On the origin of cosmic γ-ray bursts

I. L. Rozental' and B. A. Trubnikov
Institute of Space Studies, Russian Academy of Sciences, 117810 Moscow, Russia; Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

(Submitted 4 March 1999; resubmitted 1 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 417–421 (10 October 1999)

As is well known, the distribution of the number of cosmic γ-ray bursts (CGBs) as a function of their duration has two maxima. In the present letter it is suggested that the longer-duration but less powerful CGBs arise when two neutron stars merge and a magnetic field is generated, as examined below, and the shorter-duration CGBs can arise as a result of collisions of a black hole either with a neutron star or another black hole. © *1999 American Institute of Physics.*
[S0021-3640(99)00119-X]

PACS numbers: 98.70.Rz, 97.60.Jd, 97.60.Lf, 97.10.Ld

1. The distribution of cosmic γ-ray bursts (CGBs) as a function of their duration T is shown in Fig. 2 of Ref. 1. Two (and possibly three) types of CGBs can be seen in the figure: short-duration with $T\sim 1$ s and long-duration with $T\sim 10-100$ s. They also differ by the dependences of the power on their frequency, and a definite correlation is observed between the average photon energy and the total pulse duration — more powerful bursts are characterized by a shorter duration. This seems to indicate the existence of several types of CGBs, and it is reasonable to infer that black holes (HH and HN type collisions) participate in the more powerful ones, while less powerful ones arise when two neutron stars merge (NN collisions). The idea that CGBs can arise as a result of NN and HN collisions was advanced, for example, in Refs. 2 and 3.

Table I shows the parameters of some CGBs[1)] ($z=(\lambda-\lambda_0)/\lambda_0$ is the red shift, D is the bolometric distance to the source,[2)] S is the spectrum or observed energy flux, E is the true energy of the explosion, $E_\gamma^{\max}$ is the maximum photon energy, and T is the duration).

HH and HN type collisions with the participation of black holes should be of shorter duration, since the gravitational forces involved in them are stronger. The maximum γ-ray energy $E_\gamma^{\max}$ in them seems to be higher than in NN collisions, although, for example, photons with $E_\gamma^{\max}=18$ GeV were observed in an unusually long-duration CGB on February 17, 1994 with $T=90$ min.

It is evident from Fig. 1 that the number of long-duration CGBs is larger than the number of short-duration bursts. This probably indicates that the number of neutron stars in the Universe is approximately 100 times larger than the number of black holes, which could consist of quarks and gluons. It is curious that high-energy γ-ray power-law spectra of the type $dN_\gamma/dE\sim E^{-\alpha}$ with exponent $\alpha\sim 2-3$, which are indicated in Table I, are close to a cosmic-ray spectrum with exponent $\alpha\approx 2.75$. Therefore it is possible that

TABLE I.

Burst	z	D [Mpc]	S [erg/m^2]	E [erg]	$E_\gamma^{\max}$ [GeV]	T[s]
GRB 910503	no	no	$E^{-2.2}$	no	10	84
GRB 910601	data	data	$E^{-3.7}$	data	0.314	200
GRB 930131			$E^{-2.0}$		1.2	100
GRB 970508	0.24	1217	0.35	5×10^{51}	–	–
GRB 971214	3.4	31164	0.032	3×10^{53}	–	–
GRB 980329	5	52519	0.075	2×10^{54}	–	–
GRB 980425	0.009	42	0.6	10^{49}	–	–
GRB 980703	1	6328	2.6	10^{54}	–	–
GRB 990123	1.61	11473	2.4	3×10^{54}	–	–

CGBs are also sources of cosmic rays. We note that three close NN pairs, which are gradually converging, losing energy to gravitational waves, have been discovered in our galaxy, and estimates show that they could give rise to a CGB close to us within 10–100 thousand years.

2. The hypothesis of magnetic field generation, proposed in Refs. 4 and 5, can explain the appearance of γ-ray bursts (GBs) in NN collisions. Here we shall examine the corresponding mechanisms in greater detail. In Ref. 4 it is assumed that a magnetic field B with energy density $w=B^2/8\pi$ ''flows'' with the velocity of light out of a sphere with surface area $S=4\pi R^2$. This gives a power

$$W=cwS=c(BR)^2/2=1.5\times10^{52}(B/10^{15}\,\mathrm{G})^2(R/10\,\mathrm{km})^2\mathrm{erg/s}. \quad (1)$$

For example, for $B\sim10^{15}$ G this mechanism will give in 20 s a total energy $E=3\times10^{53}$ erg, which, specifically, was released in a GB on December 14, 1997 (GRB 971214). In our opinion, however, a somewhat different mechanism operates here; the

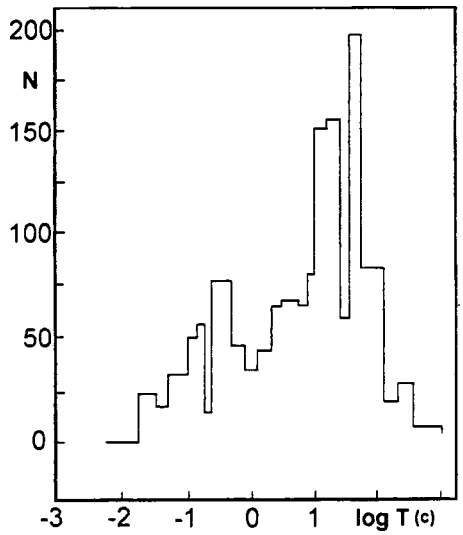

FIG. 1. Distribution of the number of GBs as a function of their duration T (sample of 400 GBs).

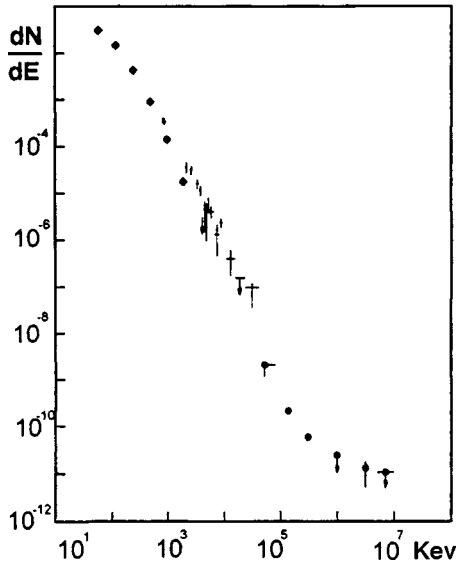

FIG. 2. Energy spectrum of GBs from Ref. 6 (dN/dE is the number of photons/keV$\cdot$s$\cdot$cm^2).

magnetic field does not simply flow ''in finished form'' out of a sphere but rather the field is generated continuously in the process of spattering of the plasma of the colliding rotating stars. We shall examine this process in greater detail.

The generation of a magnetic field is described by the magnetohydrodynamic equation $\partial\mathbf{B}/\partial t = \mathrm{curl}\,[\mathbf{v}\times\mathbf{B}]$. In what follows, we shall employ a spherical coordinate system (r,ϑ,φ), and for simplicity we shall assume that the magnetic poles lie on the rotation axis of a spherical star with radius R. We assume that the rotation of the star drags the surrounding plasma so that the φ component of the plasma velocity outside the star is $v_\varphi = r\Omega\sin\vartheta f(r)$, where $f(r)$ is a decreasing function of the radius, satisfying the condition $f=1$ on the surface of the sphere $r=R$. Rotation starts at time $t=0$, when the initial magnetic field ''frozen'' into the plasma was a purely dipole field with the components

$$B_r^0 = B_0(R/r)^3\cos\vartheta, \quad B_\vartheta^0 = (B_0/2)(R/r)^3\sin\vartheta, \quad B_\varphi^0 = 0. \tag{2}$$

where B_0 is the initial field at the poles. Solving the above-indicated generation equation, it is easily found that the two poloidal components $B_{r,\vartheta}^0$ of the magnetic field remains unchanged, and the φ component of the field grows linearly with time as

$$B_\varphi = B_0\Omega t\sin\vartheta\cos\vartheta F, \quad F = R^3 r^{-2} df/dr, \tag{3}$$

and even after several revolutions with period $\tau = 2\pi/\Omega$ it will be greater than the poloidal components. Neglecting the latter, we find that the magnetic energy density in the plasma surrounding the star grows as $w \approx B_\varphi^2/8\pi \sim t^2$, and the total magnetic energy in all space outside the star will grow as

$$E = \int w\,dV = \frac{\pi}{5} V_0 (B_0 N_{\mathrm{rev}})^2 C, \tag{4}$$

where $V_0=(4/3)\pi R^3$ is the volume of the star, $N_{\rm rev}=t/\tau$ is the number of revolutions of the star from the onset of rotation, and $C=R^3\int_a^\infty (df/rdr)^2 dr$ is a dimensionless factor. If the function f decreases as a power law $f=(R/r)^n$, then $C=n^2/(2n-3)$, where, for example, $C=3$ for $n=3$ and $C=4$ for $n=6$. The approximate formula $E \approx 2V_0(B_0N_{\rm rev})^2$ can be used.

At the moment the two neutrons stars converge, their orbital velocities can be of the order of the velocity of light. For example, for $v=c/2$ the period of relative rotation is $\tau\approx 2\times 10^{-4}$ s. The volume of a star with the radius $R=10$ km is $V_0=4\times 10^{18}$ cm^3, and with an initial field at the poles $B_0=10^{12}$ G magnetic energy equal to the rest energy of the Sun $E\approx M_\odot c^2=1.8\times 10^{54}$ erg will be produced in 5×10^5 revolutions, i.e., in 100 s, which corresponds in order of magnitude to ''long-duration'' GBs.

3. The mechanism leading to the appearance of γ rays from the magnetic field found above, in our opinion, could be as follows. The field B_φ is generated by the currents $j_{\rho,\vartheta}$, specifically, in our model $j_\vartheta=C(R/r)^{n+4}\sin 2\vartheta$, where $C=-n(n+2)cB_0\Omega t/8\pi R$. The component j_ϑ is maximum for $\vartheta=45°$, and a current $J=(n/4\sqrt{2})cRB_0\Omega t$ flows through the entire surface of this cone. For $n=5$, $R=10$ km, and $B_0=10^{12}$ G we have, numerically in amperes, $J=5.5\times 10^{19}N_{\rm rev}$ A; as a result of instability, this current can break up into individual current filaments which have relatively low currents J_1 and which are cut off as a result of growing constrictions. Numerous continuous cutoffs of the current filaments should give rise to fluxes of accelerated particles, as happens in laboratory plasma pinches. In Ref. 3 it is shown that the energy spectrum of ultrarelativistic particles, accelerated in breaks of the constrictions of a plasma pinch with current J_1, is described by the formula

$$dN_p/d\varepsilon=(C_0/\varepsilon)\exp[-(\varepsilon/\varepsilon_0)], \tag{5}$$

where C_0 and $\varepsilon_0=(e/c)J_1$ are constants.

