, and a stabilized semiconductor laser with a wavelength of 670 nm, whose radiation was transmitted to the photodiode through an optical fiber and produced a light spot with a diameter of ~1 mm on the surface of the GaAs emitter. The electronic blocks of the spectrometer included a digital lock-in amplifien for measuring the photocurrent and its derivatives with respect to electric field, the stabilized laser power sources and Helmholtz coils, and a source of bias and modulating voltages that were used for measuring the energy distributions. All electronic blocks were interfaced to a PC using standard cards.

Preliminary assessment of the quality of the fabricated emitters showed that, due to the reduced roughness of the GaAs surface, the elastic-tunneling peak in the electron energy distribution, which was observed previously only at 4.2 K [5], became "observable" at 77 K for all seven fabricated and inspected photodiodes. For this reason, the experiments in this work were conducted at liquid nitrogen temperature. As in [5], the electron spectrometer was calibrated against the position of a high-energy shoulder produced in the energy distribution by the photoelectrons emitted into vacuum with the initial energy corresponding to their kinetic energy in the conduction band at the instant of photoproduction. This energy was calculated using the known photon energy and the dispersion laws for the valence and conduction bands with allowance for their corrugation and nonparabolicity.

For the experiments on electron refraction, a photodiode with emitter showing the largest peak amplitude of elastic tunneling was chosen. The electron energy distribution $N(\varepsilon_\parallel)$ measured for this photodiode is presented in Fig. 1. Since the measurements were carried out in a uniform retarding field, the energy axis corresponds to the so-called photoelectron "longitudinal" energy, ordinarily denoted as $\varepsilon_\parallel$. One can see in the figure that the full width of energy distribution, as measured from the vacuum level E_{vac} to the conduction band bottom E_c in the semiconductor bulk, is approximately equal to 250 meV. It is also seen that the high-energy distribution edge shows a peak I_0 that is separated by approximately 15 meV from E_c. We have suggested [5] that the peak I_0 is due to the elastic tunneling of thermalized photoelectrons from the bottom of the upper size-quantization band in the near-surface region of the p^+-GaAs space charge into vacuum through the (Cs, O) layer. A weak I_{ph1} feature separated by 37 meV from the I_0 peak is its first phonon replica. The second replica was not observed for this diode because of its small amplitude and the overlap with a broad I_s band

corresponding to electrons emitting two and more optical phonons and to electrons scattered elastically at large angles to the normal to the surface. The nature of the I_v feature, which was observed for some photodiodes immediately above the vacuum level, remains to be understood.

To determine the electron angular distribution for the I_0 peak, $N(\varepsilon_\parallel)$ was measured as a function of the transverse magnetic field. When measuring $N(\varepsilon_\parallel, B)$, the blocking voltage was chosen so that the inelastically scattered electrons did not contribute to the measured current. Several measurements of $N(\varepsilon_\parallel^0, B)$ were made with various retarding voltages within 10 meV above the I_0 maximum at $\varepsilon_\parallel = \varepsilon_\parallel^0$. All these curves were essentially the same. This means that the angular distributions of all electrons at the I_0 peak coincide with a good accuracy and, hence, we deal with a single group of electrons. The $N(\varepsilon_\parallel^0, B)$ dependence, as measured for the retarding voltage cutting off electrons with $\varepsilon_\parallel < \varepsilon_\parallel^0 =$ 230 meV, is shown by circles in Fig. 2. From this figure, it follows that the electron turn in a magnetic field reduces $N(\varepsilon_\parallel^0, B)$ almost to zero in a field of 20 G. It is also seen that the decay rate of $N(\varepsilon_\parallel^0, B)$ in magnetic fields $B < 5$ G is appreciably higher than at $B > 10$ G. It is natural to assume that the main contribution to the $N(\varepsilon_\parallel^0, B)$ decay in the weak-field region comes from the field-induced turn of the ballistically emitted electrons, which are concentrated, due to refraction, in a narrow range of angles around the normal vector and, thus, are more sensitive to the transverse magnetic field. At $B >$ 10 G, the main contribution to $N(\varepsilon_\parallel^0, B)$ comes from the field-induced turn of the electrons emitted in a wide range of angles.

The $N(\varepsilon_\parallel^0, B)$ dependence was converted into the angular distribution using not the general method discussed in [9] but a simpler expedient. In essence, it amounted to fitting the approximate $N_{\text{cal}}(\varepsilon_\parallel^0, B)$ function calculated using model angular distributions with adjustable parameters to the measured $N(\varepsilon_\parallel^0, B)$ dependence. Inasmuch as the measured $N(\varepsilon_\parallel^0, B)$ dependence with a bend at $B \sim 7$ G evidenced the presence of two groups of electrons with different angular distributions, the electron angular distribution $\Phi(\theta)$ at the I_0 peak was taken in the form

$$\Phi(\theta) = \nu\Phi_b(\theta) + (1-\nu)\Phi_d(\theta), \qquad (1)$$

where the angular distribution function of ballistic electrons is denoted by $\Phi_b(\theta)$, the angular distribution function of electrons escaped into vacuum with momentum scattering is denoted by $\Phi_d(\theta)$, and ν determines the

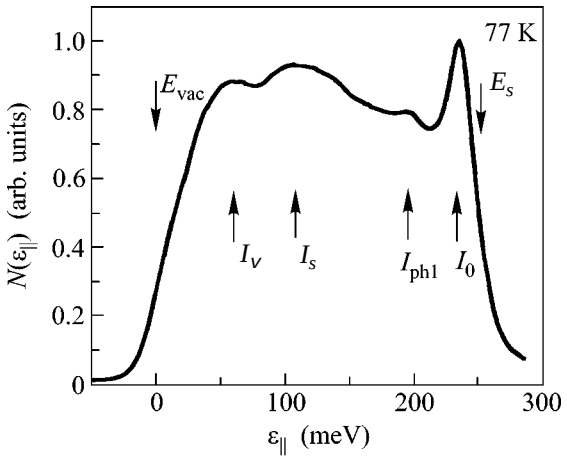


Fig. 1. Photoelectron distribution in longitudinal energy.

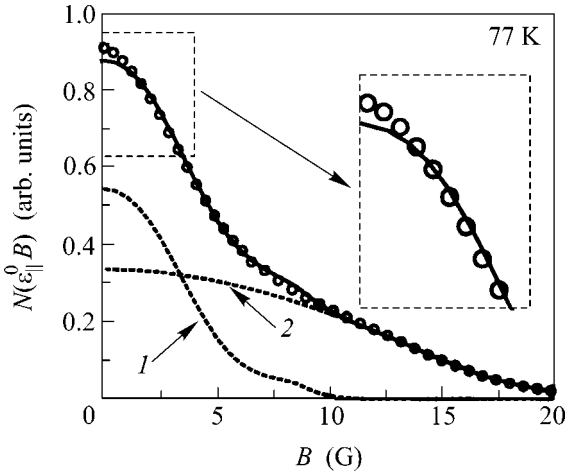


Fig. 2. Measured and calculated magnetic-field dependence $N(\varepsilon_\parallel^0, B)$.

fraction of ballistic electrons in the total flux of electrons escaping into vacuum without energy loss. The distribution functions $\Phi_b(\theta)$ and $\Phi_d(\theta)$ were taken in the same form,

$$\Phi_{b,d}(\theta) = \frac{C_{b,d}\exp[C_{b,d}(\cos\theta - 1)]}{2\pi[1-\exp(-C_{b,d})]}, \qquad (2)$$

and differed only in the values of adjustable parameters C_b and C_d. The distribution function $\Phi(\theta)$ was normalized to unity. The function $N_{\text{cal}}(\varepsilon_\parallel^0, B)$ was calculated by solving the equation of motion for electrons in crossed uniform electric and magnetic fields, with the initial conditions specified by the functions $\Phi(\theta)$ and $\rho(\varepsilon)$. The distribution $\rho(\varepsilon)$ in total electron energy at the I_0 peak, which was used in calculating $N_{\text{cal}}(\varepsilon_\parallel^0, B)$, did not

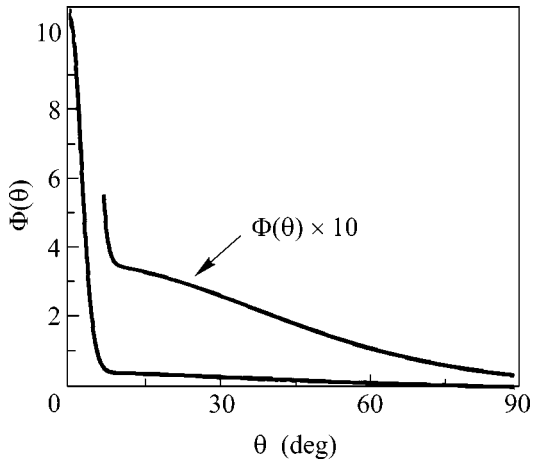


Fig. 3. Photoelectron angular distribution.

contain adjustable parameters and coincided with the shape of the I_0 peak. The unknown shape of the low-energy wing was taken to be mirror symmetric about the high-energy wing of the I_0 peak. The standard deviation of the calculated $N_{cal}(\varepsilon_\parallel^0, B)$ from the measured $N(\varepsilon_\parallel^0, B)$ was minimized by adjusting the parameters ν, C_b, and C_d. The best fit of the calculated to measured curve is shown in Fig. 2 by the solid line. The best fit was achieved for $\nu = 0.12$, $C_b = 530$, and $C_d = 2$. The systematic distinctions between these curves were the greatest at $B \le 2$ G and caused by the fact that the function $\Phi(\theta)$ given by Eqs. (1) and (2) did not account for the details of the actual angular distribution. Nevertheless, a good agreement between the measured and calculated curves in Fig. 2 is evidence that the angular distribution used in calculations is close to the actual distribution. The dotted lines in Fig. 2 are the contributions from the (curve *1*) ballistically and (curve *2*) diffusionally emitted electrons. One can see that the shape of the measured curve $N(\varepsilon_\parallel^0, B)$ in weak magnetic fields $0 \le B \le 5$ G is well described by the contribution only from the ballistic electrons, while the influence of electrons emitted with momentum scattering only gives a background that weakly depends on field. One can also see that the contribution from the ballistic electrons at $B \ge 10$ G is negligible and the experiment is well described only by the contribution from the diffusionally emitted electrons. The fact that the overlap between the ballistically and diffusionally emitted electrons to $N(\varepsilon_\parallel^0, B)$ is small in the region of weak and, especially, strong magnetic fields indicates that the obtained adjustable parameters are reliable. Using these parameters and Eqs. (1) and (2), we calculated the angular distribution of elastically emitted electrons. The results are shown in Fig. 3. From this figure, it is seen that the electron angular distribution consists of two prominent contributions, one of which decays rapidly at $\theta \le 10°$, and other fills the range of angles $0° \le \theta \le 90°$. By using the angular distributions for each group of electrons, we calculated the average transverse electron energies $\langle\varepsilon_\perp\rangle$ in these groups: $\langle\varepsilon_\perp\rangle_b \cong 1$ meV and $\langle\varepsilon_\perp\rangle_d \cong 120$ meV. The theoretical average transverse energy of electrons thermalized at 77 K and emitted ballistically into vacuum from the near-surface size-quantization state was calculated on the basis of the relations taken from [1] with allowance for the refraction effect, to arrive at $\langle\varepsilon_\perp\rangle_b = 0.45$ meV. This value is in good agreement with the measured value. At least, it is smaller than kT by a factor of seven and, in our opinion, cannot be explained without invoking the refraction effect. The distinction between the measured and calculated $\langle\varepsilon_\perp\rangle_b$ values is caused, in part, by the fact that the function $\Phi(\theta)$, which was used for the approximate description of the angular distribution of ballistically emitted electrons, does not account for the actual distribution. This follows from small systematic distinctions between the measured and calculated $N(\varepsilon_\parallel^0, B)$ dependences at weak magnetic fields, as is clearly seen in the inset in Fig. 2. One can see that the measured $N(\varepsilon_\parallel^0, B)$ function decreases with increasing field faster than the calculated function. This signifies that the actual angular distribution of ballistically emitted electrons is narrower than the width of the model function $\Phi(\theta)$ with the found adjustable parameters, and the actual value of $\langle\varepsilon_\perp\rangle_b$ is smaller than our value of 1 meV. Moreover, it is quite probable that the electron distribution near the size-quantization band bottom is not fully thermalized, because electrons were emitted into vacuum or scattered to the lower-lying size-quantization band faster than they acquired lattice temperature. A certain contribution to the broadening of the angular distribution of ballistically emitted electrons is also due to the angular resolution limit of our spectrometer. We have not determined this limit as yet. The reason for a relatively high large-angle scattering probability, as a result of which a portion of electrons is emitted along the surface, is also unclear. Nevertheless, we assume that the presence of a small group of electrons emitted from GaAs-(Cs,O) with NEA and having the angular distribution with a halfwidth of few degrees is firmly established. Since the only known explanation of the origin of such a narrow distribution amounts to invoking the refraction effect, this phenomenon can be considered proved.

We are grateful to D. Shwalm, A. Wolf, and D. Orlov for helpful discussions and support. This work was supported by the Russian Foundation for Basic Research (project no. 01-02-16912), the Ministry of Education of the Russian Federation (federal program "Integration of Science and Higher Education of Russia") via Novosibirsk State University (contract no. I765/785), and the Ministry of Industry and Science of the Russian Federation (federal program "Surface Atomic Structures," contract no. 40.012.1.1.1151).

REFERENCES

1. R. L. Bell, *Negative Electron Affinity Devices* (Clarendon Press, Oxford, 1973; Énergiya, Moscow, 1978).
2. L. W. James and J. L. Moll, Phys. Rev. **183**, 183 (1969).
3. G. Vergara, A. Herrera-Gómez, and W. E. Spicer, J. Appl. Phys. **80**, 1809 (1996).
4. B. J. Bradley, M. B. Alenson, and B. R. Holeman, J. Phys. D **10**, 111 (1977).
5. D. A. Orlov, V. É. Andreev, and A. S. Terekhov, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 220 (2000) [JETP Lett. **71**, 151 (2000)].
6. M. I. D'yakonov, V. I. Perel', and I. N. Yassievich, Fiz. Tekh. Poluprovodn. (Leningrad) **11**, 1371 (1977) [Sov. Phys. Semicond. **11**, 805 (1977)].
7. J. Smoliner, R. Heer, C. Eder, *et al.*, Phys. Rev. B **58**, 7516 (1998).
8. Yu. B. Bolkhovityanov, B. V. Morozov, A. G. Paulish, *et al.*, Pis'ma Zh. Tekh. Fiz. **16** (4), 25 (1990) [Sov. Tech. Phys. Lett. **16**, 253 (1990)].
9. V. E. Andreev, A. L. Bukhgeim, and A. S. Terekhov, J. Inv. Ill-Posed Probl. **7**, 427 (1999).

Translated by V. Sakun

JETP Letters, Vol. 77, No. 4, 2003, pp. 172–177. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 202–207.

Many-Particle Interaction in Tunneling Spectroscopy of Impurity States on the InAs(110) Surface¶

P. I. Arseev[1], N. S. Maslova[2], V. I. Panov[2], S. V. Savinov[2], and C. van Haesendock[3]

[1] *Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 117924 Russia*

e-mail: ars@lpi.ru

[2] *Department of Physics, Moscow State University, Moscow, 119899 Russia*

[3] *Laboratorium voor Vaste-Stof Fysica en Magnetisme, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium*

Received December 17, 2002

We report on the direct observation by low-temperature scanning tunneling microscopy and scanning tunneling spectroscopy of the *d* orbitals of a Mn *p*-type impurity appearing on a cleaved InAs(110) surface. We show that the crucial interplay between nonequilibrium charging effects and many-particle interaction leading to Coulomb singularities provides a consistent description of the experimental results. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 68.55.Ln; 68.37.Ef

The invention of scanning tunneling microscopy (STM) and spectroscopy (STS) was expected to provide a reliable tool to determine the local surface electronic structure with a spatial resolution comparable to the interatomic distance. For several surfaces, STM and STS measurements indeed provided crucial information about the local electronic density of states, e.g., in the case of the Si (7×7) reconstruction [1]. For other surfaces, however, discussions are still going on about the unusual tunneling characteristics that are observed. For these surfaces, more experimental data, as well as an improved understanding of nanometer-scale tunneling phenomena, are clearly needed [2]. In particular, this applies to low-temperature STM and STS measurements because of the possible presence of a nonequilibrium electron distribution in the tunnel contact area, a strong modification of the initial electronic spectrum due to tip–sample interactions, and many-particle correlation effects [3]. Qualitative changes in understanding of tunneling phenomena in STM/STS experiments revealed the significant importance of relaxation processes and nonequilibrium interaction effects in the behavior of the tunneling characteristics of ultrasmall junctions. Now it is clear that a nonequilibrium electron distribution in the tunneling contact area and interaction between tunneling particles give rise to a strong modification of the initial local density of states and tunneling conductivity spectra. Moreover, the tunneling amplitude itself can also be altered by Coulomb interaction of conduction electrons in a metallic tip with nonequilibrium charges localized at the impurity states. In this case, a Coulomb singularity in tunneling current near the threshold value of bias voltage can be observed in STM/STS experiments. In the present paper, we show that the crucial interplay between the effects connected with changes of initial density of states due to nonequilibrium charge distribution and the effects associated with modification of tunneling amplitude due to many-particle Coulomb interaction can result in new unusual features in tunneling characteristics.

Here, we present the results of low temperature STM and STS measurements on Mn-doped InAs single crystal surfaces that are obtained by *in situ* cleavage of single crystals along the (110) plane. There exists a significant difference in the local electronic properties of narrow- and wide-band-gap III–V compound semiconductors [4]. Narrow-band-gap materials such as InAs have been less intensively investigated as compared to wide-band-gap ones such as GaAs. We therefore decided to perform a more detailed study of the behavior of individual dopant atoms on an InAs surface. In STM and STS images, we observed a remarkable crosslike feature within a particular range of voltages applied between the sample and the STM tip. A similar structure was recently observed with low-temperature STM and STS for a Cr(001) surface and was explained in terms of an orbital Kondo resonance in Cr [5]. We show that the physical process causing the formation of the crosslike surface structures in our semiconducting samples is very different from the mechanism that was proposed to explain the appearance of the crosslike features in a transition metal.

Our STM and STS data were obtained with a homebuilt low-temperature microscope with an *in situ* sample cleavage mechanism. The microscope operates at liquid helium temperature and in the presence of magnetic fields up to 7 T [6]. The scanning of the STM tip (PtIr wire cut with scissors) is controlled by commer-

¶This article was submitted by the authors in English.