The quantity C_0 is difficult to estimate, since it is determined by all of the many cutoffs of the current filaments together, but approximately the same spectrum should also arise for γ rays generated by particle collisions. Thus, Fig. 2 (which was taken from Ref. 6) shows the experimentally recorded spectrum of a GB, and it is very curious that the authors of that paper[6] themselves indicate that the formula

$$dN_\gamma/dE=0.02(100/E)^{1,2}\exp[-(E/E_0)][\text{photon}\cdot\text{keV}^{-1}\text{cm}^{-2}\text{s}^{-1}], \tag{6}$$

where the photon energy E is measured in keV and $E_0=900$ keV, gives a good fit in the photon energy range from 40 keV to 3 MeV. This spectrum is virtually identical to formula (5). This likewise attests to the ''magnetic generation'' mechanism that we proposed for NN collisions.

However, a different mechanism, which is examined below, leading to the formation of the energy spectrum of the γ rays of CGBs is also possible.

4. CGB spectrum in the ''fireball'' model. The above-described idea of a filament-like structure of the pinched currents and the numerous multidirectional thin particle beams formed by them makes it easier to understand the nature of the numerous microbursts comprising each GB. Similar sequences of micropulses are observed, as is well known, in each pulse of radio pulsars, whose radiation possesses a ''pencil-like'' (more accurately ''conical'') directional pattern. However, in our model the experimental value

of the parameter $E_0 = 900\,\mathrm{keV}$ in Eq. (6) should correspond to a current in the filaments $J_1 = (c/e)E_0 = 3\times 10^4\,\mathrm{A}$, which is too low, so that the correspondence with the model is only qualitative.

More likely, the experimental parameter $E_0 = 900\,\mathrm{keV}$ corresponds to the temperature of a thermodynamically equilibrium medium consisting of electron–positron pairs and photons with energy close to $m_e c^2 = 511\,\mathrm{keV}$. In Ref. 7 the hydrodynamics of the expansion of a fireball consisting of electron–positron pairs and photons was studied, and it was shown that the radiation spectrum should be described by the formula

$$\frac{dN_\gamma}{dE} = \frac{\mathrm{const}}{2\Gamma}\int_{E/2\Gamma}^{\infty} \frac{d\varepsilon}{\varepsilon[\exp(\varepsilon/kT)-1]} \approx \frac{\mathrm{const}}{E}\exp\left(-\frac{E}{2\Gamma kT}\right). \tag{7}$$

(The expansion of the fireball was studied even earlier in a number of works by other authors, but they did not present such a formula for the spectrum.) Here Γ is the Lorentz factor of the macroscopic motion of the gas, and this formula also is in qualitative agreement with the experimental spectrum (6), but in the model of strict isotropic expansion it is difficult to explain the microbursts comprising each individual CGB.

In our opinion, the hypothesis that the long- and short-duration CGBs are of different nature can explain some of their features and can yield information about the number of neutron stars and black holes in the universe.

Addendum added after refereeing

This paper was in the refereeing stage for quite a long time, and we are grateful to the referee for calling our attention to Refs. 8 and 9 and (which appeared after our paper was submitted for publication). We wish to make a number of remarks in this connection.

In the preprint cited as Ref. 8 the authors divide all GBs into not two but three groups according to their duration T: short-duration single-peak (short $T<0.1\,\mathrm{s}$), medium-duration (middle $0.1<T<1\,\mathrm{s}$), and long-duration (long $T>1\,\mathrm{s}$). They indicate that the $\log \mathrm{N}-\log \mathrm{S}$ plots for the medium- and long-duration GBs deviate from the 3/2 law, which seems to indicate that such GBs are cosmologically far away.

However, for short-duration GBs the plot is close to the 3/2 law, which could indicate that they are close by, and therefore such bursts are of comparatively low energy. In the opinion of the authors of Ref. 8, these GBs can be produced during the rapid evaporation of ''primary'' black holes of small mass, consisting of quark–gluon plasma (according to our hypothesis, large black holes of stellar mass consist of such a plasma). However, the number of short-duration GBs analyzed in Ref. 8 is very small (~ 20), so that the conclusion that they are local is only tentative. Moreover, most investigators assume that primary black holes could hardly survive to our epoch.

A number of subtle but important features of GBs are noted in Ref. 9. In our opinion they can be explained on the basis of our model with magnetic generation of GBs:

a) Estimates show that comparatively few ultrarelativistic bosons ($\sim 10^{-5}M$) are created in a typical GB, and our model satisfies this requirement, since the number of accelerated protons is small;

b) the energy of a GB is released not instantaneously, as a single burst (as assumed in the fireball theory), but rather gradually, and both of our hypotheses can explain this;

c) a typical GB consists of many microbursts, which in our model can be explained by the sequence of cutoffs of current ''micropinches'' in the expanding volume of the plasma cloud of a GB;

d) it is necessary to explain how the energy of the explosion is transferred to photons; our model, where ultrarelativistic particle beams are produced in current pinches that are being cut off, and the subsequent collisions of the particles with the main plasma give bremsstrahlung, can explain this.

In closing, we thank A. A. Starobinskiĭ for a discussion of this work.

[1)]The shift z was first determined, and D and E were determined from z, only in 1997 after the discovery of the optical afterglow of CGBs. For this reason, the parameters z,D, and E are unknown for the first three CGBs.
[2)]Here $D=(2cz/H\Omega)[1-(2-\Omega)(1+\sqrt{1+z\Omega})^{-1}$, $E=4\pi D^2 S/(1+z)$, $\Omega=0.3$, and $H=65$ km/s·Mpc.

[1] B. I. Luchkov, I. G. Mitrofanov, and I. L. Rozental', Usp. Fiz. Nauk **166**, 743 (1996).
[2] S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, and A. G. Polmarev, Pis'ma Astron. Zh. **10**, 422 (1984) [Sov. Astron. Lett. **10**(3), 177 (1984)].
[3] B. Paczynski, Acta Astron. **41**, 257 (1991).
[4] B. Paczynski, Astrophys. J. **494**, No. 1, part 2, L45 (1998).
[5] B. A. Trubnikov, Usp. Fiz. Nauk **167**, 345 (1997).
[6] K. Hurley, B. Dingus, R. Mukherjgee *et al.*, Nature (London) **372**, 652 (1994).
[7] I. Belousova, I. L. Rozental, Astrophys. Space Sci. **249**, 117 (1997).
[8] D. B. Cline, C. Matthey, S. Otvinovski *et al.*, http://xxx.lanl.gov/abs/astro-ph/9905346 (26 May 1999).
[9] K. A. Postnov, Usp. Fiz. Nauk **169**, 545 (1999).

Translated by M. E. Alferieff

Stochastic resonance in a nonequilibrium electron–hole plasma

V. V. Osipov and E. V. Ponizovskaya
Science Center of the Russian Federation, Scientific and Industrial Association ''ORION,'' 111123 Moscow, Russia

(Submitted 17 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 422–427 (10 October 1999)

A photogenerated electron–hole plasma, heated in the process of Auger recombination, is studied. It is shown that in the plasma near the threshold for the appearance of uniform relaxational self-excited oscillations, weak noise transforms into a stochastic sequence of large-amplitude spikes. An additional optical periodic signal with amplitude approximately five times smaller than the noise variance, depending on the form of this signal, transforms these stochastic oscillations into low-amplitude quasiharmonic oscillations or into periodic spike self-excited oscillations of enormous amplitude. © *1999 American Institute of Physics.* [S0021-3640(99)00219-4]

PACS numbers: 71.35.Ee, 73.50.Td

Stochastic resonance (SR) is a striking nonlinear effect, which appears, as a rule, in bistable systems and is associated with a noise-induced transition from one state of the system into another.[1,2] A unique property of SR is that the signal/noise ratio increases with increasing noise and reaches a maximum at a quite high noise level. However, virtually all basic results on SR have been obtained primarily by analyzing axiomatic chemical and biological models. Recently, we have also shown for certain axiomatic models that SR can occur near the threshold for the appearance of relaxational self-excited oscillations.[3,4] In this case, SR is associated with the noise-stimulated transitions from a state of low-amplitude quasiharmonic oscillations into a state of spike relaxational oscillations of large amplitude, which are identical in form to the oscillations shown in Figs. 1a and 1b, respectively. At the same time, relaxational self-excited oscillations arise in a variety of concrete physical systems, including in a nonequilibrium electron–hole plasma (EHP) heated in the process of Auger recombination.[5] In the present work we discovered and studied numerically SR in an EHP uniformly photogenerated in a CdHgTe thin film.

A nondegenerate uniform EHP with sufficiently quite high density can be treated as a single system described by two equations — one for the density n of hot electrons and holes ($p=n$) and one for their effective temperature T:[5–7]

$$\frac{\partial n}{\partial t}=\kappa\Phi-G(T)n^3-\frac{n}{\tau_r}, \tag{1}$$

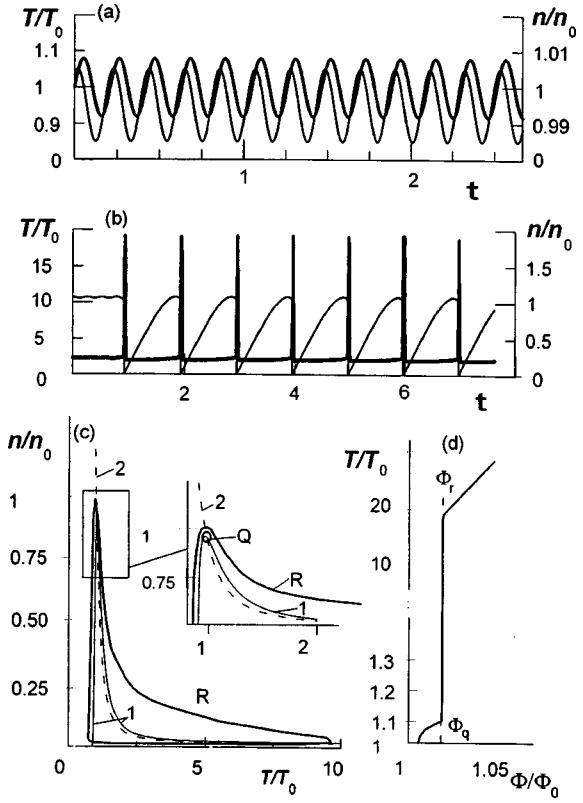

FIG. 1. Quasiharmonic and relaxational self-excited oscillations of an EHP: (a) and (b) — temporal variations in time of the temperature T (thick curve) and density n (thin curve), time is measured in units of τ_A; (c) — form of the limit cycles corresponding to these oscillations (the thick curves Q and R, respectively), and the form of the null isocline of Eqs. (8) and (9) (the thin solid curve *1* and the thin dashed curve *2*, respectively); (d) — bifurcation diagram — dependence of the amplitude of the temperature oscillations on the photogeneration rate Φ. The results of numerical investigations of Eqs. (8) and (9) for the parameters of CdHgTe presented in the text and $T_f = 30$ K, $\tau_A = 4.8\times10^{-9}$ s, $\Phi = 1.01\Phi_0$ (for a) and $\Phi = 1.05\Phi_0$ (for b).