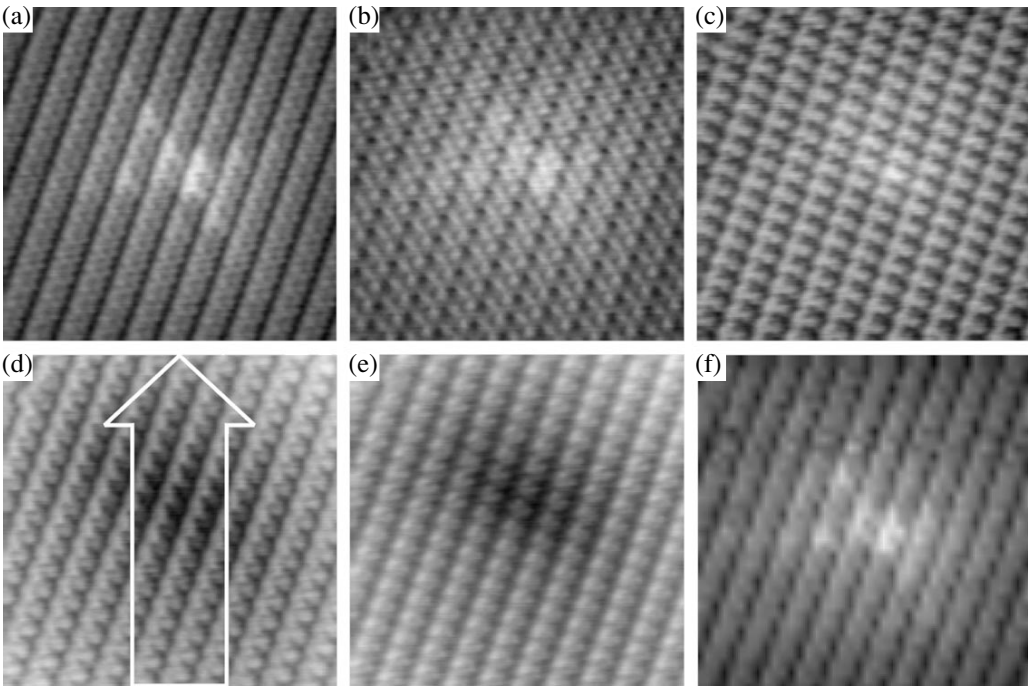


Fig. 1. Low-temperature STM images of a Mn dopant atom on the InAs(110) surface obtained at different values of the applied sample bias voltage: (a) $V = -0.5$ V, (b) $V = -0.7$ V, (c) $V = -1$ V, (d) $V = 0.5$ V, and (e) $V = 0.7$ V. In (f) we show a topographic image at $V = -0.5$ V during STS measurements. The white arrow in (d) indicates the stripe where we collected the STS data for the 1D cross section shown in Fig. 3.

cially available control electronics (Omicron Gmbh). The samples are *p*-type InAs single crystals doped with Mn atoms. The concentration of dopant atoms is relatively high and is about 5×10^{17} cm^{-3}. Mn atoms are known to act as acceptors in InAs, where they are included as substitutional atoms occupying the In sublattice. The samples are cleaved along the (110) natural cleavage plane after the temperature of the system reaches 4.5 K. The residual partial pressure for all gases, in particular for oxygen, is very low, implying that the cleaved surfaces remain atomically clean for about two weeks.

The STM/STS measurements presented below are obtained in two different modes of operation. The first mode is the usual constant-current mode STM, while the second mode is the current-imaging tunneling spectroscopy (CITS) mode, which is derived from the STS measurements. In the CITS mode, the STM tip scans a grid (typically 40×40 points), and at every point of the grid the feedback loop is interrupted for some time, allowing measurement of the current–voltage characteristic for this point. Next, the feedback is restored and the tip is moved to the next point of the grid. Therefore, one obtains the following experimental data set: a 2D array of points containing topographic information and a 2D array of data files containing the local $I(V)$ curves. Previous investigations have focused on the best way to interpret the second "spectroscopic" data set. The most commonly used measure of the local density of states is the normalized conductivity $(dI/dV)/(I/V)$ [7].

Because of the digital method used for calculating the derivative dI/dV and because of the normalization, spatial averaging of the spectroscopic data is required to obtain an acceptable signal-to-noise ratio. This is achieved by calculating average $I(V)$ curves, where the average includes the point of the grid as well as the surrounding points, implying that the spatial resolution is somewhat reduced. After this averaging procedure, it is possible to "slice" the 2D array of averaged $I(V)$ curves at specific values of the applied bias voltage, resulting in a 2D distribution of the tunneling current for the grid. In a similar way, one can calculate the average normalized conductivity for every point of the grid, and again one can "slice" the grid of normalized conductivity data at specific values of the applied bias voltage, resulting in a 2D array of normalized and averaged data. A similar procedure can also be used for a linear cross-section of an STM image (1D array of normalized conductivity data). As we will explain in more detail below, due to the complexity of both the experimental data and the physical processes occurring in nanometer-scale junctions, different ways of representing the data are required to draw reliable conclusions.

In Fig. 1, we show the results of topographical STM measurements on surface area of 6×6 nm in the vicinity of a Mn dopant atom at different values of the

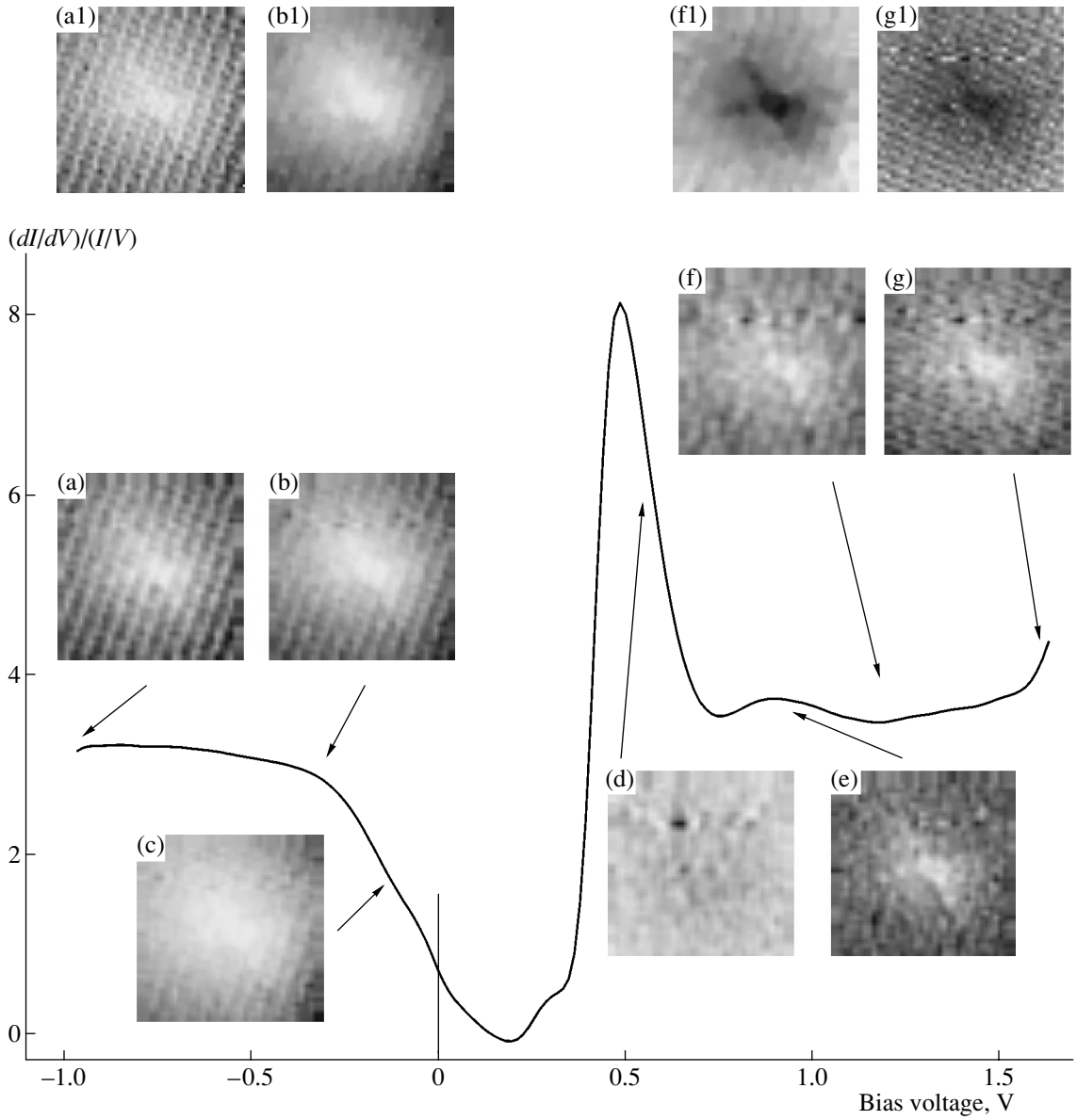


Fig. 2. Voltage dependence of the normalized tunneling conductivity $(dI/dV)(I/V)$ averaged over the whole surface area corresponding to the STM images shown in Fig. 1. The insets (a)–(g) correspond to 2D slices of the normalized tunneling conductivity data at different values of the bias voltage. The bias voltage used for the slicing is marked by the arrows. The insets (a1)–(g1) correspond to 2D slices of the $I(V)$ data. The images based on the $I(V)$ data are placed above the corresponding images obtained from the differential conductivity data.

applied bias voltage. For negative voltages, we observe a remarkable crosslike feature protruding from the atomic lattice background around the impurity atom: –0.5 V (Fig. 1a), –0.7 V (Fig. 1b), and –1 V (Fig. 1c). The brightness of the crosslike feature decreases when increasing the bias voltage. At positive sample bias voltage, a more spherical depression is observed around the impurity atom, and the depression becomes more localized when increasing the bias voltage: 0.5 V (Fig. 1d) and 0.7 V (Fig. 1e). In order to illustrate the good quality of our STS measurements, which are discussed in more detail below, we show in Fig. 1f the STM topography obtained from the STS data. The image corresponds to a 2D grid with 40×40 current values measured at $V = -0.5$ V. Atomic details and the crosslike feature around the impurity atom can clearly be observed.

In Fig. 2, we plot the normalized voltage dependence of the differential conductivity. This result has been averaged over the 40×40 points of the 2D grid (see Fig. 1f). The images that have been inserted in Fig. 2 correspond to 2D "slices" of the normalized tun-

neling conductivity at the bias voltages indicated by the arrows (Figs. 2a–2g). In order to stress the difference between the information contained in the normalized conductivity data and in the original $I(V)$ curves, we have added to Fig. 2 "slices" of the $I(V)$ data for different bias voltages Figs. 2a1–2g1. For positive voltages, the feature around the Mn dopant atom appears as a protrusion in the images based on the normalized conductivity, while the feature appears as a depression in the images based on the $I(V)$ data.

Next, we come to the physical origin of the bright crosslike feature, which is the key result of our present STM and STS investigation. For negative polarity of the sample voltage, we observe a bright crosslike feature when $-1\ \mathrm{V} < V < -0.5\ \mathrm{V}$, which corresponds to the electron density of the Mn acceptor *d*-orbital state hybridized with the InAs lattice states in the valence band (Figs. 2a–2c). For small bias voltages, the Mn impurity *d*-orbital state is close to being singly occupied. When the value of the applied voltage approaches the acceptor energy level ε, one expects an enhanced tunneling conductivity in the voltage range $|eV - \varepsilon_d| < \Gamma$ when moving the STM tip in the vicinity of the Mn atom. The parameter Γ corresponds to the lifetime broadening of the energy level ε due to the coupling with the semiconductor valence-band states. The enhanced tunneling current and local tunneling conductivity near the Mn impurity therefore result from the increased semiconductor local density of states formed in the vicinity of a Mn impurity. This density of states directly reflects the nonspherical symmetry of the localized Mn impurity *d* level.

Another factor causing an enhanced tunneling current and conductivity may be related to the nonequilibrium many-particle Coulomb interactions in the nanoscale tunneling contact. The many-particle Coulomb interaction of the conduction electrons residing at the STM tip apex with the positively charged impurity state is expected to result in a power-law singularity of the tunneling current and conductivity near a threshold value of the applied bias voltage [3]. When the tunneling rate Γ_t from the STM tip to the localized impurity state strongly exceeds the electron transition rate Γ due to the coupling of the *d*-orbital state to the semiconductor continuum states, a nonequilibrium positive charge will start to build up near the impurity atom when the negative applied bias is close to the threshold value $|eV - \varepsilon| < \Gamma$.

The Coulomb interaction between the positively charged impurity with the conduction electrons gives rise to a power-law dependence of the $I(V)$ characteristic, with a negative power law exponent $a = W\nu$, where W is the averaged value of the impurity Coulomb potential and ν is the electron density of states for the metallic tip [3]:

$$I(V) = \frac{\gamma_t}{2W\nu}\left[\frac{D^2}{(eV-\varepsilon_d)^2 + (\gamma_t+\gamma)^2}\right]^{W\nu}\sin(2W\nu\phi), \quad (1)$$

where $\phi = \mathrm{arccot}\left(\frac{eV-\varepsilon_d}{\gamma+\gamma_t}\right)$. Therefore, one expects an additional enhancement of the tunneling current near the impurity atom when the applied voltage approaches its threshold value. For bias voltages $-0.5\ \mathrm{V} < V < 0.5\ \mathrm{V}$, the bright crosslike feature is smeared out and the spatial structure of the local density of states is no longer visible in the images obtained from the STS data.

In Fig. 2, a strongly enhanced tunneling conductivity is observed in the vicinity of the impurity atom when the applied bias voltage approaches the conduction band edge. Such a behavior of the local tunneling conductivity is consistent with nonequilibrium charging effects [8]. The position of the impurity energy level depends on the electron filling numbers because of the onsite Coulomb interaction between electrons of opposite spins that are localized on the impurity. When varying the applied bias voltage and the tunneling current, one can vary the filling numbers for the localized electrons and hence shift the value of the impurity energy level. This mechanism is expected to become very effective near the band gap edges, resulting in considerable changes of the tunneling current when varying the bias voltage. As shown in more detail in [8], the dependence of the impurity energy on the electron filling numbers induces a nonmonotonic behavior of the tunneling conductivity. Electron scattering by the impurity potential causes the appearance of localized states near the band gap edges in the surface electronic structure. For a positively charged impurity, this split state might be slightly below the bottom of the conduction band, and for a negatively charged impurity the split state may be slightly above the valence band edge.

It is clear that the peak in the tunneling conductivity near the conduction band edge cannot be associated with a Kondo effect. In contrast to the experimental observations reported in [5], no specific features can be observed near the Fermi level, i.e., for bias voltages smaller than 100 meV. In the case of a Kondo resonance, an asymmetric shape of the tunneling conductivity can occur due to Fano interference effects. For a Fano resonance, the shift of the peak in the tunneling conductivity is determined by the total impurity level width Γ, which hardly exceeds 100 meV. In the present experiment, the peak in the tunneling conductivity spectra appears for a bias voltage of 400 meV and cannot be explained in terms of a Kondo resonance even when invoking the presence of Fano interference effects. For our experiments, we have to connect the tunneling conductivity peak to a localized state split

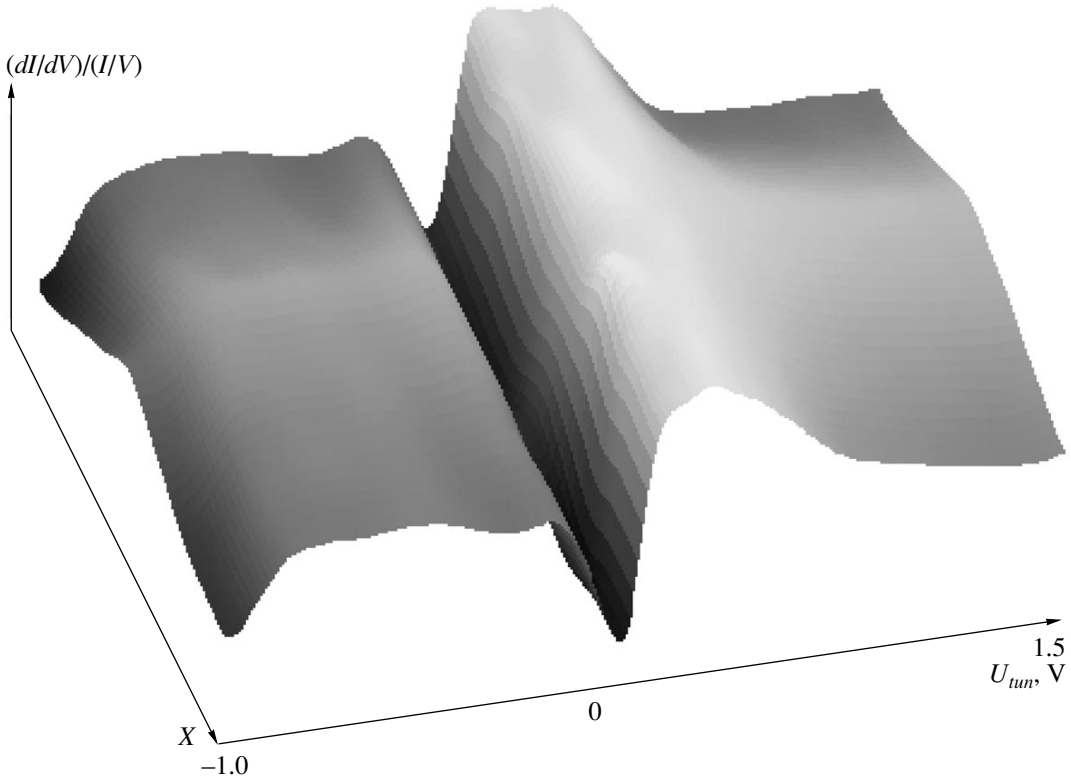


Fig. 3. Quasi-1D "cross-section" of the normalized tunneling conductivity $(dI/dV)(I/V)$ averaged along the width of the stripe indicated by the arrow in Fig. 1d.

from the conduction band due to electron scattering by the impurity potential.

For bias voltages 1 V < V < 1.5 V, the bright cross-like feature is clearly present in the spatial distribution of the local tunneling conductivity around the impurity atom. On the other hand, we note the remarkable fact that, in the spatial distribution of the tunneling current, the impurity atom appears as a dark cross, corresponding to a locally reduced tunneling current. Such unusual behavior of the tunneling characteristics can be accounted for by the many-particle nonequilibrium Coulomb interaction of a negatively charged impurity with the Fermi sea of the conduction electrons in the STM tip. When the tunneling rate Γ_t from the STM tip into the localized impurity state strongly exceeds the electron transition rate Γ for the semiconductor continuum states, a nonequilibrium negative charge appears near the impurity atom for negative tip voltages that are close to the threshold value $|eV - \varepsilon - U| < \Gamma$, where U is the onsite Coulomb repulsion between localized electrons. Near the negatively (positively) charged impurity, one expects the existence of an area with a decreased (increased) density of states deep in the continuum states (conduction and valence bands, respectively). This area with a modified density of states has a size on the order of the Debye screening radius. Consequently, the tunneling current becomes suppressed near the impurity atom when compared to a defect-free surface area.

An essential feature of our theoretical model is that, due to the many-particle interaction effects, the *tunneling current* is expected to reveal a power-law singularity with a positive exponent $a = W\nu$ [8]:

$$I(V) = \frac{\gamma_t}{2W\nu} \times \left[\frac{(eV - \varepsilon_d - U)^2 + (\gamma_t + \gamma)^2}{D^2}\right]^{W\nu} \sin(2W\nu\phi), \quad (2)$$

where $\phi = \operatorname{arccot}\left(\frac{eV - \varepsilon_d - U}{\gamma + \gamma_t}\right)$. On the other hand, this power-law singularity can be reflected as an effective enhancement of the *normalized tunneling conductivity* for the same range of bias voltages if the power-law exponent satisfies the condition $0 < W\nu < 0.5$. In this case, the singularity exceeds the decrease in the semiconductor density of states within the considered range of bias voltages:

$$(dI/dV) = \frac{\gamma_t(eV - \varepsilon_d - U)}{D^2} \times \left[\frac{D^2}{(eV - \varepsilon_d - U)^2 + (\gamma_t + \gamma)^2}\right]^{1 - W\nu} \sin(2W\nu\phi). \quad (3)$$

For $(eV - \varepsilon_d - U)\ \gamma_t + \gamma$ and rather small values of $\gamma_t + \gamma$, tunneling conductivity strongly increases. In this case, the singularity exceeds the decrease in the semiconductor density of states within the considered range of bias voltages. Finally, we want to point out that for our STS measurements we are dealing with a semiconductor density of states in the conduction and valence band that is strongly modified by the presence of the impurity atom due to the appearance of nonequilibrium impurity charges and many-particle Coulomb interaction effects between the tunneling electrons and the charged impurity. In Fig. 3, we show a 1D cross-section of the normalized conductivity along the direction indicated by the arrow in Fig. 1d. The most intriguing feature of the 1D cross-section is the remarkable enhancement of the normalized differential conductivity inside the conduction and valence bands near the impurity atoms, with almost no visible variations inside the band gap while changing the distance from the impurity. As explained above, this may be linked to the interplay between a many-particle Coulomb singularity and a modification of the nonperturbed semiconductor density of states by tunneling in the presence of a charged impurity.

In conclusion, we reported on the direct observation by low-temperature STM and STS measurements of the *d* orbitals of a Mn *p*-type impurity appearing on a cleaved InAs(110) surface. The remarkable crosslike shape of the STM image of the impurity, as well as the shape of the differential tunneling conductivity, cannot be explained in terms of a Kondo resonance. Nonequilibrium charge interaction effects provide a consistent description of our experimental results.

This work was supported by the Russian Foundation for Basic Research (project nos. 02-02-17361, 00-15-96555, and 00-02-17759) and by the Physics of Solid State Nanostructures (project no. 7) and Surface Atomic Structures (project no. 1.2) programs. Additional support was obtained from the Fund for Scientific Research–Flanders (Belgium), as well as from the Belgian Interuniversity Attraction Poles (IAP) and the Flemish Concerted Action (GOA) research programs.

REFERENCES

1. G. Binnig, H. Rohrer, and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).
2. A. Depuydt, C. Van Haesendonck, N. S. Maslova, *et al.*, Phys. Rev. B **60**, 2619 (1999).
3. P. I. Arseev, N. S. Maslova, V. I. Panov, and S. V. Savinov, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 345 (2002) [JETP Lett. **76**, 287 (2002)].
4. R. M. Feenstra, Phys. Rev. B **50**, 4561 (1994).
5. O. Yu. Kolesnychenko, R. de Kort, M. I. Katsnelson, *et al.*, Nature **415**, 507 (2002).
6. S. I. Vasil'ev, S. I. Oreshkin, S. V. Savinov, *et al.*, Instrum. Exp. Tech. **40**, 566 (1997).
7. J. A. Stroscio, R. M. Feenstra, and A. P. Fein, Surf. Sci. **181**, 295 (1987).
8. P. I. Arseev, N. S. Maslova, and S. V. Savinov, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 299 (1998) [JETP Lett. **68**, 320 (1998)].

JETP Letters, Vol. 77, No. 4, 2003, pp. 178–181. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 208–211.

Magnetoresistance of a Highly Correlated Electron Liquid¶

V. R. Shaginyan

Petersburg Nuclear Physics Institute, Russian Academy of Sciences, Gatchina, 188300 Russia
CTSPS, Clark Atlanta University, 30314 Atlanta, Georgia, USA
e-mail: vrshag@thd.pnpi.spb.ru

Received January 4, 2003

The behavior in a magnetic field of a highly correlated electron liquid approaching the fermion condensation quantum phase transition from the disordered phase is considered. We show that, at sufficiently high temperatures $T \geq T^*(x)$, the effective mass starts to depend on T, $M^* \propto T^{-1/2}$. This $T^{-1/2}$ dependence of the effective mass at elevated temperatures leads to the non-Fermi liquid behavior of the resistivity, $\rho(T) \propto T$ and at higher temperatures $\rho(T) \propto T^{3/2}$. The application of a magnetic field B restores the common T^2 behavior of the resistivity. The effective mass depends on the magnetic field, $M^*(B) \propto B^{-2/3}$, being approximately independent of the temperature at $T \leq T^*(B) \propto B^{4/3}$. At $T \geq T^*(B)$, the $T^{-1/2}$ dependence of the effective mass is reestablished. We demonstrate that this B–T phase diagram has a strong impact on the magnetoresistance (MR) of the highly correlated electron liquid. The MR as a function of the temperature exhibits a transition from negative values of MR at $T \longrightarrow 0$ to positive values at $T \propto B^{4/3}$. Thus, at $T \geq T^*(B)$, MR as a function of the temperature possesses a node at $T \propto B^{4/3}$. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 71.10.Hf; 71.27.+a

An explanation of the rich and striking behavior of a strongly correlated electron liquid in heavy fermion metals and high-temperature superconductors is, as in years before, among the main problems of condensed matter physics. There is a fundamental question about whether or not these properties can be understood within the framework of the Landau Fermi liquid theory [1]. The basis of the Landau theory is the assumption that the excitation spectrum of the Fermi liquid looks like the spectrum of an ideal Fermi gas. This excitation spectrum is described in terms of quasiparticles with an effective mass M^*, charge e, and spin 1/2. Single-particle excitations, or quasiparticles, define the major part of the low-temperature properties of Fermi liquids. The stability of the ground state of a Landau liquid is determined by the Pomeranchuk stability conditions. The stability is violated when at least one of the Landau effective interaction parameters becomes negative and reaches a critical value. The new phase at which the stability conditions are restored is again described within the framework of the same theory. The Pomeranchuk conditions do not cover all the possible instabilities. The missed instability corresponds to the situation when, at the temperature $T = 0$, the effective mass, the most important characteristic of Landau quasiparticles, can become infinitely large. Such a situation, which leads to profound consequences, can take place when the corresponding Landau amplitude, being repulsive, reaches some critical value. This leads to a completely new class of strongly correlated Fermi liquids with the fermion condensate (FC) [2, 3], which is separated from that of a normal Fermi liquid by the fermion condensation quantum phase transition (FCQPT) [4, 5].

As any phase transition, the quantum phase transition is driven by a control parameter and is related to the order parameter, which describes a broken symmetry. In our case, the control parameter is the density x of a system and the order parameter is $\kappa(\mathbf{p})$. The existence of the FC state can be revealed experimentally. Since the order parameter $\kappa(\mathbf{p})$ is suppressed by a magnetic field B, a weak magnetic field B will destroy the state with FC, converting the strongly correlated Fermi liquid into a normal Landau Fermi liquid [6]. In this case, the magnetic field plays the role of the control parameter. The transition from the strongly correlated liquid into the normal Landau liquid was observed in several experiments [7–10]. As soon as FCQPT occurs at the critical point $x = x_{FC}$, the system becomes divided into two quasiparticle subsystems: the first subsystem is characterized by quasiparticles with the effective mass M^*_{FC}, while the second one is occupied by quasiparticles with mass M^*_L. The quasiparticle dispersion law in systems with FC can be represented by two straight lines, characterized by effective masses M^*_{FC} and M^*_L and intersecting near the binding energy E_0.

The properties of these new quasiparticles with M^*_{FC} are closely related to the state of the system which is characterized by the temperature T, pressure, or the

¶ This article was submitted by the author in English.

presence of superconductivity. We may say that the quasiparticle system in the range occupied by FC becomes very "soft" and is to be considered a strongly correlated liquid. Nonetheless, the basis of the Landau Fermi liquid theory survives FCQPT: the low-energy excitations of a strongly correlated liquid with FC are quasiparticles. The only difference between the Landau Fermi liquid and Fermi liquid after FCQPT is that we have to expand the number of relevant low-energy degrees of freedom by introducing a new type of quasiparticles with the effective mass M^*_{FC} and the energy scale E_0 [4, 11].

When a Fermi system approaches FCQPT from the disordered phase, it remains a Landau Fermi liquid with the effective mass M^* strongly depending on the density $x_{FC} - x$, temperature, and magnetic field B provided that $|x_{FC} - x|/x_{FC} \ll 1$ and $T \geq T^*(x)$ [12]. This state of the system, with M^* strongly depending on T, x, and B, resembles a strongly correlated liquid. In contrast with a strongly correlated liquid, there is no energy scale E_0, and the system under consideration is a Landau Fermi liquid at sufficiently low temperatures with the effective mass $M^* \simeq$ const. Therefore, this liquid can be called a highly correlated liquid. Obviously, a highly correlated liquid has to have uncommon properties.

In this letter, we study the behavior of a highly correlated electron liquid in a magnetic field. We show that, at $T \geq T^*(x)$, the effective mass starts to depend on the temperature, $M^* \propto T^{-1/2}$. This $T^{-1/2}$ dependence of the effective mass at elevated temperatures leads to non-Fermi liquid behavior of the resistivity, $\rho(T) \sim \rho_0 + aT + bT^{3/2}$. The application of magnetic field B restores the common T^2 behavior of the resistivity, $\rho \simeq \rho_0 + AT^2$ with $A \propto (M^*)^2$. Both the effective mass and coefficient A depend on the magnetic field, $M^*(B) \propto B^{-2/3}$ and $A \propto B^{-4/3}$ being approximately independent of the temperature at $T \leq T^*(B) \propto B^{4/3}$. At $T \geq T^*(B)$, the $T^{-1/2}$ dependence of the effective mass is reestablished. We demonstrate that this B–T phase diagram has a strong impact on the magnetoresistance (MR) of the highly correlated electron liquid. The MR as a function of the temperature exhibits a transition from negative values of MR at $T \longrightarrow 0$ to positive values at $T \propto B^{4/3}$. Thus, at $T \geq T^*(B)$, MR as the function of the temperature possesses a node at $T \propto B^{4/3}$. Such a behavior is of a general form and takes place in both three-dimensional (3D) highly correlated systems and two-dimensional (2D) ones.

At $|x_{FC} - x|/x_{FC} \ll 1$ and $T \longrightarrow 0$, the effective mass M^* of a highly correlated electron liquid is given by the equation [12]

$$M^* \sim M\frac{x_{FC}}{x - x_{FC}}. \quad (1)$$

It follows from Eq. (1) that effective mass is finite provided that $|x - x_{FC}| \equiv \Delta x > 0$. Therefore, the system represents a Landau Fermi liquid. On the other hand, M^* diverges as the density x tends to the critical point of FCQPT. As a result, the effective mass strongly depends on such quantities as the temperature, pressure, and magnetic field, provided that they exceed their critical values. For example, when T exceeds some temperature $T^*(x)$, Eq. (1) is no longer valid, and M^* depends on the temperature as well. To evaluate this dependence, we calculate the deviation $\Delta x(T)$ generated by T. The temperature smoothing of the Fermi function $\theta(p_F - p)$ at p_F induces the variation $p_F\Delta p/M^*(x) \sim T$, and $\Delta x(T)/x_{FC} \sim M^*(x)T/p_F^2$, where p_F is the Fermi momentum and M is the bare electron mass. The deviation Δx can be expressed in terms of $M^*(x)$ using Eq. (1), $\Delta x/x_{FC} \sim M/M^*(x)$. Comparing these deviations, we find that at $T \geq T^*(x)$ the effective mass depends noticeably on the temperature, and the equation for $T^*(x)$ becomes

$$T^*(x) \sim p_F^2\frac{M}{(M^*(x))^2} \sim \varepsilon_F(x)\left(\frac{M}{M^*(x)}\right)^2. \quad (2)$$

Here, $\varepsilon_F(x)$ is the Fermi energy of noninteracting electrons with mass M. It follows from Eq. (2) that M^* is always finite at temperatures $T > 0$. At $T \geq T^*(x)$, the main contribution to Δx comes from the temperature; therefore,

$$M^* \sim M\frac{x_{FC}}{\Delta x(T)} \sim M\frac{\varepsilon_F}{M^*T}. \quad (3)$$

As a result, we obtain

$$M^*(T) \sim M\left(\frac{\varepsilon_F}{T}\right)^{1/2}. \quad (4)$$

Equation (4) allows us to evaluate the resistivity as a function of T. There are two terms contributing to the resistivity. Taking into account that $A \sim (M^*)^2$ and Eq. (4), we obtain the first term $\rho_1(T) \sim T$. The second term $\rho_2(T)$ is related to the quasiparticle width γ. When $M/M^* \ll 1$, the width $\gamma \propto (M^*)^3T^2/\varepsilon(M^*) \propto T^{3/2}$, where $\varepsilon(M^*) \propto (M^*)^2$ is the dielectric constant [5, 13]. Combining both of the contributions, we find that the resistivity is given by

$$\rho(T) - \rho_0 \sim aT + bT^{3/2}. \quad (5)$$

Thus, it turns out that, at low temperatures $T < T^*(x)$, the resistivity $\rho(T) - \rho_0 \sim AT^2$. At higher temperatures, the effective mass depends on the temperature, and the main contribution comes from the first term on the right hand side of Eq. (5). At the same time, $\rho(T) - \rho_0$ follows the $T^{3/2}$ dependence at elevated temperatures.

In the same way as Eq. (4) was derived, we can obtain the equation determining $M^*(B)$ [12]:

$$M^*(B) \sim M\left(\frac{\varepsilon_F}{B\mu}\right)^{2/3}, \quad (6)$$

where μ is the electron magnetic moment. We note that M^* is determined by Eq. (6) as long as $M^*(B) \le M^*(x)$; otherwise, we have to use Eq. (1). It follows from Eq. (6) that the application of a magnetic field reduces the effective mass. Note that if there exists an itinerant magnetic order in the system which is suppressed by magnetic field $B = B_{c0}$, Eq. (6) has to be replaced by the equation [6],

$$M^*(B) \propto \left(\frac{1}{B - B_{c0}}\right)^{2/3}. \quad (7)$$

The coefficient of T^2 in the expression for the resistivity $A(B) \propto (M^*(B))^2$ diverges as

$$A(B) \propto \left(\frac{1}{B - B_{c0}}\right)^{4/3}. \quad (8)$$

At elevated temperature, there is a temperature $T^*(B)$ at which $M^*(B) \simeq M^*(T)$. Comparing Eq. (4) and Eq. (7), we see that $T^*(B)$ is given by

$$T^*(B) \propto (B - B_{c0})^{4/3}. \quad (9)$$

At $T \ge T^*(x)$, Eq. (9) determines the line in the B–T phase diagram that separates the region of the B-dependent effective mass from the region of the T-dependent effective mass. At the temperature $T^*(B)$, there occurs a crossover from the T^2 dependence of the resistivity to the T dependence: at $T < T^*(B)$, the effective mass is given by Eq. (7), and at $T > T^*(B)$ M^* is given by Eq. (4).

Using the B–T phase diagram just described above, we consider the behavior of MR

$$\rho_{mr}(B, T) = \frac{\rho(B, T) - \rho(0, T)}{\rho(0, T)} \quad (10)$$

as a function of magnetic field B and T. Here, $\rho(B, T)$ is the resistivity measured at the magnetic field B and temperature T. We assume that the contribution $\Delta\rho_{mr}(B)$ coming from the magnetic field B can be treated within the low-field approximation and given by the well-known Kohler rule,

$$\Delta\rho_{mr}(B) \sim B^2\rho(0, \Theta_D)/\rho(0, T), \quad (11)$$

where D is the Debye temperature. Note that the low field approximation implies that $\Delta\rho_{mr}(B) \ll \rho(0, T) \equiv \rho(T)$. Substituting Eq. (11) into Eq. (10), we find that

$$\rho_{mr}(B, T) \sim \frac{c(M^*(B, T))^2T^2 + \Delta\rho_{mr}(B) - c(M^*(0, T))^2T^2}{\rho(0, T)}. \quad (12)$$

Here, $M^*(B, T)$ denotes the effective mass M^*, which now depends on both the magnetic field and the temperature, and c is a constant.

Consider MR given by Eq. (12) as a function of B at some temperature $T = T_0$. At low temperatures $T_0 \le T^*(x)$, the system behaves as a common Landau Fermi liquid, and MR is an increasing function of B. When the temperature T_0 is sufficiently high, $T^*(B) < T_0$, and the magnetic field is small, $M^*(B, T)$ is given by Eq. (4). Therefore, the difference $\Delta M^* = |M^*(B, T) - M^*(0, T)|$ is small and the main contribution is given by $\Delta\rho_{mr}(B)$. As a result, MR is an increasing function of B. At elevated B, the difference ΔM^* becomes a decreasing function of B, and MR as a function of B reaches its maximum value at $T^*(B) \sim T_0$. In accordance with Eq. (9), $T^*(B)$ determines the crossover from T^2 dependence of the resistivity to T dependence. Differentiating the function $\rho_{mr}(B, T)$ given by Eq. (12) with respect to B, one can verify that the derivative is negative at sufficiently large values of the magnetic field when $T^*(B) \simeq T_0$. Thus, we are led to the conclusion that the crossover manifests itself as the maximum of MR as the function of B.