$$\frac{3}{2}Kn\frac{\partial T}{\partial t} = E_g G(T) n^3 - P. \tag{2}$$

Here K is Boltzmann's constant, κ is the absorption coefficient, Φ is the photogeneration rate, E_g is the band gap, and τ_r is the Shockley–Read recombination time. To simplify, in Eq. (1) we neglected radiative recombination, which is not important for our purposes and is justified, for example, for CdHgTe.[8] In Eqs. (1) and (2) the terms proportional to n^3 describe Auger recombination, during which hot electrons arise. These hot electrons, colliding with other electrons and holes, give up to the EHP an energy of the order of E_g, thereby heating the plasma to a temperature T. We describe the power flowing from the hot electrons to the EHP by the first term on the right-hand side of Eq. (2). The second term is the power P flowing from the hot EHP to the lattice. The latter power can be represented as[7]

$$P = Kn(T - T_f)\nu_{\varepsilon a}\left(\frac{T}{T_f}\right)^{1/2} + n\nu_{\varepsilon 0}h\nu_0\left(\exp\left(-\frac{h\nu_0}{KT}\right) - \exp\left(-\frac{h\nu_0}{KT_f}\right)\right), \quad (3)$$

where T_f is the lattice temperature, $\nu_{\varepsilon a}$ and $\nu_{\varepsilon 0}$ are the characteristic relaxation frequencies of the energies of the hot carriers on acoustic and optical phonons, respectively, and ν_0 is the maximum frequency of longitudinal optical phonons.

In most semiconductors the Auger recombination rate $G(T)n^3$, more accurately, the coefficient $G(T)$, is an exponentially increasing function of the temperature of the electrons and holes. Specifically, for nondegenerate semiconductors with quadratic dispersion[9]

$$G(T) = \frac{\gamma}{n_{i0}^2}\left(\frac{KT}{E_g}\right)^{3/2}\exp\left(-\frac{E_g}{KT}\right), \quad \gamma = \frac{8(2\pi)^{5/2}q^4 m_e |F_1F_2|^2}{h^3(4\pi\varepsilon_0\varepsilon_r)^2 n_i^2}\left(\frac{E_g}{K\cdot 300\,\mathrm{K}}\right)^{-3/2}, \quad (4)$$

$$n_i^2 = n_{i0}\left(\frac{T_f}{300\,\mathrm{K}}\right)^3\exp\left(-\frac{E_g}{KT_f}\right), \quad n_{i0} = (2.5\times 10^{19}\,\mathrm{cm}^{-3})\left(\frac{m_n m_p}{m_0^2}\right)^{3/4}. \quad (5)$$

Here m_0 is the electron mass, h is Planck's constant, q is the electron charge, ε_r is the permittivity, and F_1 and F_2 are coefficients determined by the overlap integrals of the corresponding Bloch functions. We note that in order for expression (3) to be satisfied, it is sufficient in practice that only the heavy holes in CdHgTe be nondegenerate, and this holds for the EHP parameters considered here.

The spontaneous appearance of self-excited oscillations is due to positive feedback, which is due to the increasing function $G(T)$.[5,6] Indeed, as the temperature increases, the Auger recombination rate $G(T)n^3$ increases, which in turn results in a further increase of the temperature of the EHP. On account of the increasing function $G(T)$, the null isocline of Eq. (2), i.e., the function $n(T)$, which is determined by Eq. (2) with $\partial T/\partial t = 0$, is N-shaped (the thin line *1* in Fig. 1c). It is evident from Eq. (1) that this null isocline does not depend on the photoregeneration rate Φ. Conversely, the null isocline of Eq. (1) (the dashed curve *2* in Fig. 1c) is determined by Φ, and for $\Phi > \Phi_0$ it intersects the N-shaped null isocline (2) at a point on its decending section. The value $\Phi = \Phi_0$ determines the extremal point n_0 and T_0 of the N-shaped null isocline. According to Eqs. (1) and (2), the values $\Phi = \Phi_0$ and the quantities n_0 and T_0 for $T \ll E_g$ are determined by the equations

$$T_0 = T_f\left(1 + \frac{KT_f}{E_g}\right), \quad \frac{n_0}{n_{0i}} = 0.6\cdot\sqrt{\frac{\nu_{\varepsilon 0}}{\gamma}}\left(\frac{KT_f}{E_g}\right)^{7/4}, \quad (6)$$

$$\kappa\Phi_0 = 0.6 n_{0i}(\nu_{\varepsilon 0})^{3/2}(\gamma)^{-1/2}\left(\frac{KT_f}{E_g}\right)^{15/4}. \quad (7)$$

In deriving Eq. (7) we used the fact that the characteristic Auger recombination time is $\tau_A = \gamma^{-1}(n_{i0}/n_0)^2(KT_0/E_g)^{3/2} \ll \tau_r$ (see below).

It follows from Eqs. (1) and (2) that self-excited oscillations spontaneously arise for some $\Phi_c = \Phi_0(1 + \alpha KT_f E_g^{-1})$, where $\alpha = \tau_\varepsilon/\tau_A \ll 1$. It is convenient to rewrite Eqs, (1) and (2) in the dimensionless variables $\theta = T/T_0$, $\theta_f = T_f/T_0$, $\eta = n/n_0$, $\zeta = E_g/KT_0$, and $\varphi = \kappa\Phi\tau_A n_{0i}^{-1}$ as

$$\frac{\partial \eta}{\partial t} = \varphi - \frac{\theta^{3/2}}{\theta_f^3}\eta^3 \exp\left(-\zeta\left(\frac{1}{\theta} - \frac{1}{\theta_f}\right)\right) - \eta\frac{\tau_A}{\tau_r}, \tag{8}$$

$$\alpha\frac{\partial \theta}{\partial t} = \alpha\zeta\frac{\theta^{3/2}}{\theta_f^3}\eta^2 \exp\left(-\zeta\left(\frac{1}{\theta} - \frac{1}{\theta_f}\right)\right) - (\theta - \theta_f). \tag{9}$$

For numerical investigations, we used the parameters characteristic for a narrow-gap CdHgTe:[10] $E_g = 0.1$ eV, $\kappa = 10^{-4}\,\mathrm{cm}^{-1}$, $m_e = 0.01m_0$, $m_p = 0.5m_0$, $\varepsilon_r = 13.3$, $|F_1F_2| = 0.2$, $h\nu_0 = 25-30$ eV, $\nu_{\varepsilon a} = 10^{11}\,\mathrm{s}^{-1}$, $\nu_{\varepsilon 0} = 10^{13}\,\mathrm{s}^{-1}$, and $\tau_r = (1-5)\times 10^{-6}$ s. It follows from Eqs. (6) and (7) that for $T_f \simeq 30$ K at the extremal point $T_0 = 30.7$ K, $n_0 = 5.0\times 10^{14}\,\mathrm{cm}^{-3}$, and $\Phi_0 = 1.14\times 10^{19}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$, which corresponds to the radiation of a blackbody with $T \simeq 310$ K. For such EHP parameters the characteristic Auger recombination time is $\tau_A = 4.8\times 10^{-9}$ s, $\zeta = 21.9$, and $\alpha = 0.04$. Therefore, in the case at hand, $\alpha \ll 1$ and $\tau_A \simeq 10^{-8} - 10^{-9}\,\mathrm{s} \ll \tau_r$, i.e., Auger recombination dominates over other recombination channels. For $T_f \simeq 25$ K, we have $T_0 = 25.6$ K, $n_0 = 2.2\times 10^{14}\,\mathrm{cm}^{-3}$, and $\Phi_0 = 1.23\times 10^{18}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-2}$.

Numerical investigations of Eqs. (8) and (9) showed that at the threshold where the condition $\Phi > \Phi_c$ is satisfied, low-amplitude quasiharmonic oscillations arise spontaneously in the system (Fig. 1a). The limit cycle corresponding to these oscillations is represented by the curve *Q* in Fig. 1c. For very small (to the extent that the parameter α is small) excess of Φ above Φ_q such low-amplitude oscillations abruptly transform into relaxation oscillations (Fig. 1b), which for $\alpha \ll 1$ have the form of a large-amplitude spikes, proportional in magnitude to α^{-1}. The limit cycle corresponding to these oscillations is represented in Fig. 1c by the thick curve *R*.

The critical value Φ_q, for which the low-amplitude quasiharmonic oscillations abruptly transform into large-amplitude self-excited relaxation oscillations, can be found by the Bogolyubov–Mitropol'skiĭ method.[11] Routine mathematics, similar to that presented in Refs. 2 and 3, gives $\Phi_q = \Phi_c(1 + 0.1\alpha)$ to an accuracy of α^2. At the same time, the parameter $\Phi = \Phi_r$, for which the reverse transition from relaxational self-excited to quasiharmonic oscillations occurs, can be calculated using the general expression obtained in Ref. 12. As the calculations showed, Φ_r is equal to Φ_q to an accuracy of α^2. This means that the bifurcation characteristic, i.e., the dependence of the amplitude of the oscillations on the control parameter Φ, has the form shown in Fig. 1d.