We now consider MR as a function of T at some B_0. At low temperatures $T \ll T^*(B)$, it follows from Eqs. (4) and (7) that $M^*(B)/M^*(T) \ll 1$, and MR is determined by the resistivity $\rho(0, T)$. Note that B has to be comparatively high to ensure the inequality $T^*(x) \le T \ll T^*(B)$. As a result, MR tends to –1, $\rho_{mr}(B_0, T \longrightarrow 0) \simeq -1$. Differentiating the function $\rho_{mr}(B_0, T)$ with respect to B_0, we can check that its slope becomes steeper as B_0 is decreased, being proportional to $\propto(B_0 - B_{c0})^{-7/3}$. At $T = T_1 \sim T^*(B_0)$, MR possesses a node, because at this point the effective mass $M^*(B) \simeq M^*(T)$, and $\rho(B, T) \simeq \rho(0, T)$. Again, we can conclude that the crossover from T^2 resistivity to T resistivity, which occurs at $T \sim T^*(B_0)$, manifests itself in the transition from negative MR to positive MR. At $T > T^*(B)$, the main contribution in MR comes from $\Delta\rho_{mr}(B)$, and MR reaches its maximum value. Upon using Eq. (11) and taking into account that at this point T has to be determined by Eq. (9), $T \propto (B_0 - B_{c0})^{4/3}$, we find that the maximum value $\rho^m_{mr}(B_0)$ of MR is $\rho^m_{mr}(B_0) \propto (B - B_{c0})^{-2/3}$. Thus, the maximum value is a decreasing function of B_0. At $T^*(B) \ll T$, MR is a decreasing function of the temperature, and at elevated temperatures MR eventually vanishes, since $\Delta\rho_{mr}(B)/\rho(T) \ll 1$.

The recent paper [14] reports on measurements of the resistivity of $CeCoIn_5$ in a magnetic field. With increasing field, the resistivity evolves from T temperature dependence to T^2 dependence, while the field dependence of $A(B) \sim (M^*(B))^2$ displays a critical behavior best fitted by the function $A(B) \propto (B - B_{c0})^{-\alpha}$, where $\alpha \simeq 1.37$ [14]. All the data are in good agreement with the B–T phase diagram given by Eq. (9). The critical behavior displaying $\alpha = 4/3$ [12] and described by Eq. (8) is also in good agreement with the data. Transition from negative MR to positive MR with increasing T was also observed [14]. We believe that an additional analysis of the data [14] will reveal that the crossover from T^2 dependence of the resistivity to the T depen-

dence occurs at $T \propto (B - B_{c0})^{4/3}$. This analysis could also reveal additional peculiarities of MR.

In conclusion, we have described the behavior of a highly correlated electron liquid in a magnetic field. The highly correlated liquid exhibits a strong dependence of the effective mass M^* on the temperature and the magnetic field. This strong dependence is of crucial importance when describing the B–T phase diagram and such properties as MR and the critical behavior. We have also identified the behavior of the heavy fermion metal $CeCoIn_5$ in magnetic fields displayed in [14] as the highly correlated behavior of a Landau Fermi liquid approaching FCQPT from the disordered phase.

I am grateful to CTSPS for the hospitality during my stay in Atlanta. I also thank G. Japaridze for fruitful discussions. This work was supported in part by the Russian Foundation for Basic Research (grant no. 01-02-17189).

REFERENCES

1. L. D. Landau, Zh. Éksp. Teor. Fiz. **30**, 1058 (1956) [Sov. Phys. JETP **3**, 920 (1956)].
2. V. A. Khodel and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **51**, 488 (1990) [JETP Lett. **51**, 553 (1990)]; V. A. Khodel, V. R. Shaginyan, and V. V. Khodel, Phys. Rep. **249**, 1 (1994).
3. G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 208 (1991) [JETP Lett. **53**, 222 (1991)].
4. M. Ya. Amusia and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 268 (1990) [JETP Lett. **73**, 232 (2001)]; Phys. Rev. B **63**, 224507 (2001); V. R. Shaginyan, Physica B & C (Amsterdam) **312–313**, 413 (2002).
5. S. A. Artamonov and V. R. Shaginyan, Zh. Éksp. Teor. Fiz. **119**, 331 (2001) [JETP **92**, 287 (2001)].
6. Yu. G. Pogorelov and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 614 (2002) [JETP Lett. **76**, 532 (2002)].
7. A. P. Mackenzie, S. R. Julian, D. C. Sinclair, and C. T. Lin, Phys. Rev. B **53**, 5848 (1996).
8. P. Gegenwart, J. Custers, C. Geibel, *et al.*, Phys. Rev. Lett. **89**, 056 402 (2002).
9. C. Proust, E. Boaknin, R. W. Hill, *et al.*, Phys. Rev. Lett. **89**, 147 003 (2002).
10. G.-Q. Zheng, K. Yamada, M. Fujita, *et al.*, cond-mat/0211400.
11. M. Ya. Amusia, S. A. Artamonov, and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 435 (2001) [JETP Lett. **74**, 396 (2001)].
12. V. R. Shaginyan, JETP Lett. (in press); cond-mat/0212624.
13. V. A. Khodel and P. Schuck, Z. Phys. B **104**, 505 (1997).
14. J. Paglione, M. A. Tanatar, D. G. Hawthorn, *et al.*, cond-mat/0212502.

JETP Letters, Vol. 77, No. 4, 2003, pp. 182–186. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 212–216.

Influence of an Interface Double Electric Layer on the Superconducting Proximity Effect in Ferromagnetic Metals¶

V. M. Edelstein

Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow region, 142432 Russia
e-mail: edelsh@issp.ac.ru

Received January 10, 2003

The electric field within the double electric layer that occurs when two conductors with different work functions come into contact induces interface spin–orbit coupling. In the case of contact of a conventional, *s*-wave superconductor with a ferromagnetic metal, the coupling is predicted to give rise to the appearance of triplet superconductivity on both sides of the interface. The form of the triplet component of the condensation amplitude is determined, and the ability of the triplet superconductivity to penetrate into the ferromagnet on the usual coherence length is shown. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 74.45.+c; 71.70.Ej; 73.40.Jn; 74.20.-z

The phenomenon of penetration of superconductivity into a normal metal placed in contact with a superconductor (S) is now reasonably well understood [1]. However, the mechanisms through which superconductivity can penetrate into a ferromagnetic metal are still not clear. In accordance with the theory developed up to now [2], the penetration of a singlet Cooper pair into a ferromagnet (F) looks as follows. The Fermi surfaces in the F that correspond to conduction electrons with spins directed along (up states) and against (down states) the exchange field do not coincide. Therefore, upon entering the F region, the two electrons that compose the Cooper pair have to occupy different Fermi surfaces. The difference between the Fermi momenta of the surfaces makes it impossible for the total momentum of the pair be zero. As a result, the condensation amplitude, which in the case of a normal metal would exponentially decay on the length scale of the coherence length, $\xi_0 = v_f(2\pi T_c)^{-1}$ (here, v_f is the average Fermi velocity in a normal metal), acquires an additional modulation factor that oscillates with a short period proportional to $\lambda_F = 2\pi(k_{f,d} - k_{f,u})^{-1} \ll \xi_0$ (here, $k_{f,d}$ and $k_{f,u}$ are the Fermi wavevectors of down and up states of the F, respectively). So λ_F becomes the effective penetration depth of superconductivity into the ferromagnetic metal. Some experimental data published recently [3] seem to confirm the theory; there are also experimental results [4] that agree less with the theory and indicate that the superconductor may influence the ferromagnet on a distance much longer than λ_F. On the grounds of all that has been said, one can expect the condensation amplitude to cure the oscillations and hence to penetrate deep into the F region if the two electrons of the Cooper pair gain the possibility at equal energy to have opposite momenta and thus reduce the total momentum to zero. For that, both the electrons should occupy the same Fermi surface, i.e., should be in the same spin state. The very fact that it is easier for triplet than for singlet superconductivity to coexist with ferromagnetism is a simple consequence of the weak sensitivity of some of the triplet states to an external magnetic field, well known from studies of superfluid ^{3}He. The problem is to reveal mechanisms that are able to transform initially singlet superconductivity to triplet near the contact. The purpose of this paper is to point out a natural reason for such a transformation. The reason is the interface spin–orbit coupling that is always brought about by contact.

In fact, when two metals come into contact, electrons are known to transfer from the metal with the lower work function to the metal with the higher work function, so as to bring the two Fermi levels into coincidence. As a result, a double electric layer of a thickness on the order of the screening length r_D appears. When a conduction electron is in the layer, it is subject to the electric field that induces the spin–orbit (SO) coupling. Because r_D in metals is about the lattice constant, which is much shorter than the expected decay length of the triplet superconductivity, this interface SO coupling can be put into the form $H_{SO} = \alpha(\mathbf{p} \times \mathbf{c}) \cdot \boldsymbol{\sigma}\delta(\mathbf{c} \cdot \mathbf{r})$, where $\mathbf{p}$, $\boldsymbol{\sigma}$, and $\mathbf{c}$ are, respectively, the electron momentum, the Pauli matrices, and the normal directed from the F to S region. Plank's constant $\hbar$ is set to unity throughout. The interface SO coupling arises from the electric field in the double layer in the same way as the well-known $\mathbf{p}$-linear Hamiltonian in the energy spec-

¶This article was submitted by the author in English.

trum of bulk polar semiconductors [5] arises from the intracrystalline electric field [6, 7]. The presence of H_{SO} violates the total spin conservation and should result in singlet–triplet mixing. Recently, some consequences of mixing for the proximity effect in normal metals were pointed out [8]. Here, we consider how the mixing facilitates the penetration of superconductivity into the ferromagnetic metal under the assumption that the exchange field is uniform and directed along the interface.

We will assume for simplicity that the electron spectrum on both sides of the contact is isotropic. Then the one-particle Hamiltonian of the system has the form

$$H_0(\mathbf{p}) = \mathbf{p}\frac{1}{2m(z)}\mathbf{p} + \theta(-z)\left[\Gamma + \frac{b}{2}(\mathbf{h}\cdot\sigma + 1)\right] + [\beta + \alpha(\mathbf{p}\times\mathbf{c})\cdot\sigma]\delta(z), \quad (1)$$

where z is the coordinate along c, the electron mass $m(z)$ equals m_F in the F and m_S in the S (for definiteness, we assume $m_S \geq m_F$), $\theta(z)$is the unit step function, Γ is the difference between the lower conduction band edge of the F and of the S, $\mathbf{h}$ is the unit vector along the exchange field, b is the exchange energy, and $\beta\delta(z)$ models the spin-independent interface potential. The straightforward way to look into the behavior of superconductivity is to investigate the self-consistent equation for the gap-matrix $\Delta_{\alpha\beta}(\mathbf{r}, \mathbf{r}')$, which near T_c has the form [9]

$$\Delta_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = -T_c\sum_{\epsilon}\int dr_3 dr_4 \times V^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4)F_{\delta\gamma}(\mathbf{r}_4, \mathbf{r}_3|i\epsilon), \quad (2)$$

$$F_{\kappa\rho}(\mathbf{r}_1, \mathbf{r}_2|i\epsilon) = \int dr_3 dr_4 G_{\kappa\gamma}(\mathbf{r}_1, \mathbf{r}_3|i\epsilon) \times \Delta_{\gamma\delta}(\mathbf{r}_3, \mathbf{r}_4)(-1)G^t_{\delta\rho}(\mathbf{r}_4, \mathbf{r}_2|i\epsilon), \quad (3)$$

where $V^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4)$ is the interparticle interaction, the superscript t denotes transposition of the spinor indices (and also space coordinates when refering to Green's functions), G is the Green's function of independent particles, and F is the anomalous Green's function. But, before being able to enter the subject, one must ascertain the form of the G function. This can be done in the following way. (i) Since the interface is assumed to be translationally invariant in the xy plane, the electron momentum $\mathbf{k}_\parallel$ along the plane is conserved and the one-particle scattering problem is effectively one dimensional. The eigenfunctions of the 1D Schrödinger equation can be easily found and are not presented here. (ii) It is convenient to expand the G function on the spinor basis that includes the projection operators on the up and down states $G_{\alpha\beta} = \sum_{i=1}^{4} M^{(i)}_{\alpha\beta}G_{(i)}$, where $M^i = \{\Pi^{(u)}, \Pi^{(d)}\ A^{(\uparrow\downarrow)}.A^{(\downarrow\uparrow)}\}$, $\Pi^{(u,d)} = \frac{1}{2}(1 \pm \mathbf{h}\cdot\boldsymbol{\sigma})$, $A^{(\uparrow\downarrow)} = \frac{1}{2}h_+(-\mathbf{c}\cdot\boldsymbol{\sigma} - i\mathbf{h}\times\mathbf{c}\cdot\boldsymbol{\sigma})$, $A^{(\downarrow\uparrow)} = A^{(\uparrow\downarrow)+}$, $h_\pm = h_x \pm ih_y$. (iii) The chemical potential μ is supposed to be greater than $\Gamma + b$ so that both the up and down states in the F are partially filled. The spectral decomposition of the G function with definite longitudinal momentum $\mathbf{k}_\parallel$ can be constructed in line with the general theory of 1D scattering [10] with the help of the full set of Jost functions—the orthogonal and normalized scattering states of a particle corresponding to waves ingoing from the left and from the right. One should take into account that a particle with energy in the interval $(0, \Gamma)$ may be incident only from the right and that one with energy in the interval $(\Gamma, \Gamma + b)$ may be only incident from the left, being in the down state. (iv) The integration on the spectral parameters (momenta $\mathbf{k}_\parallel$ and $k_\perp$) can be performed using the stationary phase approximation, which holds on a scale that is large compared to the Fermi wavevector of any of the occupied bands. (v) Finally, one should sum over all spin states. In the sector $z < 0$, $z' > 0$, which is first of all needed for the following analysis, the result is

$$G_{(u)}(\mathbf{r}, \mathbf{r}'|i\epsilon) \cong (i\pi D_u(\mathbf{p})N_u^{1/2}(\mathbf{p}))^{-1}\exp\{i\Phi_u(\mathbf{p})\}\Big|_{\mathbf{p}=\mathbf{p}_u}, \quad (4)$$

$$G_{(\uparrow\downarrow)}(\mathbf{r}, \mathbf{r}'|i\epsilon) \cong 2V_{ud}(\mathbf{p}) \times (i\pi D_u(\mathbf{p})D_d(\mathbf{p})N_u^{1/2}(\mathbf{p}))^{-1}\exp\{i\Phi_u(\mathbf{p})\}\Big|_{\mathbf{p}=\mathbf{p}_u}, \quad (5)$$

where $\mathbf{r} = (\boldsymbol{\rho}, z)$, $\mathbf{r}' = (\boldsymbol{\rho}', z')$,

$$D_u(\mathbf{p}) = im_S^{-1}(k_{f.s}^2 - p^2)^{1/2} + im_F^{-1}(k_{f.u}^2 - p^2)^{1/2} - 2(\beta + \alpha\mathbf{p}\times\mathbf{c}\cdot\mathbf{h}), \quad (6)$$

$$N_u(\mathbf{p}) = |\boldsymbol{\rho} - \boldsymbol{\rho}'|p^{-1} \times [z'k_{f.s}^2(k_{f.s}^2(k_{f.s}^2 - p^2)^{-3/2} - zk_{f.u}^2(k_{f.u}^2 - p^2)^{-3/2})], \quad (7)$$

$$\Phi_u(\mathbf{p}) = z'(k_{f.s}^2 - p^2)^{-1/2}(k_{f.s}^2 + i\epsilon m_S) - z(k_{f.u}^2 - p^2)^{-1/2}(k_{f.u}^2 + i\epsilon m_F), \quad (8)$$

$p = |\mathbf{p}|$, $k_{f.s}$ is the Fermi momentum of the S ($k_{f.u} \leq k_{f.d} \leq k_{f.s}$), $V_{ud} = \frac{\alpha}{2i}(p_+ + p_-h_+^2)$, $p_\pm = p_x + ip_y$, $\mathbf{p}_u = [(\boldsymbol{\rho} - \boldsymbol{\rho}')/|\boldsymbol{\rho} - \boldsymbol{\rho}'|]p_u$ is the 2D vector whose value p_u is defined as the root of the equation

$$|\boldsymbol{\rho} - \boldsymbol{\rho}'| = p[z'(k_{f.s}^2 - p^2)^{-1/2} - z(k_{f.u}^2 - p^2)^{-1/2}], \quad (9)$$

and $D_d(\mathbf{p})$ is obtained from $D_u(\mathbf{p})$ if one changes in the r.h.s. of Eq. (6) $k_{f.u} \longrightarrow k_{f.d}$ and $\alpha \longrightarrow -\alpha$. An inspection of Eq. (9) reveals that $p_u \leq k_{f.u}$ at any $\mathbf{r}$ and $\mathbf{r}'$. The implicit dependence of the vector $\mathbf{p}_u$ (and hence $G_{(u)}$ and $G_{(\uparrow\downarrow)}$) on $\mathbf{r}$ and $\mathbf{r}'$ is a result of the Fermi wavevector mismatch, $k_{f.s} \neq k_{f.u}$. To obtain $G_{(u)}$at $\epsilon < 0$, one should

change in Eq. (4) $\epsilon \longrightarrow -\epsilon$, $\mathbf{p}_u \longrightarrow -\mathbf{p}_u$ and take the complex conjugation. The function $G_{(\uparrow\downarrow)}$ at $\epsilon < 0$ is obtained in the same way except for the complex conjugating of V_{ud}. The second pair of G functions, $G_{(d)}$ and $G_{(\downarrow\uparrow)}$, has an analogous form. They are given by the r.h.s. of Eqs. (4) and (5), respectively, where all quantities denoted by the subscript u ($k_{f.u}$, D_u, N_u, p_u, and Φ_u) are replaced by the quantities denoted by a subscript d with the prescription $V_{ud} \longrightarrow V_{du} = (V_{ud})^*$. There is an asymmetry between $G_{(\uparrow\downarrow)}$ and $G_{(\downarrow\uparrow)}$. The former contains the function $D_u(\mathbf{p}_d)$, whereas the latter contains $D_d(\mathbf{p}_u)$. The defining equation for p_d (which is nothing but Eq. (9) where $k_{f.u}$ is replaced by $k_{f.d}$) says only that $p_d \leq k_{f.d}$. But it may appear that $p_d > k_{f.u}$ at some $\mathbf{r}$ and $\mathbf{r}'$. One can show that in such a case the equation $(k_{f.u}^2 - p_d^2)^{1/2}$, which enters $D_u(\mathbf{p}_d)$, should be replaced by $i(p_d^2 - k_{f.u}^2)^{1/2}$. In contrast, the quantity $k_{f.d}^2 - p_u^2$, which enters $D_d(\mathbf{p}_u)$, is always positive. This asymmetry reflects the fact that at $k_{f.u} \leq k_{f.d}$, not every electron incident from the S being in the down state can pass to the up states of the F even if it is allowed to pass into the down state of the F.

The interparticle interaction (in both metals) includes all spherical harmonics; however, the conventional character of the S assumed means that the strongest attraction takes place for electron pairs in the singlet s-wave state. Inspite of the changes in the G function induced by H_{SO}, upon solving Eqs. (2) and (3) with

$$V^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4) = V_s^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4)$$
$$= \lambda_s(\mathbf{r}_1) g_{\alpha\beta} g^t_{\gamma\delta} \delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(\mathbf{r}_3 - \mathbf{r}_4)\delta(\mathbf{r}_1 - \mathbf{r}_3),$$

where $g = i\sigma_y$, $\lambda_s(\mathbf{r}) = \theta(-z)\lambda_s(F) + \theta(z)\lambda_s(S)$, $\lambda_s(F)$ and $\lambda_s(S)$ are the coupling constants in the F and S, respectively, one obtains the singlet gap-matrix $\Delta^{(s)}_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')g_{\alpha\beta}\Delta_{(s)}(z)$ with the usual behavior of $\Delta_{(s)}(z)$ described at the beginning of the paper up to small corrections on the order of $(\alpha m)^2 \ll 1$—it is a smooth function at $z > 0$ that approaches a constant value $\Delta_{(s)}(\infty)$ in the bulk of the S, and it is a rapidly oscillating function at $z < 0$. At the same time, $F_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2)$ given by the r.h.s. of Eq. (3) with $\Delta_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = \Delta^{(s)}_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2)$ gains a different quality—it acquires a triplet component, i.e., a component that is even at $\alpha \rightleftarrows \beta$ and odd at $\mathbf{r}_1 \rightleftarrows \mathbf{r}_2$, and, what is more, the component does not suffer from the rapid oscillations. To verify this fact, one needs an appropriate projector, i.e., an operator such that a nonzero result of its application to the F function would definitely indicate the presence of the triplet part. As such a projector, one can choose the operator of convolution of the F function with the interparticle interaction in the triplet p-wave channel

$$-T\sum_{\epsilon}\int dr_3 dr_4 V_p^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4)F_{\delta\gamma}(\mathbf{r}_4, \mathbf{r}_3|i\epsilon), \quad (10)$$

Here,

$$V_p^{\alpha\beta|\gamma\delta}(\mathbf{r}_1, \mathbf{r}_2|\mathbf{r}_3, \mathbf{r}_4) = \lambda_p(\mathbf{r}_1)(\sigma^k g)_{\alpha\beta}$$
$$\times (g^t\sigma^k)_{\gamma\delta}\delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(\mathbf{r}_3 - \mathbf{r}_4)\delta(\mathbf{r}_1 - \mathbf{r}_3)\nabla^i_{12}\nabla^i_{34}$$

is the e–e interaction in the triplet p-wave channel [11] and $\nabla^i_{12} = (\partial/\partial r_2^i - \partial/\partial r_1^i)/2ik_f$, where k_f is a representative wavevector on the order of $k_{f.s}$. Its precise value is not important, because it always enters equations being multiplied by λ_p. If one considers the convolution only from a mathematical viewpoint, λ_p can take any nonzero value. However, because the actual e–e interaction in the metals surely has a nonzero triplet p-wave component, one can choose λ_p equal to its true value, $\lambda_p(\mathbf{r}) = \lambda_p(S)\theta(z) + \lambda_p(F)\theta(-z)$. Then the convolution will define the triplet component of the gap-matrix, $\Delta^{(t)}_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2)$, just as the convolution with the singlet interaction defines the singlet component of the matrix.

Being represented as a function of the Cooper pair center of mass $\mathbf{r}$ and the momentum of the relative motion $\mathbf{p}$, the convolution takes the form $\Delta^{(t)}_{\alpha\beta}(\mathbf{r}, \mathbf{p}) = A_{ij}(\mathbf{r})\dfrac{p_f}{k_f}(\sigma_i g)_{\alpha\beta}$ (it will be seen below that only conserved components of $\mathbf{p}$ parallel to the interface enter this equation), where

$$A_{ki}(\mathbf{r}) = \lambda_p(\mathbf{r})\int d^3\mathbf{r}' K^{ik}_{(st)}(\mathbf{r}, \mathbf{r}')\Delta_{(s)}(\mathbf{r}'), \quad (11)$$

$$K^{ik}_{(st)}(\mathbf{r}_1, \mathbf{r}_2) = \lim_{\mathbf{r}'_1 \to \mathbf{r}_1} \nabla^i_{11'} T$$
$$\times \sum_{\epsilon_n} \mathrm{Tr}[g^t\sigma^k G_{\epsilon_n}(\mathbf{r}_1, \mathbf{r}_2) g G^t_{-\epsilon_n}(\mathbf{r}_2, \mathbf{r}'_1)]. \quad (12)$$

In evaluating the kernel $K^{ik}_{(st)}$ by means of the obtained G function, one should remember that rapidly oscillating terms should be removed; i.e., one should consider the kernel averaged over distances large compared to k_f but small compared to $\xi_0 = v_f/(2\pi T_c)$ (here and below, v_f is the characteristic Fermi velocity of the system) and also that, since we did not introduce a frequency cutoff on the order of the Debye frequency ω_D in deriving Eq. (12), the equation is valid at $|z_1 - z_2| > \delta \sim v_f/\omega_D$ [12]. By expanding the Green's functions on the spinor basis M^i, one can see that major contributions come from the pairs of Green's functions $G_{(u)}(\mathbf{r}_1, \mathbf{r}_2|i\epsilon)$, $G_{(\uparrow\downarrow)}(\mathbf{r}_2, \mathbf{r}'_1| - i\epsilon)$ and $G_{(d)}(\mathbf{r}_1, \mathbf{r}_2|i\epsilon)$ $G_{(\downarrow\uparrow)}(\mathbf{r}_2, \mathbf{r}'| - i\epsilon)$. In these contributions, the largest parts of the phases Φ_u

and Φ_d (see Eq. (8)) that do not depend on the imaginary frequencies $i\epsilon_n = i\pi T(2n + 1)$ exactly cancel each other, thereby avoiding the oscillating behavior of $A_{ji}(\mathbf{r})$. The first pair of the functions forms a part of $K^{ik}_{(st)}$ that gives rise to the triplet component with $\mathbf{S} \cdot \mathbf{h} = 1$ (where $\mathbf{S}$ is the total spin of the Cooper pair) and the second one results in formation of the component with $\mathbf{S} \cdot \mathbf{h} = -1$. The triplet component with $\mathbf{S} \cdot \mathbf{h} = 0$ can be shown to rapidly oscillate like $\Delta_{(s)}(z)$. So one gets $A_{ji}(\mathbf{r}) = h_i(h \times c)_j\Delta_{(t1)}(z) + ih_ic_j\Delta_{(t2)}(z)$. The explicit form of the kernel $K^{ik}_{(st)}$ is rather awkward at arbitrary parameters of the energy spectrum of the F. Therefore, we consider below the limiting case of equal masses $m_S = m_F \equiv m$ and weak exchange energy $\xi_0^{-1} \ll k_{f.d} - k_{f.u} \ll k_{f.d} + k_{f.u}$. The difference between $k_{f.s}$ and $1/2(k_{f.d} + k_{f.u})$ will be ignored as well. Then, with an accuracy up to corrections on the order of $(b/\mu)^2$, at $z < 0$, one gets

$$\Delta_{(t1,2)}(z) = -\alpha m\lambda_p(N)N(0)\xi_0^{-1} \times \int_0^\infty dz' \int_0^\infty du L_{(1,2)}(z, z', u)\Delta_{(s)}(z'), \tag{13}$$

$$L_{(1)}(z, z', u) = \frac{B}{2}u(1 + u)^{-1}[1 + B^2(1 + u)]^{-2} \times \sinh^{-1}\left(\frac{\zeta}{\xi_0}\sqrt{1 + u}\right), \tag{14}$$

$$L_{(2)}(z, z', u) = \frac{b}{2\mu}\frac{|z|}{|z| + z'} \times \left[1 + \frac{|z| + z'}{4\xi_0}\sqrt{1 + u}\coth\left(\frac{|z| + z'}{\xi_0}\sqrt{1 + u}\right)\right]L_{(1)}(z, z', u). \tag{15}$$

Here, $N(0) = mk_f(2\pi^2)^{-1}$ is the electron density of states per unit energy interval at the Fermi level and $B = m\beta/k_F$. It is seen that the functions $\Delta_{(t1,2)}(z)$ and hence the triplet part of the F function concentrate in a domain of width ξ_0. The contribution to the r.h.s. of Eq. (13) from integration over negative z' is small because of the oscillating behavior of $\Delta_{(s)}(z)$ at $z < 0$. One can show that A_{ji} on the S side of the contact also concentrates in a domain of the same order and at any energy spectrum of the F has only $h_i(h \times c)_j$ and $(c \times h)_ih_j$ components but not a ih_ic_j component. In the weak exchange energy limit, the components combine to $e_{jin}c_n$, yielding the same triplet order parameter as in the case of N/S contact [8] plus corrections on the order of $(b/\mu)^2$.

In principle, one ought to introduce both singlet and triplet components of the gap-matrix and the *s*- and *p*-wave interparticle interactions from the very beginning. Then the spinor self-consistency equation reduces to a symmetrical set of two simultaneous equations for $\Delta_{(s)}(\mathbf{r})$ and $A_{ji}(\mathbf{r})$. However, the conventional character of superconductivity assumed in the bulk of the S means that the *s*-wave interaction is attractive and larger in magnitude. Therefore, in the bulk, where the influence of H_{SO} disappears, only the *s*-wave interaction is relevant in the sense that the convolution of the F function with the *p*-wave interaction vanishes. Thus, there is only $\Delta_{(s)}(\mathbf{r})$ in the bulk. The *p*-wave interaction becomes relevant near the interface. If λ_p is small, the set of equations transforms into the usual equation for $\Delta_{(s)}$ and Eq. (11). Thus, although both components of the gap-matrix, $\Delta^{(s)}_{\alpha\beta}$ and $\Delta^{(t)}_{\alpha\beta}$, are simultaneously initiated at the critical temperature T_c, they do not possess equal rights—the singlet order parameter is the source for the triplet one. The important difference between the triplet part of the F function, which is the basic conceptual element of the pairing theory, and that of the gap-matrix, which is only one of characteristics of the F function, should also be stressed. The latter is proportional to λ_p, whereas the former at small λ_p is independent of λ_p.

Since $B \sim U_b/\mu$, where U_b is the height of the interface barrier, one may estimate $B \sim 1$. The small quantity in Eq. (13) is $\alpha m \sim \alpha_V/v_F$, where α_V is the "bulk" SO constant within the double electron layer. A tentative estimate of the order of magnitude of α_V (and hence αm) performed previously [8] shows that αm can be on the order of 10^{-2} under favorable conditions. Thus, the F function in the F has a rapidly oscillating singlet component and a small but smooth triplet component [13]. (In the general case, there is no reason for λ_p/λ_s to be anomalously small. So $\Delta_{(t1,2)}/\Delta_{(s)} \sim (\alpha m)(\lambda_p N(0))$ can well appear to be on the order of 10^{-2}–10^{-3}. However, discussion of possible effects of $\Delta_{(t1,2)}$, in particular, its nonunitarity in the F, is beyond the scope of this paper.)

Finally, we remark that the above shows that the problem of the F/S contact admits rigorous consideration within the Gor'kov–Nambu formulation of pairing theory. The triplet–singlet mixing found is the consequence of inversion symmetry breaking due to the *p*-linear interface SO coupling. Still more interesting effects of simultaneous breaking of T and P symmetry can apparently be brought out in the nonequilibrium properties of the contact. So this system certainly merits further examination.

This work was supported in part by the Russian Foundation for Basic Research (grant no. 01-02-16506).

REFERENCES

1. P.-G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966; Mir, Moscow, 1968).
2. E. A. Demler, G. B. Arnold, and M. R. Beasley, Phys. Rev. B **55**, 15174 (1997).
3. T. Kontos, M. Aprili, J. Lesueur, and X. Grison, Phys. Rev. Lett. **86**, 304 (2001); V. V. Ryazanov, V. A. Oboz-

nov, A. Yu. Rusanov, *et al.*, Phys. Rev. Lett. **86**, 2427 (2001).
4. V. T. Petrashov, I. A. Sosnin, I. Cox, *et al.*, Phys. Rev. Lett. **83**, 3281 (1999).
5. E. I. Rashba and V. I. Scheka, Fiz. Tverd. Tela (Leningrad) **2**, 162 (1958) [Sov. Phys. Solid State **2** (1958)]; R. C. Gasella, Phys. Rev. Lett. **5**, 371 (1960).
6. R. Romestain, S. Geshwind, and G. E. Devlin, Phys. Rev. Lett. **39**, 1583 (1977).
7. Note that in the article by V. L. Geshkenbein and A. I. Larkin [Pis'ma Zh. Éksp. Teor. Fiz. **43**, 306 (1986) [JETP Lett. **43**, 395 (1986)]] a term of the same form was assumed to be present in the interface transmission amplitude owing to a difference in SO coupling on either side of the interface. However, the ability of the difference alone, i.e., without any electric field, to give rise to the term was never proved.
8. V. M. Edelstein, Phys. Rev. B **67**, 020505(R) (2003).
9. L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **37**, 1407 (1960) [Sov. Phys. JETP **10**, 998 (1960)].
10. P. Deift and E. Trubowitz, Commun. Pure Appl. Math. **32**, 121 (1979).
11. H. Kleinert, Fortschr. Phys. **26**, 565 (1978).
12. C.-R. Hu and V. Korenman, Phys. Rev. **178**, 684 (1969).
13. It should be mentioned that the triplet superconductivity on the F side of the F/S interface was also considered in the article by F. S. Bergeret, A. F. Volkov, and K. B. Efetov [Phys. Rev. Lett. **86**, 4096 (2001)] as a result of the magnetization rotation near the interface like that in the Bloch domain wall. In addition to the difference between models studied in the present and that paper, there is a fundamental difference between approaches applied. In that paper, the Usadel equation for the anomalous Green's matrix-function (AGF) as a function of the Cooper pair center of mass coordinates was solved. The antisymmetric components of the AGF were, as usually, attributed to the singlet superconductivity whereas the symmetric ones were interpreted as the appearance of the triplet superconductivity. However, the total AGF, $F(\mathbf{r}_1, \mathbf{r}_2)_{\alpha\beta}$ (as well as its singlet and triplet parts separately) must change sign at $(\alpha, \mathbf{r}_1) \rightleftarrows (\beta, \mathbf{r}_2)$ due to the Pauli principle. It is unclear how one can assure this property for the symmetric component found within the Usadel approach, i.e., without a consideration of the relative coordinate dependence of the AGF.

JETP Letters, Vol. 77, No. 4, 2003, pp. 187–190. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 217–221.

Spin Dynamics of Semiconductor Electrons in a Hybrid Ferromagnet/Semiconductor Structure¶

V. N. Gridnev

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

e-mail: gridnev@mail.ioffe.ru

Received December 11, 2002; in final form, January 20, 2003

We analyze the recent experimental study by R.J. Epstein *et al.* [Phys. Rev. B **65**, 121202 (2002)] on the spin dynamics of semiconductor electrons in a hybrid ferromagnet/semiconductor structure by using a simple model based on the Bloch equations. A comparison between the model calculations and the experimental observations shows that the spin relaxation rate is strongly anisotropic. We interpret this anisotropy as a manifestation of the exchange interaction between metallic and semiconductor electrons at the ferromagnet/semiconductor interface. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 72.25.Mk; 72.25.Rb; 73.40.-c; 75.70.-i

In recent years, much interest has been aimed at the physics of electron spin dynamics in bulk semiconductors and heterostructures [1]. An understanding of spin relaxation mechanisms is important because of the potential use of spin degrees of freedom in magnetoelectronics and quantum computation [2].

A hybrid ferromagnet/semiconductor (F/S) structure is often considered as a key element of magnetoelectronics. Usually, when considering the transport of spin-polarized electrons from a ferromagnetic metal into a semiconductor, it is tacitly assumed that the spin relaxation rate in the semiconductor is independent of the proximity of the ferromagnet. In most cases it is a good approximation, but it is questionable for samples in which the exchange coupling between metal and semiconductor electrons reaches a noticeable magnitude. It is well known that an exchange interaction is able to provide an efficient channel for the spin relaxation. The BAP mechanism of the spin relaxation in semiconductors is a well-known example [3].

In the recent experiment [4], it was observed by the method of time-resolved Faraday rotation that in the hybrid F/S structures, namely, Fe/*n*-GaAs and MnAs/*n*-GaAs, the proximity of the ferromagnetic metal affects the dynamical behavior of the spin polarization of semiconductor electrons. It was found that, in an external in-plane magnetic field, a short circularly polarized (CP) or linearly polarized (LP) pump laser pulse, with energy tuned near the GaAs band gap, induces the component of the spin polarization of semiconductor electrons $S_x(t)$:

$$S_x(t) = A_0 e^{-t/T_2}\cos(\omega t + \phi), \qquad (1)$$

where A_0 is the amplitude, T_2 is the effective transverse spin lifetime, $\omega = g\mu_B B/\hbar$ is the Larmor precession frequency, ϕ is the phase of the spin precession, and the x axis is directed along the sample normal. In the case of a CP pump pulse, which creates spin-polarized electrons (optically oriented) along the pump path, the well-known [5] result with $\phi = 0$ was reproduced. Although a LP pump creates spin-unpolarized carriers, the spin oscillations given by Eq. (1) with $\phi \approx -90°$ were also observed. This phenomenon was qualitatively explained in [4] as a manifestation of the exchange interaction between the ferromagnet and semiconductor electrons. Despite the absence of spin polarization in the semiconductor just after the arrival of a pump pulse, it arises due to the interlayer exchange interaction and oscillates with the Larmor frequency ω for several nanoseconds (to be exact, we note that the term "spin interaction" rather than "exchange interaction" was used in [4]).