This result agrees with our numerical calculations and signifies that the low-amplitude quasiharmonic oscillations and large-amplitude relaxation oscillations for $\alpha \ll 1$ transform abruptly into one another in a very narrow range near the point $\Phi = \Phi_q = \Phi_r$ (Fig. 1d). Hence it follows that if Φ is chosen to be close to Φ_q, then small perturbations of the photogeneration rate Φ (Fig. 2a) can transform into explosive noise in the form of stochastic spike-shaped self-excited oscillations (Fig. 2b), associated with random noise-induced transitions between low-amplitude and self-excited relaxation oscillations. To take account of the small fluctuations of the photogeneration rate, in the numerical calculations Φ was expressed as $\Phi = \Phi_q + \delta\Phi$, where $\delta\Phi$ is Gaussian noise with variance σ with respect to both amplitude and time. It was found that such stochastic self-excited oscillations arose in the EHP only if the variance of the fluctuations Φ was in the range $8\times 10^{-5} < \sigma/\Phi_0 < 5\times 10^{-4}$. When the variance of the fluctuations was chosen to be less than $8\times 10^{-5}\Phi_0$, quasiharmonic oscillations were observed in the

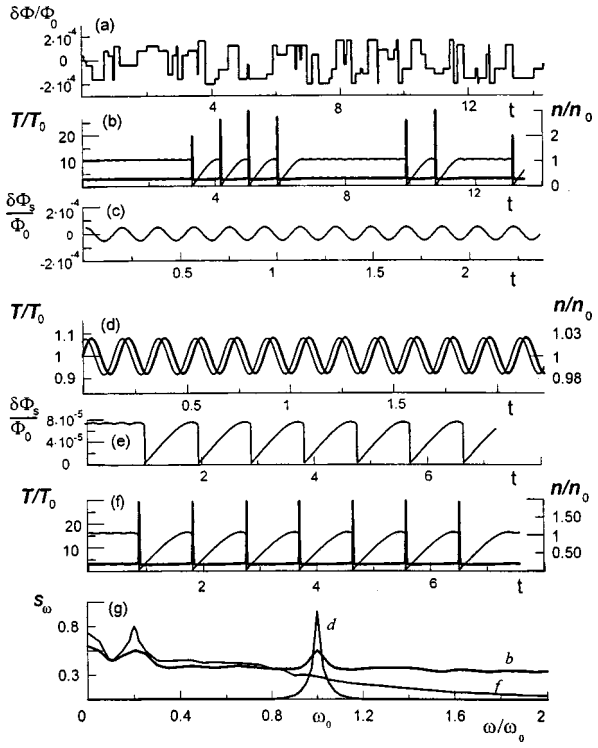

FIG. 2. Self-excited oscillations of an EHP in the presence of noise and a signal. The form of the small random variations $\delta\Phi$ of the photogeneration rate (a) and the corresponding form of the self-excited oscillations (b), time is measured in units of τ_A; (c) and (e) — form of the signal oscillations $\delta\Phi_S$, and (d) and (f) — the corresponding self-excited oscillations of temperature and density of the EHP; (g) — spectral density of the self-excited oscillations: the curves *b*, *d*, and *f* in Fig. 2g correspond to self-excited oscillations shown in Figs. 2b, 2d, and 2f, respectively. The results of numerical investigations of Eqs. (8) and (9) for the parameters of CdHgTe presented in the text and $T_f = 30$ K, $\tau_A = 4.8\times10^{-9}$ s, $\Phi = 1.01\Phi_0$ (for a) and $\Phi = 1.05\Phi_0$ (for b).

system (Fig. 1a), and when σ was greater than $5\times10^{-4}\Phi_0$, the system transformed into a regime of virtually periodic relaxational self-excited oscillations, identical in form to those shown in Fig. 1b.

Numerical investigations showed that the stochastic oscillations (Fig. 2b) formed as a result of weak fluctuations of the photogeneration rate Φ are very sensitive to weak signals. To analyze this effect, an additional weak optical harmonic signal $\delta\Phi_s = a\cdot\cos\omega t$ (Fig. 2c) was added on the right-hand side of Eq. (1) to the parameter Φ, together with $\delta\Phi$. Such a signal with frequency ω_0 approximately equal to the frequency ω_q of quasiharmonic oscillations caused the stochastic oscillations (Fig. 2b) to be transformed into self-excited oscillations of the form shown in Fig. 2d. The spectrum corresponding to them has the form of the curve *d* in Fig. 2g. It is significant that this effect was observed even for signal amplitudes approximately one-fifth of the noise variance of the parameter Φ. When the frequency of the external optical signal differed substantially but by less than a factor of 2 from ω_q, the quasiharmonic temperature and concentration oscillations were somewhat distorted, and when ω_0 differed from ω_q severalfold the

quasiharmonic oscillations started to alternate with infrequent spikes.

When the weak signal had the form of periodic relaxation oscillations of low amplitude (Fig. 2e), the stochastic spikes of self-excited oscillations transformed into periodic self-excited relaxation oscillations of large amplitude. The corresponding spectrum of oscillations is shown by the curve *f* in Fig. 2g. We note that such a gigantic amplification of a weak periodic signal was also observed even when the amplitude of the signal was approximately one-fifth of the noise variance.

One of the significant properties of SR is the dependence of the signal/noise ratio on the variance of the fluctuations of the control parameter. The signal/noise ratio is defined as the ratio of the spectral power of the oscillations to the spectral noise power at the signal frequency.[2] Calculations of the signal/noise ratio as a function of the variance of the fluctuations of Φ showed that for quasiharmonic signals and relaxational signals this ratio has a maximum, characteristic for SR,[1,2] approximately at the center of the range $8\times10^{-5}<\sigma/\Phi_0<5\times10^{-4}$, i.e., in the range of existence of stochastic self-excited oscillations. At the edges of this range the signal/noise ratio decreases, and for quasiharmonic signals it once again starts to grow as the variance decreases, while for relaxational oscillations, conversely, it grows with increasing noise variance.

In summary, we have shown that stochastic resonance occurs in a photogenerated electron–hole plasma for certain values of the plasma parameters, as a result of which a weak subthreshold signal whose amplitude is several times smaller than the noise variance can be observed. The gain of the photosignal can reach several orders of magnitude.

It is our pleasant duty to thank A. Yu. Selyakov for stimulating discussions.

[1] M. I. Dykman, D. G. Luchinsky, R. Mannella *et al.*, Nuovo Cimento D **17**, 661 (1995).
[2] V. S. Anishchenko, A. B. Neĭman, F. Moss, and L. Shimanskiĭ-Gaĭer, Usp. Fiz. Nauk **169**, 1 (1999).
[3] V. V. Osipov and E. V. Ponizovskaya, Radiotekh. Elektron. (Moscow) **186**, 2 (1998).
[4] V. V. Osipov and E. V. Ponizovskaya, Phys. Lett. A **369**, 369 (1998).
[5] N. N. Degtyarenko and V. F. Elesin, JETP Lett. **13**, 326 (1971); N. N. Degtyarenko, V. F. Elesin, and V. A. Furmanov, Fiz. Tekh. Poluprovodn. **7**, 1716 (1973) [Sov. Phys. Semicond. **7**, 1147 (1973)].
[6] B. S. Kerner and V. V. Osipov, Zh. Éksp. Teor. Fiz. **83**, 2201 (1982) [Sov. Phys. JETP **56**, 1275 (1982)].
[7] F. G. Bass and Yu. G. Gurevich, *Hot Electrons and Strong Electromagnetic Waves in Plasma in Semiconductors and Gas Discharges* [in Russian] (Nauka, Moscow, 1975).
[8] A. A. Grugova and V. V. Osipov, Fiz. Tekh. Poluprovodn. **15**, 2384 (1981) [Sov. Phys. Semicond. **15**, 1384 (1981)].
[9] A. R. Beattie and P. T. Landsberg, Proc. R. Soc. London, Ser. A **249**, 16 (1959); L. Gel'mont, Zh. Éksp. Teor. Fiz. **75**, 536 (1978) [Sov. Phys. JETP **48**, 268 (1978)].
[10] A. Rogalski, Infrared Phys. **28**(3), 139 (1988).
[11] N. N. Bogolyubov and Yu. A. Mitropol'skiĭ, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Greach, New York, 1962) [Russian original, Nauka, Moscow, 1983].
[12] S. M. Baer and T. Erneux, SIAM J. Appl. **46**, 721 (1986); S. M. Baer and T. Erneux, SIAM J. Appl. **52**, 1651 (1992).

Translated by M. E. Alferieff

Scenario of the electric breakdown and UV radiation spectra in single-bubble sonoluminescence

N. García
Laboratorio de Física de Sistemas Pequeños y Nanotecnología Consejo Superior de Investigaciones Cientificas, 28006 Madrid, Spain

A. P. Levanyuk
Departamento de Física de la Materia Condensada, Universidad Autónoma de Madrid, 28049 Madrid, Spain

V. V. Osipov
Laboratorio de Física de Sistemas Pequeños y Nanotecnología Consejo Superior de Investigaciones Cientificas, 28006 Madrid, Spain; Department of Theoretical Physics, ORION Russian Science Center, 111123 Moscow, Russia

(Submitted 5 April 1999; resubmitted 2 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 428–433 (10 October 1999)

In a preceding paper we put forward the hypothesis that the single-bubble sonoluminescence (SBSL) is caused by strong electric fields arising near the surface of a collapsing gas bubble on account of the flexoelectric effect in water. Here we argue that these fields can indeed provoke a multiple electric breakdown in water, in a micron-size region near the surface of the collapsing gas bubble, and show that the main numerical characteristics of the SBSL can be naturally explained within this mechanism. The SBSL spectra are determined by radiative transitions between high-energy levels of noble-gas atoms excited by hot electrons produced by the strong flexoelectric field in ''cold'' water.
 [S0021-3640(99)00319-9]

PACS numbers: 43.35.Hl, 78.60.Mq, 47.55.Dz

Sonoluminescence refers to the phenomenon of light emission during acoustic irradiation of a liquid and is associated with cavitation bubbles present in the liquid. The most controllable and promising experimental data have been obtained for the single-bubble sonoluminescence (SBSL): picosecond UV radiation from a single bubble pulsating in the field of a sound wave.[1] One of the specific features of SBSL is a decisive role of the noble gases dissolved in water in the phenomenon. Recently a theory has been developed[2] which explains the very existence of stable pulsations of a bubble in terms of diffusion of the dissolved gas and chemical reactions in the bubble gas.

In this paper we argue that the noble-gas atoms play a decisive role in the process of light emission in bright SBSL. We assume, as in Ref. 3, that the flexoelectric effect,[4] i.e., polarization of the water due to pressure gradients, is essential for the SBSL. According to estimates[3] the depolarizing field can reach 10^7 V/cm, i.e., it can exceed the breakdown field of water. However, the breakdown scenario has been far from clear. Indeed, the

mechanism of the breakdown in water involves ''lucky electrons'' whose acceleration in the electric field leads to the development of an avalanche (see, e.g., Ref. 5). At ambient temperature the concentration of free electrons in water is quite negligible, and it is hardly probable to find a ''lucky electron'' in a small volume near the bubble surface, i.e., in the region of high electric field, during the short time of existence of this field. Moreover, the origin of the optical radiation in the transparency window of water remained unexplained in Ref. 3. The reason for the spectacular synchronization of the emission pulses[1,6] also remained uncertain. The aim of the present paper is to answer these questions.

Let us describe briefly the proposed scenario of the breakdown and the radiation. At a certain short time interval when the bubble radius is near its minimum, the acceleration of the bubble surface and the pressure gradient have gigantic values (10^{14} cm/s^2 and 10^{14} dyne/cm^3).[1] Because of the flexoelectric effect the field at the bubble surface can reach a value of 10^7 V/cm.[3] It is important that in a certain time interval, $\tau_c \sim 1$ ns, when the bubble surface decelerates, the field is directed toward the center of the bubble. The strong electric field is localized in a thin spherical layer, of about 1 μm size, near the bubble surface and its value can exceed substantially the static breakdown field in water. However, breakdown starts only if ''lucky electrons'' appear in the thin layer during the deceleration period. These electrons indeed appear, because of a sharp increase in the temperature of the bubble gas and of a water layer very close to the bubble surface, up to at least several thousand kelvin, when the bubble radius is close to its minimum value.[7,8] The free electrons diffuse into the thin water layer in which the strong electric field is localized and is directed toward the center of the bubble. Then they are accelerated to energies sufficient for the generation of additional free electrons (as a result of the electric breakdown of the water) and for excitation of the noble-gas atoms. The light emission comes mainly from intense optical transitions between high-energy excited states of the noble-gas atoms. Because of the spatial inhomogeneity of the strong electric field, the radiation spectrum is continuous and resembles the blackbody radiation spectrum for a temperature equal to the effective temperature of the hot electrons. The characteristic time of the electric breakdown is much less than τ_c, i.e., the polarization can continue to change after the first breakdown, and the electric field can once again reach the breakdown threshold. As a result, several breakdowns can take place during the time interval τ_c. The multiplicity of the breakdowns determines both the energy of the SBSL pulse and its duration.