Although the physical origin of the observed effect was justified in [4] in sufficient detail, many features of the effect remain unexplained. First, we should understand why the oscillation frequency ω is determined solely by the external magnetic field **B** and unaffected by the exchange interaction, whereas the amplitude A_0 of the spin oscillations varies (in the case of an LP pump) with the angle α between **B** and ferromagnet's magnetization **M** as $\sin\alpha$. Second, it should be explained why the spin relaxation time for an LP pump (~4 ns) differs from that for a CP pump (~2 ns). Finally, the obtained spin relaxation times are comparable by the order of magnitude with those obtained for *n*-GaAs without a ferromagnetic layer [6]. It looks as if, despite the pronounced exchange interaction between the metal and semiconductor electrons, this interaction has little effect on the spin relaxation rate. Of course, we keep in

¶This article was submitted by the author in English.

mind that the spin relaxation rate is strongly influenced by impurities and defects and, as a consequence, is sample-dependent, so that the comparison between different samples should be made with care. Nonetheless, the experiment [4] raises a question about the influence of the exchange interaction at the F/S interface on the spin relaxation rate of semiconductor electrons.

In this paper, we consider the spin dynamics of semiconductor electrons in the vicinity of the F/S interface and analyze the experimental observations of [4] by using the semiphenomenological Bloch equations. In accordance with the conditions of the experiment, we assume the external magnetic field **B** and the ferromagnet's magnetization **M** to be parallel to the F/S interface. We choose the x and y axes to be directed along the interface normal and the magnetization **M**, respectively (see [4] for details). The thickness of the n-GaAs layer (100 nm) in the heterostructure used in the experiment was much less than the laser wavelength (≈820 nm). For this reason, we neglect the variation of the spin polarization $\mathbf{S}(x, t)$ of semiconductor electrons across the layer and replace it with some effective value $\mathbf{S}(t)$. Next, we take into account the exchange interaction between the metal and semiconductor electrons by means of an effective exchange field $\mathbf{B}_{\text{ex}}$, which is parallel (or antiparallel) to **M** and independent of x. As a consequence of these simplifications, we obtain a model in which homogeneous spin polarization $\mathbf{S}(t)$ varies under the action of the external magnetic field **B** and the exchange field $\mathbf{B}_{\text{ex}}$ from an initial value $\mathbf{S}_0$ to the equilibrium one $\mathbf{S}_e = \chi(\mathbf{B}_{\text{ex}} + \mathbf{B})/g\mu_B$, where χ is the magnetic susceptibility, g is the electron g factor, and μ_B is the Bohr magneton. This model is rather crude and cannot be applied to the analysis of those properties of the spin dynamics which depend on the distribution of the spin polarization across the film (see the discussion below). However, we believe that some experimentally observed features of the spin dynamics are determined mostly by the geometrical characteristics of the system, such as the relative orientation of the external and exchange magnetic fields.

The equation of motion for the spin polarization $\mathbf{S}(t)$ is

$$d\mathbf{S}/dt = \mathbf{\Omega} \times \mathbf{S} - \hat{R} \cdot (\mathbf{S} - \mathbf{S}_e), \tag{2}$$

where $\mathbf{\Omega} = g\mu_B(\mathbf{B}_{\text{ex}} + \mathbf{B})/\hbar$ is the vector of the electron spin precession frequency and $\hat{R}$ is the relaxation tensor. In [4] it was shown that the external magnetic field **B** is modified by a nuclear hyperfine field $\mathbf{B}_n$. Because $\mathbf{B}_n$ is parallel to **B**, we assume that the nuclear field is included in **B**. The form of the relaxation tensor is crucial for the following consideration. We choose the relaxation tensor $\hat{R}$ to be diagonal for a given coordinate system: $R_{ij} = \delta_{ij}\Gamma_i$. This assumption physically means that the exchange field $\mathbf{B}_{\text{ex}}$ has a much larger influence on the spin relaxation than the external magnetic field **B**. We shall justify this by comparison of our model calculations with the experimental observations. It might seem that the anisotropy of the spin relaxation is redundant here, since the solution of Eq. (2) for $S_x(t)$ with the isotropic relaxation tensor $R_{ij} = \delta_{ij}/T_2$ has the required form, Eq. (1). However, for such a solution, $\omega = \Omega \propto |\mathbf{B} + \mathbf{B}_{\text{ex}}|$, which contradicts to the experiment. Moreover, as we shall show, for an LP pump ($\mathbf{S}_0 = 0$) and isotropic tensor $\hat{R}$, Eq. (2) gives $S_x(t) \equiv 0$. Thus, the anisotropy of the spin relaxation is essential for the explanation of the experimental observations.

We are interested in the experimentally measured component of the spin polarization $S_x(t)$. The anisotropy of the spin relaxation somewhat complicates the solution of Eq. (2) for $S_x(t)$, which differs now from Eq. (1) and has the form

$$S_x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + C_3 e^{\lambda_3 t}, \tag{3}$$

where λ_i are the eigenvalues of the matrix $H_{ij} = e_{ikj}\Omega_k - \Gamma_i\delta_{ij}$ and e_{ikj} is the unit antisymmetric tensor of rank three. The coefficients C_i should be determined from the initial conditions for the spin polarization $\mathbf{S}(t)$. Expression (3) for $S_x(t)$ describes both LP and CP pumps, with different initial conditions in either case and, as a consequence, with different coefficients C_i. Assuming $S_x(0) = S_{0,x}$ and $S_y(0) = S_z(0) = 0$, we obtain

$$C_1 = \frac{S_{0,x}D_1 + \Omega_z S_{e,y}(\Gamma_y - \Gamma_z)}{(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)}, \tag{4}$$

where

$$D_1 = \Gamma_x^2 - \Omega^2 + \Gamma_x(\lambda_2 + \lambda_3) + \lambda_2\lambda_3. \tag{5}$$

The coefficients C_2 and C_3 can be obtained from Eqs. (4) and (5) by cyclic permutation of the indices 1, 2, and 3. The eigenvalues λ_i can be determined from the third-order algebraic equation

$$\begin{aligned}(\lambda + \Gamma_x)(\lambda + \Gamma_y)(\lambda + \Gamma_z) + \Omega_y^2(\lambda + \Gamma_y) \\ + \Omega_z^2(\lambda + \Gamma_z) = 0.\end{aligned} \tag{6}$$

Analytical expressions for the eigenvalues λ_i are rather cumbersome and therefore not given here. In the following, we shall calculate λ_i approximately for particular relations between the spin relaxation rates Γ_i that are relevant to the experiment.

As is seen from Eq. (4), there are two physically different sources for the nonzero spin polarization $S_x(t)$. The first one is well known and originates from the initial spin polarization (optical orientation) $S_{0,x}$ created by a pump pulse. This is the case of a CP pump. The second source is not related with the optical orientation and is relevant to an LP pump. It comes into play when the external magnetic field is not parallel to the magnetization ($\Omega_z \neq 0$) and, additionally, the spin relaxation in the interface plane is anisotropic, that is, $\Gamma_y \neq \Gamma_z$. Physically, this means that under these conditions the emerging spin polarization $\mathbf{S}(t)$ is not parallel to $\mathbf{\Omega}$,

thus creating $S_x(t)$. Our primary interest here lies in the case of the LP pump, since it reveals some important properties of the exchange coupling between the ferromagnet and the semiconductor.

A distinctive feature of the spin oscillations induced by an LP pump is the dependence of their amplitude A_0 on the angle α between the external magnetic field **B** and the exchange field $\mathbf{B}_{ex}$. It was found experimentally that $A_0 \propto \sin\alpha$. Our model calculations also predict some dependence of A_0 on α. As is seen from Eq. (4), the form of this dependence is determined by the product $\Omega_z S_{e,y}$, which, in turn, depends on the relation between the external magnetic field and the exchange field. Since $\Omega_z \propto \sin\alpha$ and the same angular dependence was observed for $S_x(t)$ in the experiment, we conclude that $S_{e,y} = \chi(B_{ex} + B\cos\alpha)$ must be independent of α (within the experimental accuracy). This occurs only if $B_{ex} \gg B$. Thus, the observed dependence of $S_x(t)$ on α indicates that the exchange field is much larger than the external magnetic field ~1 T that was used in the experiment.

Now we turn to the temporal behavior of $S_x(t)$. As we have already noted, in the general case, Eq. (3) cannot be reduced to the decaying oscillations, Eq. (1), observed in the experiment. However, we shall show that, under special relations between the parameters of the considered system, Eq. (3) leads to almost the same temporal behavior of $S_x(t)$ as Eq. (1) and, at the same time, gives the true dependence of the Larmor frequency and the amplitude of the spin oscillations on the external magnetic field. Let us consider this point in more detail. As we consider the oscillatory behavior of the spin polarization $S_x(t)$, two (say, λ_1 and λ_2) of the three eigenvalues in Eq. (3) must be complex and inevitably complex conjugate, $\lambda_1 = \lambda_2^*$. Taking into account that $C_1 = C_2^*$, we conclude that the first two terms in Eq. (3) give decaying oscillations similar to Eq. (1) with $1/T_2 = -\mathrm{Re}\lambda_1 = -\mathrm{Re}\lambda_2$, $\omega = \mathrm{Im}\lambda_1 = -\mathrm{Im}\lambda_2$, and $\phi = \arg C_1$. In order for Eq. (3) to be reduced to Eq. (1), we must have a reason to neglect the third term in Eq. (3). This is possible in two cases. The first one is trivial and takes place if $C_3 = 0$. One can easily see from Eqs. (4)–(6) that this occurs if the in-plane relaxation is isotropic ($\Gamma_y = \Gamma_z$). As we have already noted, this case is not relevant to the experiment. Another situation, when the third term in Eq. (3) can be neglected, occurs if $|\lambda_3| \gg 1/T_2$. Then, this term rapidly goes to zero with time. In order to find the conditions under which λ_3 is large, we should analyze the dependence of the eigenvalues λ_i on the parameters Γ_i and Ω_i in the case of strong anisotropy of the spin relaxation. To this end, we compare the behavior of the terms in Eq. (6) when Γ_z increases. We notice that the second term in Eq. (6) can be neglected when $\Gamma_z \longrightarrow \infty$, and then $\lambda_i^{(0)} = \lambda_i(\Gamma_z \longrightarrow \infty)$ are given by

$$\lambda_{1,2}^{(0)} \approx -\frac{\Gamma_x + \Gamma_y}{2} \pm i\sqrt{\Omega_z^2 - \frac{(\Gamma_x - \Gamma_y)^2}{4}}, \qquad (7)$$
$$\lambda_3^{(0)} \approx -\Gamma_z.$$

Thus, for a large Γ_z, we obtain λ_i with the required properties: λ_3 is large and the Larmor frequency $\omega \approx \Omega_z \propto B_z$. It is this dependence of ω on the external magnetic field that was observed in the experiment. Note that the effective spin relaxation rate $1/T_2 = (\Gamma_x + \Gamma_y)/2$ is independent of Γ_z. It follows from Eq. (4) that the coefficients C_1 and C_2 become purely imaginary when λ_3 increases. Consequently, Eq. (3) reduces to Eq. (1) with $\phi = \pm 90°$, where the sign, plus or minus, depends on the sign of the exchange field, i.e., on whether the exchange coupling at the interface is ferromagnetic or antiferromagnetic.

In order to obtain the restrictions, imposed on the parameters Γ_i and Ω_i, under which $\lambda_i \simeq \lambda_i^{(0)}$, we estimate corrections $\delta\lambda_i$ to $\lambda_i^{(0)}$ for large but finite Γ_z and require that $|\delta\lambda_i| \ll |\lambda_i^{(0)}|$. Substituting $\lambda_i = \lambda_i^{(0)} + \delta\lambda_i$ into Eq. (6) and assuming $\Omega_z \gg \Gamma_x \simeq \Gamma_y$, we obtain $\delta\lambda_3 = -\delta\lambda_{1,2} \simeq \Omega_y^2/\Gamma_z$. The condition $|\delta\lambda_3| \ll |\lambda_3^{(0)}|$ gives $\Gamma_z \gg \Omega_y$, but from the second condition $|\delta\lambda_{1,2}| \ll |\lambda_{1,2}^{(0)}|$ we obtain the stronger inequality $\Gamma_z \gg \Omega_y^2/\Gamma_x$. The last relation is a strict mathematical definition of the term "the strong anisotropy of the spin relaxation" for a given problem.

What temporal behavior of $S_x(t)$ do we obtain if Γ_y rather than Γ_z increases? Interchanging the subscripts y and z in Eqs. (7), we again obtain Eq. (1), but with the Larmor frequency $\omega = \Omega_y \propto B_{ex} + B_y$. This is inconsistent with the experiment, where $\omega \propto B_z$.

Thus, the observation of the oscillations of the spin polarization for an LP pump, given by Eq. (1), indicates that the relaxation rate Γ_z significantly exceeds Γ_x and Γ_y. This fact is not evident from the experimental data, since the rate of decay of the Faraday rotation $1/T_2 = (\Gamma_x + \Gamma_y)/2$ is independent of Γ_z. Hence, the increase of Γ_z leaves T_2 unchanged.

To get a more quantitative characterization of the spin relaxation anisotropy, we consider a set of inequalities which are satisfied by Γ_i and Ω_i: (i) $\Gamma_z \gg \Omega_y^2/\Gamma_x$, (ii) $\Omega_y \gg \Omega_z$, and (iii) $\omega \gg 1/T_2$. The first of these inequalities allows us to reduce Eq. (3) to Eq. (1), the second is obtained from the fit of the angular dependence of $S_x(t)$, and the third follows from the experiment, where $\omega \simeq 10$ ns^{-1} at $B = 0.25$ T and $1/T_2 \simeq 0.25$ ns^{-1}. Since the restrictions imposed on the parameters Γ_i and

Ω_i are determined by the three inequalities (i)–(iii), the lower bound for the relaxation rate Γ_z is allowed to vary over the whole range, which we estimate as $\sim(10^3$–$10^5)/T_2$, and the corresponding range for the relaxation time $T_z \sim 0.01$–1 ps.

The physical picture of the motion of the spin polarization is very simple. When Γ_z is large, S_z rapidly approaches its equilibrium value $S_{e,z}$ due to the large damping. This means that $\mathbf{S}(t)$ precesses around the z axis. The frequency of such precession is determined by the z component of the magnetic field and is independent of the exchange field.

So far, we have considered the case of an LP pump. The case of a CP pump can be considered quite analogously. Now the coefficients C_i in Eq. (3) should be calculated taking into account that the initial spin polarization $S_{0,x}$ is nonzero, since a CP pump creates spin-polarized electrons. Assuming again the validity of Eqs. (7), we find that $S_x(t)$ is given by Eq. (1) with the phase ϕ depending on the magnitude of the initial spin polarization $S_{0,x}$. When $S_{0,x}$ increases, then $\phi \longrightarrow 0$, and Eq. (1) reproduces the well-known decaying spin oscillations in a transverse magnetic field [5]. These oscillations were observed in [4] for a CP pump. Although, in the case of a CP pump, Eq. (1) follows from the Bloch equation for the isotropic in-plane relaxation, the frequency ω remains dependent on the anisotropy. It was found in the experiment that the oscillation frequencies are the same for both pump polarizations. This means that the anisotropy of the spin relaxation is equally essential for the explanation of the experimental observations for both pump polarizations.

If we compare the relaxation times for LP and CP pumps, we see a distinction between our model calculations and the experimental data. Our calculations give equal relaxation times $T_2 = 2(\Gamma_x + \Gamma_y)^{-1}$ for LP and CP pumps, while the relaxation times measured experimentally differ: $T_2 = 4$ ns for an LP pump and $T_2 = 2$ ns for a CP pump. A possible reason for this discrepancy is the inhomogeneity of the spin polarization combined with the different origin of the spin polarization for LP and CP pumps. As opposed to the case of an LP pump, the spin polarization for a CP pump exists without the exchange coupling due to the optical orientation. This leads to the different profiles of $S_x(x, t)$ across the film for LP and CP pumps. Since the exchange coupling decreases with distance from the interface, the variation of $S_x(x, t)$ with x affects the relaxation rates and, in turn, leads to the different T_2 for LP and CP pumps. Our model fails to describe this effect, since it does not take into account the spatial variations of the exchange coupling and the spin polarization.

In conclusion, we have analyzed the experimental results of [4] on the spin dynamics in the hybrid F/S structure by using semiphenomenological Bloch equations. It has been shown that qualitative agreement between the model calculations and the experimental observations can be achieved only under certain conditions imposed on the parameters of the system. First, we have shown that the exchange field, representing the effect of exchange coupling between the ferromagnet and the semiconductor on semiconductor electrons, significantly exceeds the external magnetic field. Second, the spin relaxation rate of semiconductor electrons near the F/S interface is strongly anisotropic. The spin relaxation rate depends on the orientation of the spin polarization in the interface plane and reaches its maximum when the spin polarization is perpendicular to the ferromagnet's magnetization. In a sense, this anisotropy is hidden, since it weakly affects the rate of decay of the Faraday rotation beating and manifests itself only through the form of the beating. For this reason, we have obtained only a rough estimate of the anisotropy. The maximum value of the spin relaxation rate exceeds the bulk one approximately by a factor of 10^3–10^5. The corresponding relaxation times are on the order of 0.01–1 ps. Such fast spin relaxation is typical for magnetic semiconductors. This indicates that the exchange coupling between the ferromagnet and the semiconductor provides an additional channel for the spin relaxation. We may speculate that the anisotropy of the spin relaxation is related with the anisotropy of the surface spin excitation spectrum of the ferromagnet. This point requires additional theoretical and experimental investigations.

The author thanks K.A. Chao for reading the manuscript and for criticism. This work was supported by the Russian Foundation for Basic Research and by programs of Russian Ministry of Science.

REFERENCES

1. L. Vina, J. Phys.: Condens. Matter **11**, 5929 (1999).
2. Special issue on semiconductor spintronics, Semicond. Sci. Technol. **17**, 275 (2002).
3. G. L. Bir, A. G. Aronov, and G. E. Pikus, Zh. Éksp. Teor. Fiz. **69**, 1389 (1975) [Sov. Phys. JETP **42**, 705 (1975)].
4. R. J. Epstein, I. Malajovich, R. K. Kawakami, *et al.*, Phys. Rev. B **65**, 121202(R) (2002).
5. D. D. Awschalom and J. M. Kikkawa, Phys. Today **52**, 33 (1999).
6. J. M. Kikkawa and D. D. Awschalom, Phys. Rev. Lett. **80**, 4313 (1998).

JETP Letters, Vol. 77, No. 4, 2003, pp. 191–195. From Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 222–226.