To support the above scenario let us first discuss the sign of the depolarizing electric field E at the bubble surface. It depends on the sign of the flexoelectric coefficient f, defined by $E = f\nabla p = -f\rho a$ (Ref. 3), where p is the pressure, ρ is the mass density, $a = -d^2R/dt^2$, and R is the radius of the bubble. Water molecules are highly asymmetric, so the negative oxygen ion tends to be located in the region of higher pressure, while the positive hydrogen ions tend to be located in the region of lower pressure. In other words, the polarization vector is directed opposite to the gradient and the depolarizing field is parallel to the gradient, i.e., $f > 0$ for water.

The most essential features of dynamics of the pulsating bubble are presented, e.g., in Ref. 1. Within a short time interval (1–10 ns) when the bubble radius R is close to its minimum value $R_{\min}$, the bubble surface first accelerates and then sharply decelerates, i.e., the velocity $v = -dR/dt$ first reaches its maximum and then goes to zero at the point

$R=R_{\min}$. That means that at R somewhat exceeding $R_{\min}$ the acceleration $a=-d^2R/dt^2$ reverses its sign and reaches huge values. During the acceleration period the pressure gradient $\nabla p=-\rho a$ is directed from the center of the bubble toward its periphery, i.e., in the same direction as the field. During the deceleration period, $\tau_c\sim 1$ ns, the flexoelectric field that arises is directed toward the bubble center.

As we have noted, the field E can substantially exceed the static threshold field for breakdown of water, E_{th}, but the electric breakdown occurs only if ''lucky'' electrons which could provoke an avalanche are available. This situation resembles that of laser breakdown.[5] The presence of ''lucky'' electrons in a volume of 10^{-11}–10^{-12} cm^3 is practically negligible at ambient temperature within the time interval of interest, $\tau_c\sim 1$ ns. The source of the ''lucky'' electrons is, in our case, the bubble gas and a water layer very close to the bubble surface. These electrons are accelerated if the field is directed toward the bubble center. This situation, as we have shown above, is realized during the deceleration period, when the bubble radius goes to its minimum value $R_{\min}\simeq 1$ μm.[1] Just when the bubble radius is close to $R_{\min}$ the gas temperature sharply increases.[7,8] Even relatively moderate temperatures are sufficient to ensure breakdown. Indeed, if we assume, as in Ref. 9, that the bubble gas has a temperature of 7000 K, then a water layer of thickness equal to the thermal penetration length, $\delta_T\sim 0.1$ μm, has a temperature of about 3000 K. Water can be considered an amorphous semiconductor with a band gap $E_g\simeq 6.5$ eV[5,10] and an effective density of states $N^*\simeq 10^{21}$ for $T>3000$ K. The equilibrium concentration of conduction electrons at $T\geqslant 3000$ K is $n=N^*\exp(-E_g/2kT)>3\times 10^{15}$ cm^{-3} and is established in the thermal ionization time $\tau_T=(N^*\sigma v_T)^{-1}\exp(E_g/2kT)<10$ ns, where $\sigma\simeq 10^{-15}$ cm^{-2} is the cross section for free-carrier recombination, and v_T is the thermal velocity of conduction electron in water. Therefore the number of conduction electrons N thermally generated within the time interval of interest, $\tau_c\sim 1$ ns, and in a surface layer of volume 10^{-11}–10^{-12} cm^3 exceeds 30–300. Note that a similar number of free electrons can be generated by thermal ionization of the bubble gas. Indeed, the gas temperature (7000 K) is about two times larger and the ionization energy of both the water vapor and Ar, Kr, Xe is about 12–16 eV, i.e., also about twice as large as E_g. Thus the SBSL flash is emitted at a well-defined moment when, during the existence of high deceleration and a strong electric field in cold water, the gas bubble becomes very hot. This can explain the cause of the synchronization of SBSL observed experimentally.[1,6]

Let us now elucidate the decisive role of the noble gases in the bright SBSL. In the typical situation the mean electron energy is of the order of several eV (see below). There is some probability that an electron has an energy exceeding 10 eV, which is about the lowest excitation energy of noble-gas atoms, and it excites these atoms. An important feature of the noble-gas atoms is that their metastable states have lifetimes of up to milliseconds (see, e.g., Ref. 11). So, once excited, the noble-gas atom can remain in the metastable state during the entire time interval of positive acceleration and multiple breakdowns, $\tau_c\sim 1$ ns, and, possibly, during many periods of the acoustic wave (~ 30 μs).[1] The hot electrons collide with these noble-gas atoms in the metastable state and transfer them to higher excited states. The radiative transitions between high-energy excited states produce radiation in the transparency window of water ($h\nu<E_g\simeq 6.5$ eV) and govern the SBSL spectrum.

The energy distribution of the hot electrons with energy $\varepsilon<E_g\simeq 6.5$ eV is given by

the Maxwell distribution. Indeed, in the strong electric fields an electron acquires, on average, an energy $qE\lambda \gg \varepsilon_{ph}$, where λ is the electron energy relaxation length, which for water is about 20–100 Å, and ε_{ph} is the characteristic energy of local oscillations in water, which is practically equal to the energy of optical phonons in ice, $\varepsilon_{ph} \simeq 80$–$100$ meV.[5] Since in the process of acceleration an electron radiates many optical phonons, the electron energy distribution is nearly isotropic in momentum space and is approximated by the Maxwell function with an effective electron temperature $kT_e = (eE\lambda)^2/3\varepsilon_{ph}$ (Ref. 12). Under these conditions the effective thermal velocity of electrons $v_T = \sqrt{3kT_e/m_e}$ is much higher than the drift velocity $v_d = \sqrt{\varepsilon_{ph}/m_e}$. For $E \simeq (2-10)\times 10^6$ V/cm the effective electron temperature $T_e \simeq (2-5)\times 10^4$ K.

The high-energy excited states of the noble-gas atoms are split by the Stark effect in the strong electric field. In addition, the electric field in the active region goes through several variations, at least (see below). The inhomogeneous broadening of all the lines is very strong, and one can consider the density of the high-energy excited states as a constant. At every collision the hot electron transfers the metastable noble-gas atom to a state with energy ε, the reference point of energy being the energy of the metastable state (for our estimations we consider only one metastable state). The concentration of atoms excited during the breakdown time τ_b to energies within an interval $d\varepsilon$ reads

$$dn_n^{**} = n_n^* (\sigma_{ex} v_T n)\, \tau_b \exp\left(-\frac{\varepsilon}{kT_e}\right) \cdot \frac{d\varepsilon}{kT_e}, \tag{1}$$

where n_n^* is the concentration of the noble-gas atoms in the metastable state, and σ_{ex} is the cross section for impact excitation of an atom from the metastable state to a state with energy ε. Atoms excited to states with the energy ε go to the ground state preferentially through intermediate excited states, radiating mainly phonons with energy $h\nu \simeq \varepsilon$ (Ref. 11). Thus the spectral density of energy per unit volume radiated in a single breakdown can be written as

$$\tilde{P}(h\nu)\, d(h\nu) = h\nu w_r n_n^* (\sigma_{ex} v_T n)\, \tau_b \exp\left(-\frac{h\nu}{kT_e}\right) \cdot \frac{d(h\nu)}{kT_e}, \tag{2}$$

where w_r is the probability of the spontaneous radiative transition. Making use of the fact that $w_r = (4(2\pi)^4 \nu^3/3c^3 h) D^2$, where D is the absolute value of the matrix element of the dipole moment of the transition,[13] we find

$$\tilde{P}(h\nu)\, d(h\nu) = \frac{4(2\pi)^4 D^2}{3c^3} n_n^* (\sigma_{ex} v_T n)\, \tau_b \nu^4 \exp\left(-\frac{h\nu}{kT_e}\right) \cdot \frac{d(h\nu)}{kT_e}. \tag{3}$$

Note that T_e in Eq. (3) is a function of the position, since $T_e \sim E^2$ (see above) and $E = E(r)$. To find $E(r)$ approximately one can consider water as an incompressible liquid in which $\nabla p \sim r^{-2}$. Since $E \propto \nabla p$, one has $E \simeq E_s R_s^2/r^2$, where E_s is the field at the bubble surface. Integrating Eq. (3), we find approximately the spectral density of radiation of a single breakdown:

$$P = \int \tilde{P}(h\nu, r)\, dV = \frac{4(2\pi)^4 D^2}{3c^3 h} n_n^* (\sigma_{ex} v_T N_t)\, \tau_b \nu^3 \exp\left(-\frac{h\nu}{kT_{es}}\right), \tag{4}$$

where $kT_{es}=(eE_s\lambda)^2/3\varepsilon_{ph}$, and N_t is essentially the total number of electrons participating in the breakdown. This number can be found by taking into account that as a result of the breakdown the depolarizing flexoelectric field becomes screened. The total charge transported in the screening process is

$$Q_t=PS=4\pi R_s^2\varepsilon_0 E_s=eN_t, \tag{5}$$

where we have used the fact that in our case $D=P+\varepsilon_0 E=0$, i.e., $P=-\varepsilon_0 E$.

The observed spectra are cut off in the short-wavelength region (at $h\nu\simeq E_g\simeq 6.5$ eV) because of absorption by the water. Taking this into account, we find the energy radiated in a single breakdown:

$$U_r=\int Pw_r^{-1}d(h\nu)\simeq h\overline{\nu}(\sigma_{ex}v_T N_t)\tau_b n_n^*\geqslant h\overline{\nu}\,\frac{v_T}{v_d}(\sigma_{ex}R_s n_n^*)N_t, \tag{6}$$

where $h\overline{\nu}$ is the characteristic photon energy, close to the energy of the maximum of the observed spectrum, $h\overline{\nu}\sim 5$–6 eV in the case of a strong electric field. In Eq. (6) we made use of the fact that the breakdown time $\tau_b\simeq R_s/v_d$ (see below). For the electric fields considered, $E_s\simeq 3\times(10^6$–$10^7)$ V/cm, we find from Eq. (5) that $N_t\simeq(2$–$10)\times 10^5$. Using $v_T/v_d=eE\lambda/\varepsilon_{ph}=20$–$100$, $\sigma_{ex}\simeq 10^{-15}\,\mathrm{cm}^2$, and $R_s\simeq 1$ μm, and putting $n_n^*\simeq 10^{17}$–$10^{18}\,\mathrm{cm}^{-3}$, which is one order of magnitude less than the concentration of the noble-gas atoms near the bubble,[2] and assuming that there are about 10 breakdowns during the time interval τ_c (see below), we find from Eq. (6) that the total photon number is 10^6–10^8 and the total energy $W_r=10^{-13}$–10^{-11} J. The observed energies of the bright SBSL pulses[1,14,15] are approximately equal to the energy of 10^6–10^7 photons with $h\omega\simeq 5$ eV.

An estimate of the upper limit of the radiation energy due to a single breakdown can also be made in another way. The energy of the electric field in the medium (water) is

$$U_t=\frac{\epsilon\epsilon_0}{2}\int_{R_s}^{\infty}E^2dV. \tag{7}$$

We emphasize that not all of this electric energy but only a certain part of it can in principle be spent on generation of light. Indeed, the light arises because of excitation of the noble-gas atoms by electrons accelerated by the flexoelectric depolarizing field. In other words, the radiation is a part of the Joule power, which is

$$W_J=\int_{R_s}^{\infty}jE\,dV=\int_{R_s}^{\infty}E\left(\frac{d}{dt}P\right)dV=\frac{\epsilon_0}{2}\,\frac{d}{dt}\left(\int_{R_s}^{\infty}E^2dV\right)\equiv\frac{d}{dt}U_e, \tag{8}$$

where $j=dP/dt$ is the current due to the screening of the depolarizing field, and we take into account that in our case $P=-\varepsilon_0 E$. From Eq. (8) follows that

$$U_e=\frac{\epsilon_0}{2}\int_{R_s}^{\infty}E^2dV=\frac{1}{2}\int_{R_s}^{\infty}EP\,dV. \tag{9}$$

One can see that U_e is only the energy of the electric field created by the dipole moments. The difference of the energies U_t-U_e is the free energy of ordering of the dipole moments. So only a part of the total electric energy U_t can be transformed into light.

Equation (9) can be rewritten approximately as

$$U_e = \frac{\epsilon_0}{2} \int_{R_s}^{\infty} E^2 dV \simeq 2\pi R_s^3 \varepsilon_0 E_s^2 \simeq 2\pi R_s^3 \varepsilon_0 (f \nabla p)^2. \quad (10)$$

Here for estimation we have used, as above, the relation $E \sim r^{-2}$. To estimate the maximum of U_e we put $E_s \simeq 10^7$ V/cm and $R_s \simeq 1$ μm and find that $U_e \simeq (10^{-11} - 10^{-10})$ J. The total Joule energy per SBSL pulse is higher because of multiplicity of the breakdown. This total Joule energy is much larger than the observed energy of the SBSL pulse.[1,14,15]

We emphasize that the spectrum given by Eqs. (4) resembles the blackbody one but the role of the temperature is played here by the effective temperature of the hot electrons T_e in water, which can be much higher than temperature of the bubble gas. The experimentally observed spectra can be fitted in the wavelength interval 200–700 nm to the blackbody spectra with temperatures $(2-5)\times 10^4$ K.[14] As we have mentioned above, such T_e are reached at electric fields $E_s \simeq (2-10)\times 10^6$ V/cm.

The pulse duration is limited by the breakdown time. Its value is $\tau_b = R_s / v_d$, where v_d is the drift velocity of the electrons, which in the strong electric fields saturates at a value $v_d = \sqrt{\varepsilon_{ph}/m_e} \simeq 10^7$ cm/s (see above). Assuming $R_s \sim 1$ μm, we find $\tau_b \sim 10$ ps. The whole deceleration period is about two orders of magnitude larger than $\tau_c \sim 1$ ns. Therefore, after the breakdown is finished and the depolarization field is screened, the polarization continues to change on account of the change in acceleration, and the field arises once again. As a result, a new breakdown may take place. This situation can be repeated several times. In effect, it manifests itself in an increase of the pulse duration. Therefore, within our scenario the larger the pulse duration, the greater is the energy of the pulse. Such a correlation is observed in experiment.[1,16,17] Also note that, within our scenario of SBSL, the pulse width does not depend on the spectral range of the radiation. This agrees with experiment as well.[16]

We emphasize also that our mechanism of the SBSL operates in water near the bubble surface, without any assumptions of extraordinary conditions, and not in the bubble gas, as is assumed in the most popular models of SBSL.[1] Our scenario provides an answer to the question:[1] ''Why is water the friendliest liquid for the SBSL?''

One sees from Eq. (6) that the radiation energy is proportional to n_n^*. The smaller the excitation energy to the metastable level, the larger is n_n^*. This explains the increasing influence on the SBSL in the series He–Xe.[1] Strong magnetic fields are known to hamper the heating of the electrons.[18] This might be the reason for the decreasing of the SBSL intensity with magnetic field observed experimentally.[19]

We have neglected the conductivity of the water. This is justified because, as a rule, the dielectric relaxation time τ_D is much greater than $\tau_c \simeq 1$ ns, which is the longest characteristic time of our problem. By increasing the ionic conductivity of the water by adding NaCl, for example, one can, according to our estimates, decrease τ_D down to 0.1 ns. In such an electrolyte the mechanism of SBSL discussed in this paper might be less effective. Observation of this effect would support the proposed mechanism of bright SBSL.

N.G. and V.V.O. thank EU ESPRIT, the Spanish CSIC, and NATO for Linkage Grant Ref. OUTRLG 970308. Also, APL thanks NATO for Linkage Grant HTECHLG 971213.

[1] B. P. Barber, R. A. Hiller, R. Löfstedt *et al.*, Phys. Rep. **281**, 65 (1997), and references therein.
[2] D. Lohse and S. Hilgenfeldt, J. Chem. Phys. **107**, 6986 (1997).
[3] N. García and A. P. Levanyuk, JETP Lett. **64**, 907 (1996).
[4] A. K. Tagantsev, Usp. Fiz. Nauk **152**, 423 (1987) [Sov. Phys. Usp. **30**, 588 (1987)].
[5] C. A. Sacchi, J. Opt. Soc. Am. B **8**, 337 (1991) and references therein.
[6] B. P. Barber and S. J. Putterman, Nature **352**, 318 (1991); L. A. Crum, Phys. Today No. 9, pp. 22–29 (September 1994).
[7] K. R. Weninger, B. P. Barber, and S. J. Putterman, Phys. Rev. Lett. **78**, 1799 (1997); B. P. Barber and S. J. Putterman, Phys. Rev. Lett. **69**, 3839 (1992).
[8] A. Madrazo, N. García, and M. Nieto-Vesperinas, Phys. Rev. Lett. **80**, 4590 (1998).
[9] R. Löfstedt, B. P. Barber, and S. J. Putterman, Phys. Fluids A **5**, 2911 (1993); E. B. Flint and K. B. Suslick, Science **253**, 1397 (1991).
[10] F. Williams, S. P. Varma, and S. Hillenius, J. Chem. Phys. **64**, 1549 (1976).
[11] M. J. Beesley, *Lasers and Their Applications*, Taylor and Francis, London, 1972.
[12] P. A. Wolf, Phys. Rev. **95**, 1415 (1954); G. A. Baraff, Phys. Rev. **128**, 2507 (1962); L. V. Keldysh, Zh. Éksp. Teor. Fiz. **48**, 1692 (1965) [Sov. Phys. JETP **21**, 1135 (1965).
[13] H. A. Bethe, *Intermediate Quantum Mechanics*, Benjamin, New York–Amsterdam (1964).
[14] R. Hiller, S. J. Putterman, and B. P. Barber, Phys. Rev. Lett. **69**, 1182 (1992).
[15] B. P. Barber, C. C. Wu, R. Lofstedt *et al.*, Phys. Rev. Lett. **72**, 1380 (1994).
[16] B. Gompf, R. Günther, G. Nick *et al.*, Phys. Rev. Lett. **79**, 1405 (1997).
[17] R. Hiller, S. J. Putterman, and K. R. Weninger, Phys. Rev. Lett. **80**, 1090 (1998).
[18] F. G. Bass and Yu. G. Gurevich, *Hot Electrons and Strong Electromagnetic Waves in Semiconductors and Gas Discharge Plasma* [in Russian], Nauka, Moscow (1975).
[19] J. B. Young, T. Schmiedel, and Woowon Kang, Phys. Rev. Lett. **77**, 4816 (1996).

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

On the dominant role of two-photon relaxation in photonic crystals in external fields

A. M. Basharov*)
Moscow Engineering-Physics Institute, 115409 Moscow, Russia

(Submitted 16 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 434–438 (10 October 1999)

It is shown that resonant interaction of a coherent wave with impurity atoms leads to filling of levels of an impurity atom that lie in the gap in the photon density of states and do not belong to resonant transitions, while the interaction of impurity atoms with a nonresonant coherent wave results in effective deactivation of the indicated levels. The main mechanism determining the pumping and decay of an impurity level in a gap are two-photon radiative relaxation processes previously investigated by the present author (Zh. Éksp. Teor. Fiz. **102**, 1126 (1992) [Sov. Phys. JETP **75**, 611 (1992)]). © *1999 American Institute of Physics.* [S0021-3640(99)00419-3]

PACS numbers: 42.70.Qs, 78.40.−q

The realization of photonic crystals with gaps in the photon density of states that lie in the ranges from microwave to visible and have widths up to 30% of the center frequency of the gap[1–3] opens up unique possibilities for controlling fundamental electromagnetic processes in a medium. Of enormous interest to investigators is the situation where an excited level of an impurity atom, coupled with the main electric dipole transition, lies in the gap in the photon density of states. Depending on the position of the level in the gap with respect to the edge, non-Markovian relaxation,[4–6] the role of resonant dipole–dipole interaction[7,8] quadrupole relaxation,[7] and two-atom relaxation[9] are discussed.

If the atomic level lies sufficiently deep in a wide gap in the photon density of states, then its relaxation is strongly suppressed, and such a level is of special interest for use as an element in various kinds of optical memory devices,[10] computational schemes,[11] and so on. In such problems coherent waves are used as the controlling waves. The interaction of coherent fields with such impurity atoms also exhibits important physical effects, for example, optical bistability.[12] The concept of a localized photon[13] and its role in radiative problems[14,15] is also being widely discussed, though localized photons, gap solitons, and light-induced transparency are studied mainly in the context of the interaction of electromagnetic waves with a periodic lattice of resonant atoms. However, in all such problems the interaction of coherent fields with impurity atoms is studied on the basis of taking account of and/or determining more accurately the same relaxation mechanisms that occurred in the absence of coherent fields.