Surface Instability of a Multicomponent Condensate and Andreev–Bashkin Effect¶

D. A. Abanin

Low-Temperature Laboratory, Helsinki University of Technology, FIN-02015 HUT, Finland
Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117334 Russia
e-mail: abanin@itp.ac.ru

Received January 20, 2003

It is shown that the surface of a liquid consisting of several interpenetrating superfluids moving with different velocities becomes unstable at some threshold. We demonstrate that the criterion for the onset of the instability changes in the presence of dissipative interaction between the surface and the environment. Possible physical applications of the surface instability are discussed. © *2003 MAIK "Nauka/Interperiodica"*.

PACS numbers: 47.37.+q; 47.20.-k; 67.60.-g; 03.75.Mn

1. INTRODUCTION

Analogues of the classical Kelvin–Helmholtz (KH) instability are relevant for various physical systems. For instance, it was recently proposed [1] that the analogue of KH instability for a mixture of two superfluids can provide a possible mechanism for pulsar glitches [2]. Interest in KH instabilities has also been revived in connection with the observation of the instability arising on the interface separating two superfluids, ^{3}He-A and ^{3}He-B [3]. The results of this experiment are in excellent agreement with the criterion for the onset of the instability derived in [4]. Free surface instability of this class was discussed in [5].

However, in most interesting physical situations, such as neutron stars and spin Bose–Einstein condensates [6], several interpenetrating superfluids are present. Such mixtures of superfluids are known to exhibit the Andreev–Bashkin effect [7]; that is, the mass current of each condensate is carried by the velocities of the others:

$$\mathbf{j}_i = \sum_k \rho_{ik}\mathbf{v}^k, \quad i = 1, \ldots, n, \tag{1}$$

where $\mathbf{v}^i$ and $\mathbf{j}_i$ are the superfluid velocity and the mass current of the ith component of the condensate. The density of kinetic energy of such condensates can be written in the following form:

$$E_k = \frac{1}{2}\sum_{i,j} \rho_{ij}\mathbf{v}^i\mathbf{v}^j. \tag{2}$$

In the present letter, we show that the free surface of a multicomponent condensate exhibiting the Andreev–Bashkin effect becomes unstable when the velocities of its constituents reach certain values. There exist two different thresholds characterizing the instability, the first resulting from the dissipative interaction of the surface with the environment, which is typically small, and the second corresponding to the classical dynamic instability. When the system is between these thresholds, the time of instability development is determined by the interaction with the environment; after passing through the second threshold, the characteristic time of instability evolution quickly decreases. In some cases, it is possible that there is no intermediate region, that is, that the instability develops very quickly from its very beginning.

We propose that this surface instability should arise earlier than the bulk instability described in [1]. Indeed, the critical value of velocity obtained in [1] is comparable with the sound velocity in the bulk, while the critical velocity for arising surface instability is determined by the field stabilizing the surface and should therefore be much smaller.

There is another possible application of the surface instability arising in mixtures of several interpenetrating superfluids: this mechanism could be relevant for the instability of a Bose–Einstein condensate with a multicomponent order parameter (see [6] and references therein).

The paper is organized as follows. First we derive the criterion for the thermodynamic instability, then the condition for the emergence of dynamic instability is obtained. We analyze how the presence of the interaction of the surface with the environment affects the instability criterion. After that, we demonstrate that in the "shallow water" limit the physics of the instability remains the same but can be described in terms of the effective ripplon metric.

¶This article was submitted by the author in English.

0021-3640/03/7704-0191$24.00 © 2003 MAIK "Nauka/Interperiodica"

2. THERMODYNAMIC INSTABILITY

In this section, we consider a thermodynamic instability arising on a free surface of several interpenetrating superfluids moving with different superfluid velocities $\mathbf{v}^i$. We suppose that the surface in equilibrium coincides with the plane $z = 0$ and that superfluid velocities $\mathbf{v}^i$ are equal to $\mathbf{u}^i$. As we will show in Section 3, this instability results from the interaction between the surface of the liquid and the environment. The instability emerges when the energy of static perturbations becomes negative in the container frame (or, equivalently, in the rest frame of the normal components, where normal velocity $\mathbf{v}_n = 0$).

Let us find the energy of the perturbation with wave vector $\mathbf{k} = (k_x, k_y)$. First of all, the deformation $\zeta = a\sin\mathbf{kr}$ of the surface causes the perturbation $\delta\mathbf{v}^i = \nabla\Phi^i$ of the superfluid velocity fields $\mathbf{v}^i$. Therefore, the free energy functional for the perturbation is given by

$$\mathcal{F} = \frac{1}{2}\int dxdy[F\zeta^2 + \sigma((\partial_x\zeta)^2 + (\partial_y\zeta)^2)] + \frac{1}{2}\int_{-\infty}^{\zeta} dz\int dxdy\left[\sum_{i,j}\rho_{ij}\delta\mathbf{v}^i\delta\mathbf{v}^j\right], \quad (3)$$

where F is the external field stabilizing the surface, e.g., the gravitational field, and σ is surface tension. Since the perturbation is stationary, the continuity equation for each component reduces to $\Delta\Phi^i = 0$. Hence, Φ^i has the following form:

$$\Phi^i = A^i\exp kz\cos\mathbf{kr}. \quad (4)$$

We can express constants A^i via a using the boundary conditions

$$\mathbf{u}^i\frac{\partial\zeta}{\partial\mathbf{r}} = \frac{\partial\Phi^i}{\partial z}. \quad (5)$$

This gives $A^i = a(\mathbf{u}^i\hat{\mathbf{k}})$, where $\hat{\mathbf{k}} = \mathbf{k}/k$. Substituting these values of A^i into the free energy (3), one obtains the energy of the perturbation:

$$E(k) = \frac{a^2}{4}\left[F + k^2\sigma - k\sum_{i,j}\rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})(\mathbf{u}^j\hat{\mathbf{k}})\right]. \quad (6)$$

Thus, the criterion for the onset of the thermodynamic instability is given by

$$\frac{1}{2}\max_{\hat{\mathbf{k}}}\sum_{i,j}\rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})(\mathbf{u}^j\hat{\mathbf{k}}) = \sqrt{\sigma F}. \quad (7)$$

3. DYNAMIC INSTABILITY

Now we proceed to the investigation of the analogue of the classical KH dynamical instability for several interpenetrating superfluids.

This instability emerges when the frequency of some surface wave acquires a positive imaginary part. It appears that the threshold for arising of the instability depends substantially on the presence of dissipation. Even a very small interaction of the superfluid with the environment (for example, friction between the surface of the superfluid and the container walls or between the surface and normal component) moves the threshold to another value. Let us first suppose that the interaction of the surface with the environment is absent.

The Euler equation in our case reads

$$\sum_i\frac{\partial j^i_\beta}{\partial t} + \sum_{i,j}\rho_{ij}v^i_\alpha\nabla_\alpha v^j_\beta = -\nabla_{\beta p} - F_\beta. \quad (8)$$

The solution of (8) corresponding to a small-amplitude surface wave with frequency ω and wave vector $\mathbf{k} = (k_x, k_y)$ can be chosen in the following form:

$$v^i_x = u^i_x + ik_xe^{i(\mathbf{kr}-\omega t)}e^{kz}A^i, \quad (9)$$

$$v^i_y = u^i_y + ik_ye^{i(\mathbf{kr}-\omega t)}e^{kz}A^i, \quad (10)$$

$$v^i_z = ke^{i(\mathbf{kr}-\omega t)}e^{kz}A^i, \quad (11)$$

$$p = -Fz + ie^{i(\mathbf{kr}-\omega t)}e^{kz} \times\left[\omega\sum_{i,j}\rho_{ij}A^i - \sum_{i,j}\rho_{ij}(\mathbf{u}^i\mathbf{k})A^j\right], \quad (12)$$

where A^i, $i = 1, \ldots, n$ are some constants. Apart from equation (8), our system must satisfy boundary conditions:

$$-p = \sigma\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\zeta, \quad (13)$$

$$\delta v^i_z - \mathbf{u}^i\frac{\partial\zeta}{\partial\mathbf{r}} = \frac{\partial\zeta}{\partial t}. \quad (14)$$

Using Eq. (13), we can express the surface deformation ζ in terms of the constants A^i:

$$\zeta = ie^{i(\mathbf{kr}-\omega t)}e^{kz}/(F + k^2\sigma) \times\left[\omega\sum_{i,j}\rho_{ij}A^i - \sum_{i,j}\rho_{ij}(\mathbf{u}^i\mathbf{k})A^j\right]. \quad (15)$$

Substituting (15) into (14) and making use of (11), we obtain a set of equations for the constants A^i:

$$k(F + k^2\sigma)A^i = (\omega - \mathbf{u}^i\mathbf{k})\left[\sum_{j,l}\rho_{jl}A^j(\omega - \mathbf{u}^l\mathbf{k})\right], \quad (16)$$

$$i = 1, \ldots, n.$$

It is convenient to rewrite equation (16) in the form

$$\sum_j \phi_{ij} A^j = 0, \quad i = 1, \ldots, n, \tag{17}$$

where

$$\phi_{ij} = k(F + k^2\sigma)\delta_{ij} - (\omega - \mathbf{u}^i\mathbf{k})\sum_l \rho_{jl}(\omega - \mathbf{u}^l\mathbf{k}). \tag{18}$$

For system (17) to have a nontrivial solution, it is necessary that

$$\det\phi_{ij} = 0. \tag{19}$$

It can be shown (see *Appendix*) that condition (19) is equivalent to the following equation:

$$k(F + k^2\sigma) = \sum_{ij} \rho_{ij}(\omega - \mathbf{u}^i\mathbf{k})(\omega - \mathbf{u}^j\mathbf{k}). \tag{20}$$

Equation (20) determines the ripplon spectrum. The frequency of the ripplon with the wave vector directed along $\hat{\mathbf{k}}$ acquires an imaginary part when

$$\rho\sum_{i,j} \rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})(\mathbf{u}^j\hat{\mathbf{k}}) - \left(\sum_{i,j} \rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})\right)^2 = 2\rho\sqrt{F\sigma}, \tag{21}$$

where $\rho = \sum_{i,j}\rho_{ij}$. Equation (21) provides the criterion for the onset of the instability, which is analogous to the classical KH instability. In terms of "mean velocity"

$$\mathbf{v} = \frac{\mathbf{j}_{\text{tot}}}{\rho} = \frac{\sum_{i,j}\rho_{ij}\mathbf{u}^i}{\rho}, \tag{22}$$

Eq. (21) becomes more transparent:

$$\sum_{i,j} \rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})(\mathbf{u}^j\hat{\mathbf{k}}) - \rho(\mathbf{v}\hat{\mathbf{k}})^2 = 2\sqrt{F\sigma}. \tag{23}$$

As was pointed out in [4], criterion (23) will change if one takes into account the dissipative force arising when the surface is moving with respect to the container walls. This force can be written in the following form:

$$F_{\text{fr}} = -\Gamma\partial_t\zeta. \tag{24}$$

In the presence of friction, Eq. (20) for the spectrum of ripplons is modified as follows:

$$k(F + k^2\sigma) - i\Gamma\omega k = \sum_{ij} \rho_{ij}(\omega - \mathbf{u}^i\mathbf{k})(\omega - \mathbf{u}^j\mathbf{k}). \tag{25}$$

When the superfluid velocities reach critical values, the imaginary part of the frequency crosses zero. Then, from Eq. (25) it follows that the real part of the frequency must also cross zero. Consequently, the instability condition is

$$\frac{1}{2}\max_{\hat{\mathbf{k}}}\sum_{i,j} \rho_{ij}(\mathbf{u}^i\hat{\mathbf{k}})(\mathbf{u}^j\hat{\mathbf{k}}) = \sqrt{\sigma F}. \tag{26}$$

This equation coincides with condition (7) for the thermodynamic instability. From Eqs. (23) and (26), it can be easily seen that thermodynamic instability always arises earlier than the classical dynamic instability. In the region between these two thresholds, the instability develops rather slowly (because the time of instability development is proportional to the friction parameter Γ, which is usually relatively small), and after passing through the dynamic threshold it develops quickly.

Criteria (23) and (26) can be represented in another form, which sometimes proves more suitable. In order to obtain this form, we note that instability occurs first for the ripplons with wave vector $k_0 = \sqrt{F/\sigma}$. We further rewrite Eq. (25) with $\mathbf{k} = k_0\hat{\mathbf{k}}$ in terms of the effective ripplon metric:

$$\begin{aligned} g^{\mu\nu}\hat{k}_\mu\hat{k}_\nu &= i\Gamma\frac{\omega}{k_0}, \\ \hat{k}_\mu &= (-\omega/k_0, \hat{k}_\alpha), \quad \alpha = x, y. \end{aligned} \tag{27}$$

The effective ripplon metric in equation (27) is given by

$$\begin{aligned} g^{00} &= -\rho, \quad g^{0\alpha} = -j_{\text{tot}_\alpha}, \\ g^{\alpha\beta} &= 2\sqrt{F\sigma}\delta^{\alpha\beta} - \sum_{i,j}\rho_{ij}u^i_\alpha u^j_\beta. \end{aligned} \tag{28}$$

It can be easily seen that criterion (23) for dynamic instability is equivalent to the condition

$$\det g_{\mu\nu} = \infty, \tag{29}$$

while criterion (26) for thermodynamic instability is equivalent to

$$g_{00} = 0. \tag{30}$$

4. CASE OF TWO INTERPENETRATING SUPERFLUIDS

Now let us consider in more detail the case when only two interpenetrating superfluids are present. Equation (25) for the spectrum of ripplons reads

$$\begin{aligned} \frac{F + k^2\sigma}{k} - i\Gamma\frac{\omega}{k} &= \rho_{11}\left(\frac{\omega}{k} - \mathbf{u}^1\hat{\mathbf{k}}\right)^2 \\ &+ \rho_{22}\left(\frac{\omega}{k} - \mathbf{u}^2\hat{\mathbf{k}}\right)^2 + 2\rho_{12}\left(\frac{\omega}{k} - \mathbf{u}^1\hat{\mathbf{k}}\right)\left(\frac{\omega}{k} - \mathbf{u}^2\hat{\mathbf{k}}\right). \end{aligned} \tag{31}$$

It is convenient to rewrite Eq. (31) in terms of mean velocity (22) and relative velocity $\mathbf{u} = \mathbf{u}^1 - \mathbf{u}^2$:

$$\frac{F + k^2\sigma}{k} - i\Gamma\frac{\omega}{k} = \rho\left[\left(\frac{\omega}{k} - \mathbf{v}\hat{\mathbf{k}}\right)^2 + \alpha(\mathbf{u}\hat{\mathbf{k}})^2\right], \quad (32)$$

where $\alpha = (\rho_1\rho_2 - \rho_{12}^2)/\rho^2$. Thus, the frequency of the surface wave with wave vector $\mathbf{k}$ satisfies the following equation:

$$\frac{\omega}{k} = \mathbf{v}\hat{\mathbf{k}} \pm \sqrt{\frac{F + k^2\sigma}{\rho k} - i\frac{\Gamma}{\rho}\frac{\omega}{k} - \alpha(\mathbf{u}\hat{\mathbf{k}})^2}. \quad (33)$$

Using (33), one immediately finds that the first instability threshold, corresponding to the onset of the thermodynamic instability, is determined by the equation

$$\frac{2\sqrt{F\sigma}}{\rho} = \max_{\hat{\mathbf{k}}}[\alpha(\mathbf{u}\hat{\mathbf{k}})^2 + (\mathbf{v}\hat{\mathbf{k}})^2], \quad (34)$$

and the second instability threshold is given by

$$\frac{2\sqrt{F\sigma}}{\rho} = \alpha u^2. \quad (35)$$

Note that, in the absence of the Andreev–Bashkin effect, results (34) and (35) coincide with those for the instability of the interface separating two one-component superfluids [4]. The only difference between these situations consists in the fact that the stabilizing field F in the case of surface instability is proportional to the total density of the mixture, while in the case of interface instability it is proportional to the difference of the densities of the separated liquids.

As we have already mentioned, the thermodynamic instability always occurs before the dynamic one. However, under certain circumstances, these instabilities can arise simultaneously. Indeed, if u is perpendicular to v and $v \leq \sqrt{\alpha}u$, then Eqs. (34) and (35) coincide. Therefore, the instability starts at $u^2 = 2\sqrt{F\sigma}/\rho$, and it is "strong" in the sense that the time of its development quickly decreases as u grows. This scenario is analogous to the violation of the "cosmic censorship" principle in quantum liquids [8], when the effective ripplon metric acquires a naked singularity.

5. SHALLOW WATER LIMIT.

In the previous sections, we supposed that the depth of the liquid h was large, so that

$$\sqrt{\frac{F}{\sigma}}h \gg 1. \quad (36)$$

In this section, we demonstrate that, in the limit of "shallow water" (i.e., $\sqrt{F/\sigma}\,h \ll 1$), the mechanism of the instability remains the same. Indeed, the equation for the ripplon spectrum in this limit has the following form:

$$\sum_{i,j}\rho_{ij}(\omega - \mathbf{u}^i\mathbf{k})(\omega - \mathbf{u}^j\mathbf{k}) = Fhk^2 + \sigma hk^4 - i\Gamma(k)\omega. \quad (37)$$

It can be rewritten in terms of the effective ripplon metrics:

$$g^{\mu\nu}k_\mu k_\nu = i\Gamma(k)\omega - \sigma hk^4, \quad (38)$$

$$k_\mu = (-\omega, k_\alpha), \quad \alpha = x, y, \quad (39)$$

$$g^{00} = -\rho, \quad g^{0\alpha} = -j_{\mathrm{tot}_\alpha}, \quad (40)$$

$$g^{\alpha\beta} = Fh\delta^{\alpha\beta} - \sum_{ij}\rho_{ij}u_\alpha^i u_\beta^j. \quad (41)$$

Note that, in contrast to metric (28), metric (40), (41) is valid for all long-wavelength ripplons.

Again, there would be two thresholds for the onset of the instability, which are completely analogous to (23) and (26), the only difference being that the ripplons that are responsible for the instability have wave vectors close to zero. The thermodynamic instability threshold corresponds to the condition (30) $g_{00} = 0$, which ensures the appearance of the horizon, as was recently proposed by Schutzhold and Unruh [9]; the condition for the onset of the dynamic instability is det $g_{\mu\nu} = \infty$, in agreement with [8].

6. CONCLUSIONS

To conclude, we have found that the free surface of several interpenetrating superfluids, moving with different velocities, becomes unstable at some threshold. The criterion for such instability contains the off-diagonal densities ρ_{ij}, characterizing the Andreev–Bashkin effect. This kind of instability could provide a triggering mechanism for pulsar glitches. It also probably takes place in atomic Bose–Einstein condensates.