In the present letter attention is drawn to the fact that in photonic crystals in coherent

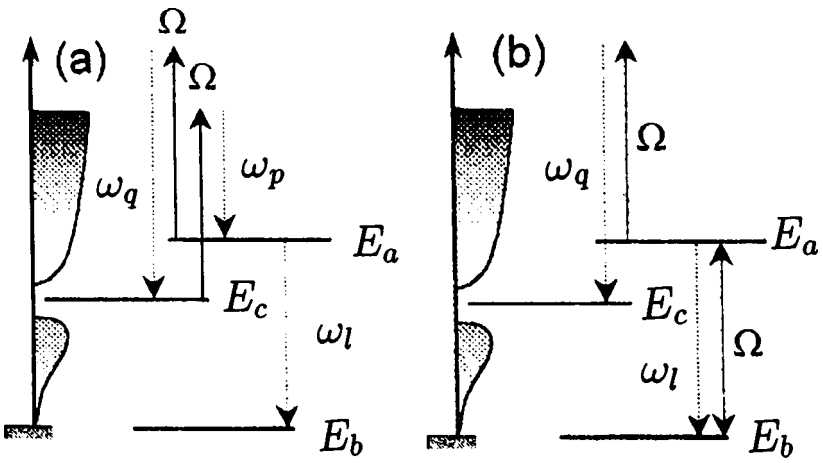

FIG. 1. Arrangement of the levels of an impurity with respect to the gap in the photon density of states and the frequencies of the fields participating in radiation processes.

fields (both resonant and nonresonant) under conditions where some of the conventional relaxation channels of impurity atoms are suppressed, new relaxation channels associated with two-photon relaxation processes,[16] which do not appear in ordinary media because they are weak compared with the spontaneous radiative decay, appear at the forefront. However, in photonic crystals they not only refine known results but they also completely change the picture of the interaction of coherent fields with impurity atoms in a wide range of values of the parameters. As a result, many suggestions for using impurity atoms in various devices and results of investigations of the interaction of coherent fields with impurity atoms must be re-examined.

In what follows, two very typical situations will be examined as examples illustrating the assertions made above. In these examples the energy level E_c lies deep inside a wide gap in the photon density of states and is coupled with the ground-state level E_b by an electric dipole transition. Other atomic energy levels coupled with the ground state by an electric dipole transition have energies greater than E_c. Let E_a be a typical such level. In the first situation (Fig. 1a) an impurity atom excited into the level E_c is in a field of a nonresonant coherent wave with frequency Ω. Then, in a time $\tau_{\text{two-photon}}$ much shorter than the lifetime of the level E_c in the absence of external fields, the impurity atom will pass into the ground state E_b for a wide range of frequencies $\Omega > (E_c - E_a)/\hbar$. In the second situation (Fig. 1b) the impurity atom initially in the ground state E_b interacts with a coherent wave, resonant with the transition $E_a \rightarrow E_b$. Then, in a time of the order of $\tau_{\text{two-photon}}$ the impurity atom will end up in the level E_c in the gap in the photon density of states, and a transparency induced by a resonant interaction with the coherent wave will occur at the transition $E_a \rightarrow E_b$. Evidently, this light-induced transparency is the opposite of the transparency caused by saturation of absorption at the transition $E_a \rightarrow E_b$, since this transition becomes ''idle'' because the impurity atom leaves the levels E_b and E_a and goes into the state E_c, which here is a kind of trap. It should be underscored that the resonant transition $E_a \rightarrow E_b$ can be any electric dipole transition whose frequency does not fall in the gap in the photon density of states and which

contains the ground state level. In addition, the level E_a can lie above or below E_c, lying in the gap in the photon density of states. The effects described are of a general character for the processes in which impurity atoms of photonic crystals in electromagnetic fields participate, and they must be taken into consideration everywhere.

The equations for the density matrix ρ of an impurity atom in the field of a nonresonant coherent wave with intensity

$$E=\mathcal{E}\exp[i(\mathbf{K}\cdot\mathbf{r}-\Omega t)+\text{c.c.} \tag{1}$$

can be written, taking account of the results of Ref. 16, as

$$\frac{d}{dt}\rho_{cc}=-(\gamma_c^{(2)}+\gamma_c|\mathcal{E}|^2)\rho_{cc}+\gamma_a|\mathcal{E}|^2\rho_{aa}, \quad \frac{d}{dt}\rho_{bb}=\gamma_c^{(2)}\rho_{cc}+\gamma_a^{(0)}\rho_{aa},$$

$$\frac{d}{dt}\rho_{aa}=-(\gamma_a^{(0)}+\gamma_a|\mathcal{E}|^2)\rho_{aa}+\gamma_c|\mathcal{E}|^2\rho_{cc}, \tag{2}$$

$$\gamma_c=\kappa_p|\Pi_{ca}(-\Omega)|^2/\hbar^2, \quad \gamma_a=\kappa_q|\Pi_{ac}(-\Omega)|^2/\hbar^2, \quad \gamma_a^{(0)}=\kappa_l|d_{ab}|^2/2\hbar^2,$$

$$\Pi_{ac}(\Omega)=\sum_\beta \frac{d_{a\beta}d_{\beta c}}{\hbar}\left(\frac{1}{\omega_{\beta a}+\Omega}+\frac{1}{\omega_{\beta c}-\Omega}\right), \quad \omega_{\alpha\beta}=(E_\alpha-E_\beta)/\hbar.$$

Here κ_p, κ_q, and κ_l are the coupling constants with the photonic thermostats with center frequencies $\omega_p=\Omega-\omega_{ac}$, $\omega_q=\Omega+\omega_{ac}$, and ω_{ab}; $d_{\alpha\beta}$ is the dipole moment of the transition $E_\alpha\to E_\beta$; $\gamma_c^{(2)}$ is the decay rate of a level inside the gap in the photon density of states as a result of two-atom relaxation.[9] We assume that the carrier frequency Ω of the wave (1) lies far from the spectral features of the photonic crystal.

Let the impurity atom initially be in the level E_c in the gap in the photonic density of states (Fig. 1a). It is evident that in the absence of a nonresonant coherent field (1) the lifetime of the atom in the level E_c is $1/\gamma_c^{(2)}$, while in the coherent field, for $\gamma_c^{(2)}\ll\gamma_{a,c}|\mathcal{E}|^2<\gamma_a^{(0)}$, it is determined by

$$\tau_{\text{two-photon}}\sim 1/\gamma_c|\mathcal{E}|^2\sim 1/\gamma_a|\mathcal{E}|^2\ll 1/\gamma_c^{(2)}.$$

The relaxation mechanism described occurs virtually for any frequency of the wave (1). One restriction is that there must not exist in the energy range $E_b<E_a<E_c+\hbar\Omega$ an impurity level E_a which is an excited level for some electric dipole transition that is unaffected by the spectral features of the photonic crystal.

When the two-photon relaxation[16] is taken into account, the equations for the density matrix of an impurity atom in the field of a coherent wave (1), resonant with respect to the transition $E_a\to E_b$ (Fig. 1b), at first glance not coupled with the level E_c in the gap in the photonic density of states, are

$$\frac{d}{dt}r_{ab}-i(\Delta-\Delta_{ab}|\mathcal{E}|^2)r_{ab}=\frac{i}{\hbar}(\rho_{bb}-\rho_{aa})d_{ab}\mathcal{E}-(\gamma_a^{(0)}/2+\gamma_{ab}|\mathcal{E}|^2)r_{ab},$$

$$\frac{d}{dt}\rho_{bb}=\gamma_a^{(0)}\rho_{aa}+\gamma_c^{(2)}\rho_{cc}+\frac{i}{\hbar}(r_{ab}^*d_{ab}\mathcal{E}-r_{ab}d_{ab}^*\mathcal{E}^*), \tag{3}$$

$$\frac{d}{dt}\rho_{aa} = -(\gamma_a^{(0)} + \gamma_a|\mathcal{E}|^2)\rho_{aa} - \frac{i}{\hbar}(r_{ab}^* d_{ab}\mathcal{E} - r_{ab} d_{ab}^* \mathcal{E}^*),$$

$$\frac{d}{dt}\rho_{cc} = \gamma_a|\mathcal{E}|^2\rho_{aa} - \gamma_c^{(2)}\rho_{cc}.$$

Here

$$\Delta = \Omega - \omega_{ab}, \quad \Delta_{ab} = \Pi_a(\Omega) - \Pi_b(\Omega), \quad \gamma_{ab} = \kappa_\Omega(\Pi_a(\Omega) - \Pi_b(\Omega))^2/2\hbar^2 + \gamma_a/2,$$

$$\gamma_a = \kappa_q|\Pi_{ac}(-\Omega)|^2/\hbar^2, \quad \Pi_\alpha(\Omega) = \sum_{\alpha'} \frac{|d_{\alpha\alpha'}|^2}{\hbar}\left(\frac{1}{\omega_{\alpha\alpha'} + \Omega} + \frac{1}{\omega_{\alpha\alpha'} - \Omega}\right).$$

Two-quantum relaxation couples the resonant transition under study with the level E_c. If this level does not lie in the gap in the photon density of states, then the weak ''arrival'' in the level E_c and the rapid ''exit'' from it would have virtually no effect on the dynamics of the resonant transition. However, since the ''exit'' from the level E_c is suppressed, this level is occupied after a time of the order of $\tau_{\text{two-photon}}$ after the onset of the resonant interaction of the wave (1) with the transition $E_a \to E_b$, and the density of atoms in the resonant levels becomes much less than in the level in the gap in the photonic density of states:

$$\rho_{bb} \cong \epsilon(1 + \gamma_a^{(0)}/\Gamma), \quad \rho_{aa} \cong \epsilon, \quad \rho_{cc} \cong 1 - \epsilon(2 + \gamma_a^{(0)}/\Gamma),$$

where

$$\epsilon = \frac{\gamma_c^{(2)}}{\gamma_a|\mathcal{E}|^2} \ll 1, \quad \Gamma = 2\frac{|d_{ab}\mathcal{E}/\hbar|^2(\gamma_a^{(0)}/2 + \gamma_{ab}|\mathcal{E}|^2)}{(\gamma_a^{(0)}/2 + \gamma_{ab}|\mathcal{E}|^2)^2 + (\Delta - \Delta_{ab}|\mathcal{E}|^2)^2}.$$

It should be underscored that, in contrast to an ordinary medium, in the present situation the two-photon exit from the level E_c is suppressed by the fact that the frequency ω_p falls in the gap in the photon density of states.