I am grateful to G.E. Volovik for suggesting this problem and for many useful discussions. I would also like to thank the Low-Temperature Laboratory of Helsinki University of Technology, where this work was done, for kind hospitality. This work was supported in part by the Russian Foundation for Basic Research (project no. 02-02-16218), the International Center for Fundamental Physics, and the Dynasty Foundation.

APPENDIX

Here, we present a calculation of the determinant of the matrix (18). It proves useful to introduce the following notations:

$$\alpha^i = \omega - \mathbf{u}^i\mathbf{k}, \quad (42)$$

$$\beta^l = \sum_i \rho_{il}(\omega - \mathbf{u}^i\mathbf{k}), \tag{43}$$

$$\gamma = k(F + k^2\sigma). \tag{44}$$

Then matrix (18) takes the form

$$\phi = \begin{pmatrix} \gamma - \alpha^1\beta^1 & -\alpha^1\beta^2 & \dots & -\alpha^1\beta^n \\ -\alpha^2\beta^1 & \gamma - \alpha^2\beta^2 & \dots & -\alpha^2\beta^n \\ \dots & \dots & \dots & \dots \\ -\alpha^n\beta^1 & -\alpha^n\beta^2 & \dots & \gamma - \alpha^n\beta^n \end{pmatrix}. \tag{45}$$

The determinant of matrix ϕ_{ij} can be written as follows:

$$\det\phi = \prod_{i=1}^{n}\alpha^i \times \det\begin{pmatrix} \gamma/\alpha^1 - \beta^1 & -\beta^2 & \dots & -\beta^n \\ -\beta^1 & \gamma/\alpha^2 - \beta^2 & \dots & -\beta^n \\ \dots & \dots & \dots & \dots \\ -\beta^1 & -\beta^2 & \dots & \gamma/\alpha^n - \beta^n \end{pmatrix}. \tag{46}$$

In order to calculate the determinant of the matrix on the r.h.s. of (46), let us subtract its $(n-1)$th line from its nth line, its $(n-2)$th line from its $(n-1)$th line, etc. This gives

$$D_n = \det\begin{pmatrix} \gamma/\alpha^1 - \beta^1 & -\beta^2 & \dots & -\beta^n \\ -\gamma/\alpha^1 & \gamma/\alpha^2 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & -\gamma/\alpha^{n-1} & \gamma/\alpha^n \end{pmatrix}. \tag{47}$$

One can easily obtain a recurrent formula for D_n:

$$D_n = \gamma/\alpha^n D_{n-1} - \beta^n\prod_{i=1}^{n-1}\gamma/\alpha^i. \tag{48}$$

Thus,

$$\det\phi = \gamma^n - \gamma^{n-1}\sum_{i=1}^{n}\alpha^i\beta^i, \tag{49}$$

and it equals zero when

$$\gamma = \sum_{i=1}^{n}\alpha^i\beta^i. \tag{50}$$

REFERENCES

1. N. Andersson, G. L. Comer, and R. Prix, astro-ph/0211151.
2. D. Langlois, in *Vortices in Unconventional Superconductors and Superfluids*, Ed. by R. P. Huebener, N. Schopohl, and G. E. Volovik (Springer, Berlin, 2002), Springer Ser. Solid-State Sci. **132**, 65 (2002).
3. R. Blaauwgeers, V. B. Eltsov, G. Eska, *et al.*, Phys. Rev. Lett. **89**, 155 301 (2002).
4. G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 491 (2002) [JETP Lett. **75**, 418 (2002)].
5. S. E. Korshunov, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 496 (2002) [JETP Lett. **75**, 423 (2002)].
6. A. Fetter and A. Svidzinsky, J. Phys.: Condens. Matter **13**, R135 (2001).
7. A. F. Andreev and E. Bashkin, Zh. Éksp. Teor. Fiz. **69**, 319 (1975) [Sov. Phys. JETP **42**, 164 (1975)].
8. G. E. Volovik, gr-qc/0301043.
9. R. Schutzhold and W. G. Unruh, Phys. Rev. D **66**, 044019 (2002).

JETP Letters, Vol. 77, No. 4, 2003, pp. 196–198. Translated from Pis'ma v Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 77, No. 4, 2003, pp. 227–229.

Practical Implementation of a Nontrivial Evolution "Unravelling" for a Resonantly Fluorescing Atom

L. V. Il'ichev

Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences, Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia

e-mail: leonid@iae.nsk.su

Received November 15, 2002

One of the consequences of the implementation of the so-called nontrivial "unravelling" of the resonance fluorescence of a two-level atom is considered. Namely, we consider the artificial re-aiming of the environment at the detection of a new set of events that are different from the photoemission to the components of a resonance fluorescence triplet. The atomic states obtained after the detection of new events are studied, and the element providing the maximum probability of the atom occurring in the excited state is found for the simplest unravelling class. © *2003 MAIK "Nauka/Interperiodica".*

PACS numbers: 32.50.+d; 03.65.Yz

At present, the progress of scientific technology allows one to set up experiments on single atoms. This is a very important achievement, especially in connection with the development of quantum computer science. In essence, these experiments represent quantum-state "engineering." In the approximation where an atom can be considered as a closed system, engineering amounts to the implementation of the appropriate Hamiltonian atomic evolution, e.g., by using the desired spatiotemporal configuration of light fields. In reality, however, an atom is an open quantum system that exchanges energy and information with its environment. The scope for quantum engineering is then extended, because one can rearrange the atomic environment by varying the character of energy and information exchange. In this communication, the version of "minimal" (in a sense that will be explained below) environment rearrangement is considered for a resonantly fluorescing atom.

Let us introduce the necessary notions and notation. The simplest model of a two-level atom, with the ground $|g\rangle$ and excited $|e\rangle$ states interacting with a monochromatic electromagnetic field, is considered. The so-called "dressed" states diagonalizing the atomic Hamiltonian are

$$\begin{aligned} |1\rangle &= \cos\theta|e\rangle + \sin\theta|g\rangle, \\ |2\rangle &= \cos\theta|g\rangle - \sin\theta|e\rangle, \end{aligned} \tag{1}$$

where the displacement angle θ is related to the Rabi frequency Ω and the detuning Δ of light frequency from the resonance as $\tan 2\theta = -\Omega/\Delta$. The corresponding atomic Hamiltonian is $\hat{H} = \sqrt{\Omega^2 + \Delta^2}(|1\rangle\langle 1| - |2\rangle\langle 2|)/2$. The equation for the evolution of the atomic density matrix has the form

$$\partial_t \hat{\varrho} = -i[\hat{H}, \hat{\varrho}] + \mathcal{D}[\hat{\varrho}]. \tag{2}$$

The term responsible for the spontaneous transitions is taken in the Lindblad form [1]

$$\mathcal{D}[\hat{\varrho}] = \hat{L}\hat{\varrho}\hat{L}^\dagger - \frac{1}{2}\{\hat{L}^\dagger\hat{L}, \hat{\varrho}\}_+, \tag{3}$$

where $\hat{L} = \sqrt{\gamma}|g\rangle\langle e|$, and γ is the spontaneous-decay rate of the excited state. In terms of dressed states (1), the Lindblad spontaneous-decay operator is written as

$$\hat{L} = \sum_\alpha \hat{L}_\alpha, \tag{4}$$

where $\alpha \in \{-, 0, +\}$ and

$$\begin{aligned} \hat{L}_0 &= \sin\theta\cos\theta(|1\rangle\langle 1| - |2\rangle\langle 2|), \\ \hat{L}_- &= -\sin^2\theta|1\rangle\langle 2|, \quad \hat{L}_+ = \cos^2\theta|2\rangle\langle 1|. \end{aligned} \tag{5}$$

In the so-called secular approximation [2]

$$\gamma \ll \sqrt{\Omega^2 + \Delta^2}, \tag{6}$$

spontaneous-decay superoperator $\mathcal{D}$ (3) can be replaced by

$$\sum_\alpha \mathcal{D}_\alpha, \tag{7}$$

where

$$\mathcal{D}_\alpha[\hat{\varrho}] = \hat{L}_\alpha\hat{\varrho}\hat{L}_\alpha^\dagger - \frac{1}{2}\{\hat{L}_\alpha^\dagger\hat{L}_\alpha, \hat{\varrho}\}_+. \tag{8}$$

The appearance of three operators instead of one Lindblad operator $\hat{L}$ is crucial for the further consideration.

To illustrate this, we make a short excursion to note that the energy and information exchange between the atom and its macroscopic environment occurs as a series of events in the form of abrupt and well time-localized events that correspond to a change in the environment state [3] as a result of the absorption of spontaneously emitted photons. In the case of an experimentally organized environment, the event appears as a detector snap (photocount). The Lindblad operator $\hat{L}$ corresponds to this event. Clearly, one cannot speak about the frequency of a detected photon, because the photocount is localized in time. However, within the framework of secular approximation (6), one can perform crude frequency selection for the detected spontaneous photons, because the superoperator $\mathcal{D}_\alpha$ is responsible for the photoemission into the left ($\alpha = \{-\}$), central ($\alpha = \{0\}$), and right ($\alpha = \{+\}$) components of the resonance fluorescence components. From the operational point of view, strong inequality (6) allows one to place a system of spectral filters between the atom and its environment and channel spontaneous photons into one of the three detectors, depending on the photon frequency (Fig. 1). The inverse Q factor (spectral width of filters) f should satisfy the condition

$$\gamma \ll f \ll \sqrt{\Omega^2 + \Delta^2}. \quad (9)$$

The right inequality in this expression provides a reliable spectral selection of the spontaneous photons among the triplet components, while the left inequality guarantees that photons will not be accumulated in filters. Otherwise, one should introduce the field of spontaneously emitted photons, and it would be incorrect to identify the photoemission and photon-detection events within the common notion of event.

Due to the appearance of three Lindblad operators $\hat{L}_\alpha$ ($\alpha \in \{-, 0, +\}$) and the corresponding three types of events, transformations of the $\{\hat{L}_\alpha\}$ set which do not change spontaneous-decay superoperator (7) (and, hence, the equation for the evolution of density matrix [4]) appear:

$$\hat{L}_\alpha \longrightarrow \hat{L}_\alpha(U) = \sum_{\alpha'} U_{\alpha\alpha'}\hat{L}_{\alpha'}, \quad (10)$$

where $U \in \mathcal{SU}(3)$ is an arbitrary unitary matrix of the third rank. The choice of one or another transformation (10) will be referred below (following Carmycle) to as unravelling of the evolution equation. The information aspect of the freedom in choosing the unravelling in quantum optics was touched on in [5]. Transformation (10) can be practically implemented using a system of mirrors and beam splitters placed between the spectrum- selecting and detector systems. At each beam splitter, two paths cross, along which the spontaneously emitted photons are channeled, so that to realize arbitrary transformation (10) three splitters with the appropriately chosen transmission and reflection coefficients would suffice in the general case. Then, the snap of detector α can no longer be interpreted as the detection of a photon from a certain triplet component. This new event, which corresponds to the operator $\hat{L}_\alpha(U)$, is, in a sense, a "superposition" with respect to the set of original events.

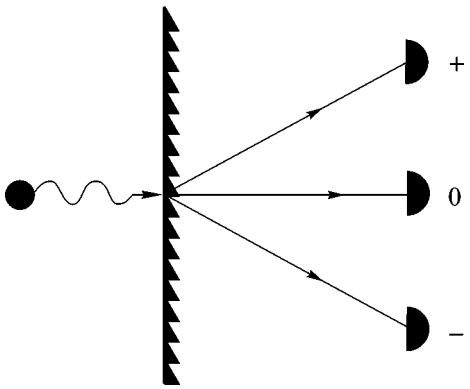


Fig. 1. Schematic of the detection of photoemission into the components of resonance fluorescence triplet. Spectral filter is drawn as a diffraction grating.

We are interested in the state $\hat{\rho}(\alpha, U)$, in which the atom occurs immediately after the event that was detected by the detector α for unravelling U:

$$\hat{\rho}(\alpha, U) = \frac{\hat{L}_\alpha(U)\hat{\rho}^{(st)}\hat{L}_\alpha^\dagger(U)}{\mathrm{Tr}(\hat{L}_\alpha^\dagger(U)\hat{L}_\alpha(U)\hat{\rho}^{(st)})}. \quad (11)$$

The density matrix before detection is assumed to be stationary, i.e., diagonal in the basis of dressed states and having elements $\rho_{11}^{(st)} = \sin^4\theta/(\cos^4\theta + \sin^4\theta)$ and $\rho_{22}^{(st)} = \cos^4\theta/(\cos^4\theta + \sin^4\theta)$. This choice corresponds to averaging over all possible atomic prehistories. Assume, for definiteness, that $\cos^2\theta \geq \sin^2\theta$. Note that, after the event was detected by the detector $\alpha = \{-\}$ under the conditions of trivial unravelling $U = id$ (i.e., after the absorption of a photon belonging to the left triplet component), the atom finds itself in the $\hat{\rho}(\{-\}, id) = |1\rangle\langle 1|$ state. The probability of finding the atom in the excited state, according to Eq. (1), is

$$\langle e|\hat{\rho}(\{-\}, id)|e\rangle = \cos^2\theta. \quad (12)$$

If our goal is to obtain the highest possible value of ρ_{ee} for a given θ, then Eq. (12) is limiting for the trivial unravelling. However, it turns out that the organization of a nontrivial unravelling allows the atom to be found, after the detection of a certain event, in a state that is more inverted than in Eq. (12). We now demonstrate how it happens.

Let us consider the unravelling of the form (Fig. 2)

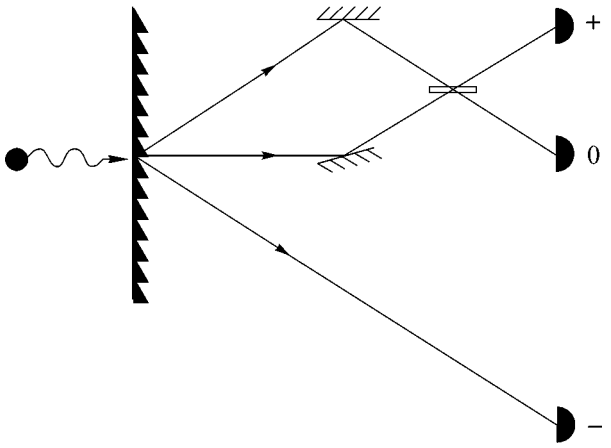


Fig. 2. Unravelling with the superposed emissions into the right and central components of the resonance fluorescence triplet.

$$\begin{aligned}\hat{L}_+(U) \equiv \hat{L}_+(\phi) &= \cos\phi\hat{L}_+ + \sin\phi\hat{L}_0,\\ \hat{L}_0(U) \equiv \hat{L}_0(\phi) &= -\sin\phi\hat{L}_+ + \cos\phi\hat{L}_0,\\ \hat{L}_-(U) \equiv \hat{L}_-(\phi) &= \hat{L}_-.\end{aligned} \tag{13}$$

After the event was detected by the detector $\alpha = \{+\}$, the atom occurs in the $\hat{\varrho}(\{+\}, U) \equiv \hat{\varrho}(\{+\}, \phi)$ state, for which

$$\begin{aligned}&\langle e|\hat{\varrho}(\{+\}, \phi)|e\rangle\\ &= \frac{(1-2\sin 2\phi)\cos^2\theta\sin^4\theta + \sin^2\phi\cos^4\theta\sin^2\theta}{\cos^2\theta\sin^2\theta + \sin^2\phi(\cos^6\theta + \sin^6\theta)}.\end{aligned} \tag{14}$$

The right-hand side of Eq. (14) reaches maximum at $\phi = \phi_{max}$ that is determined by the equation

$$\begin{aligned}\tan\phi_{max} &= \frac{1}{2(2-\sin^2\theta)}((\cos 2\theta - 1)\cos 2\theta\\ &- \sqrt{(\cos 2\theta - 1)^2\cos^2 2\theta + 2\sin^2 2\theta(2-\sin^2 2\theta)}).\end{aligned} \tag{15}$$

For $\theta = \pm\pi/4$, where, according to Eq. (12), the inversion cannot be achieved for the trivial unravelling, we have $\tan\phi_{max} = -1/\sqrt{2}$ and $\langle e|\hat{\varrho}(\{+\}, \phi_{max})|e\rangle = (2 + \sqrt{2})/4 \simeq 0.854$. With an increase in $\cos^2\theta$, the value of $\langle e|\hat{\varrho}(\{+\}, \phi_{max})|e\rangle$ decreases, and the gain, with respect to the value given by Eq. (12) for the trivial unravelling, vanishes at $\cos^2\theta \simeq 0.637$.

It is worth noting that, after the detection of events corresponding to the Lindblad operators $\hat{L}_+(\phi)$ and $\hat{L}_0(\phi)$ with $\phi \neq 0 \pmod{\pi}$, the initially stationary atomic density matrix becomes nondiagonal in the basis set of dressed states. This fact cardinally distinguishes unravelling (13) from the trivial unravelling or from the unravelling for which the photons emitted into the side triplet components mix at a single beam splitter. In both latter cases, the atom (initially in the stationary state) finds itself, after the detection of any event, in one of the dressed states with a certain probability.

Therefore, we considered the unravelling effect, i.e., the effect of a certain environmental rearrangement for a fluorescing atom. The effect consists in the creation of a state with the maximal value of ϱ_{ee}. As was already pointed out, the unravelling does not change the condition for spontaneous atomic decay (it is this reason for which such a rearrangement was called minimal at the beginning of this article). One can imagine that the atom "does not know" how an experimenter unravelled its evolution, i.e., how he disposed of the information carried by the spontaneously emitted photons. Due to the unravelling, an atom can be found in the desired state.

We considered the simplest case of unravelling with a single beam splitter. It is conceivable that the general unitary transformation (10), which realizes the superposition of all three original events, can be used to achieve higher values of ϱ_{ee}. Next, we considered the effect of only a single event, rather than the effect of longer series. When solving this problem, one must take into account the Hamiltonian's atomic evolution in the intervals between the events, because the off-diagonal elements appear in the density matrix in the dressed basis set for any series of events. Finally, unravelling localization is of particular interest for dynamic quantum systems.

REFERENCES

1. G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
2. C. Cohen-Tannoudji and S. Reynaud, J. Phys. B **10**, 345 (1977).
3. Ph. Blanchard and A. Jadczyk, Rep. Math. Phys. **36**, 235 (1995).
4. M.-D. Choi, Linear Algebr. Appl. **10**, 285 (1975).
5. J. K. Breslin, G. J. Milburn, and H. M. Wiseman, Phys. Rev. Lett. **74**, 4827 (1995); J. K. Breslin and G. J. Milburn, J. Mod. Opt. **44**, 2469 (1997).

Translated by V. Sakun