The two-photon relaxation times can be estimated numerically using the formula

$$\tau_{\text{two-photon}} \sim \frac{\hbar c^4}{4\pi I \omega_q^3 |\Pi_{ac}(-\Omega)|^2},$$

where I is the intensity of the coherent wave, averaged over the period $2\pi/\Omega$ of fast oscillations. However, the following expression is more convenient for judging the efficiency of the two-photon relaxation processes:

$$\tau_{\text{two-photon}} \sim \frac{1}{\gamma_a^{(0)}}\left(\frac{\Delta\omega_{\alpha a}}{\Lambda}\right)^2\left(\frac{\omega_{ab}}{\omega_q}\right)^3,$$

where Λ is the characteristic Rabi frequency, and $\Delta\omega_{\alpha a} = \Omega - \omega_{\alpha a}$ is the detuning from a quasiresonant level E_α. Under quasiresonance conditions $|\Delta\omega_{\alpha a}/\Lambda|^2 \sim 10^1 - 10^2$, and because of the choice of frequencies the parameter $(\omega_{ab}/\omega_q)^3$ can be of the order of $10^{-1} - 1$. Thus the spontaneous two-photon relaxation time is an order of magnitude longer than the conventional spontaneous emission time of an isolated atom, which is several orders of magnitude shorter than the lifetime of an excited level in the gap in the

photon density of states.[7,9] The quasiresonance condition can be easily satisfied for a nonresonant case for a wide range of intensities by adjusting the frequency Ω. For a resonant situation, satisfaction of quasiresonance conditions is limited by the structure of the energy levels of the impurity atom and, for this reason, for unfavorable parameters the two-photon relaxation times can be another one or two orders of magnitude longer.

In conclusion, I underscore that the two-photon relaxation operator and the corresponding relaxation parameters used here were obtained in Ref. 16 by switching from the total Hamiltonian, describing the electric dipole interaction of classical and quantum fields with an atom, to an effective Hamiltonian and then to the Bloch kinetic equations, while many investigations of spontaneous radiative relaxation processes in external fields[17–20] proceed from kinetic equations of a different kind, as a result of which the indicated two-photon processes are neglected. The complete picture of two-photon relaxation processes likewise is hidden in the quantum jump technique[21] in application to the initial Hamiltonian, since an infinite number of terms corresponding to various resonant combinations must be summed there. The approach developed in Ref. 16 is best suited for studying relaxation processes in photonic crystals in external fields.

This work was performed as part of INTAS Project No. 96-0339.

*)e-mail: ashat@amicom.ru; ashat@ashat.mephi.su

[1] C. M. Soukoulis (Ed.), *Photonic Gap Materials*, Vol. 315 of NATO ASI Ser. E (Kluwer Academic, Dordrecht, 1996).
[2] C. T. Chan, S. Datta, K. M. Ho, and C. M. Soukoulis, Phys. Rev. B **49**, 1988 (1994).
[3] U. Gruning, V. Lehmann, and C. M. Engelhardt, Appl. Phys. Lett. **66**, 3254 (1995).
[4] S. John and T. Quang, Phys. Rev. A **50**, 1764 (1994).
[5] S. Bay, P. Lambropoulos, and K. Molmer, Phys. Rev. Lett. **79**, 2654 (1997).
[6] N. Vats and S. John, Phys. Rev. A **58**, 4168 (1998).
[7] S. John and J. Wang, Phys. Rev. B **43**, 12772 (1991).
[8] S. John and T. Quang, Phys. Rev. A **52**, 4083 (1995).
[9] A. M. Basharov, Zh. Eksp. Teor. Fiz. **115**, 30 (1999) [JETP **88**, 16 (1999)].
[10] T. Quang, M. Woldeyohannes, S. John, and G. S. Agarval, Phys. Rev. Lett. **79**, 5238 (1997).
[11] D. M. Greenberger and A. Zeilinger (Eds.), *Fundamental Problems in Quantum Theory*, Vol. 755 (Academy of Sciences, New York, 1995).
[12] S. John and T. Quang, Phys. Rev. A **54**, 4479 (1996).
[13] S. John, Phys. Rev. Lett. **53**, 2169 (1984); Phys. Rev. B **31**, 304 (1984).
[14] T. Quang and S. John, Phys. Rev. A **56**, 4273 (1997).
[15] Hu Huang, Xing-Hua Lu, and Shi-Yao Zhu, Phys. Rev. A **57**, 4945 (1998).
[16] A. M. Bashsarov, Zh. Éksp. Teor. Fiz. **102**, 1126 (1992) [Sov. Phys. JETP **75**, 611 (1992)].
[17] L. M. Narducci, M. O. Scully, G.-L. Oppo *et al.*, Phys. Rev. A **42**, 1630 (1990).
[18] O. Kocharovskaya, Shi-Yao Zhu, M. O. Scully *et al.*, Phys. Rev. A **49**, 4928 (1994).
[19] O. Kocharovskaya, P. Mandel, and M. O. Scully, Phys. Rev. Lett. **74**, 2451 (1995).
[20] G. S. Agarwal, Phys. Rev. A **54**, R3734 (1996).
[21] H. J. Carmichael, Phys. Rev. A **56**, 5065 (1997).

Translated by M. E. Alferieff

New friction force due to spontaneous light pressure

O. N. Prudnikov,*) A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin
Novosibirsk State University, 630090 Novosibirsk, Russia

(Submitted 19 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 439–444 (10 October 1999)

We have discovered a new friction force, acting on an atom in the field of two oppositely propagating elliptically polarized waves of low intensity. In contrast to the well-known friction forces, the new force does not vanish at zero detuning of the field from resonance, and the direction of the kinetic process (heating or cooling) is determined by the relative orientation and the ellipticity of the polarization vectors of the oppositely propagating waves. © *1999 American Institute of Physics.* [S0021-3640(99)00519-8]

PACS numbers: 32.80.−t, 42.50.Vk

1. As is well known, an atom in a resonant light field is subjected to induced and spontaneous light pressure.[1,2] For example, in the region of sub-Doppler laser cooling in nonuniformly polarized low-intensity fields, the induced light pressure force is conventionally attributed to the gradient of the ellipticity (in a field with $lin\perp lin$ configuration), and the spontaneous light pressure force is attributed to the gradient of the orientation of the polarization ellipse of the field (in a field with $\sigma_+ - \sigma_-$ configuration).[3] The induced force can be interpreted as being due to the spatial nonuniformity of optical shifts of magnetic sublevels, which results in a Sisyphean friction mechanism. The spontaneous light pressure force also leads to sub-Doppler cooling, but the friction mechanism is different here. Specifically, the motion of the atom gives rise to elliptically orientation of the ground state, as a result of which an imbalance of the spontaneous light pressure forces due to the oppositely propagating waves arises. For all known mechanisms of laser cooling, characteristically, the friction force in the light wave is an odd function of the detuning and therefore vanishes at exact resonance. This is valid for a simple model of a two-level atom, in which the polarization aspects of the interaction are completely neglected,[1,2] and for an atom with degenerate energy levels, which moves in a uniformly polarized field[4] and in nonuniformly polarized fields of known configuration.[3,5] The question of the existence of other friction mechanisms and forces, which depend differently on the detuning, in light fields with a more general configuration remains open.

In the present letter we examine the one-dimensional motion of atoms with an optical transition $j_g = 1/2 \rightarrow j_e = 3/2$ in the field of a standing light wave in which all possible gradients (of the phase, polarization, and intensity) are present. A new friction force, different from the forces described previously in the literature, has been found. By its very nature this force is a spontaneous light pressure force, but the friction mechanism, similarly to the Sisyphean mechanism, is based on the delay of the orientation (difference

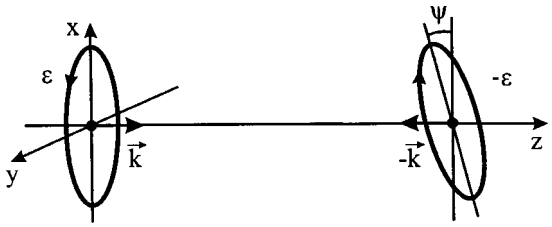

FIG. 1. ε–*angle*–$\overline{\varepsilon}$ spatial field configuration produced by two oppositely traveling waves with opposite elliptic polarizations ε and $-\varepsilon$; ψ is the angle between the semimajor axes of the polarization ellipses of the oppositely traveling waves.

of the populations of the magnetic sublevels) of the ground state with respect to variations of the optical potentials. In contrast to the well-known Sisyphean friction,[3] in our case the spatial nonuniformity of the delay effect plays a large role, i.e., the coordinate dependence of the relaxation rate of the orientation of an atom. A remarkable feature of the new friction force is that it is an even function of the detuning, and the friction does not vanish at exact resonance. We note that this property, though unexpected, is consistent with the general symmetry relations with respect to time reversal.[6] The direction of the kinetic process (heating or cooling) depends on the field configuration, i.e., on the ellipticity and relative orientation of the polarization vectors of the oppositely propagating waves.

2. Let us consider one-dimensional (along the z axis) motion of atoms with total angular momenta $j_g=1/2$ in the ground state and $j_e=3/2$ in the excited state in a field of two oppositely propagating waves with the same frequency and intensity, i.e., the light field at the point z can be written as

$$\mathbf{E}(z)=E_0\mathbf{e}(z)e^{-i\omega t}+\text{c.c.} \tag{1}$$

Here E_0 is the amplitude of each of the oppositely propagating waves, and the vector $\mathbf{e}(z)=a_+\mathbf{e}_+ + a_-\mathbf{e}_-$ with cyclic components $a_+(z)$ and $a_-(z)$ determines the local polarization ellipse and the phase of the field. Next, for greater definiteness, we shall consider the case where the oppositely propagating waves have the same degree of ellipticity but opposite directions of rotation of the polarization vectors, while the semimajor axes of the polarization ellipses make an angle ψ (Fig. 1). In this case we have

$$a_+=-\cos(\varepsilon-\pi/4)e^{ikz}-\cos(\varepsilon+\pi/4)e^{-i\psi}e^{-ikz},$$
$$a_-=\cos(\varepsilon+\pi/4)e^{ikz}+\cos(\varepsilon-\pi/4)e^{i\psi}e^{-ikz}. \tag{2}$$

The parameter ε characterizes the degree of ellipticity of the oppositely propagating waves ($|\tan\varepsilon|$ is the ratio of the semiminor axis of the polarization ellipse to the semimajor axis). Without loss of generality it can be assumed that $-\pi/4\leqslant\varepsilon\leqslant\pi/4$, where $\varepsilon=\pm\pi/4$ corresponds to circular polarization of the waves and $\varepsilon=0$ corresponds to linear polarization. In what follows, we shall use the notation ε–*angle*–$\overline{\varepsilon}$ for the field configuration in Fig. 1. We note that the well-known $lin\perp lin$ ($\varepsilon=0$, $\psi=\pi/2$) and $\sigma_+ - \sigma_-$ ($\varepsilon=\pi/4$) configurations are particular cases of the configuration considered here.

In the resonance approximation, the atom–field interaction operator (1) can be written in the form $\hat{H}_{af}=\hat{V}+\hat{V}^\dagger$, where

$$\hat{V}(z)=\hbar\Omega\left[a_-\left(\frac{1}{2}|j_e,-3/2\rangle\langle j_g,-1/2|+\frac{\sqrt{3}}{6}|j_e,-1/2\rangle\langle j_g,1/2|\right)\right.$$
$$\left.+a_+\left(\frac{1}{2}|j_e,3/2\rangle\langle j_g,1/2|+\frac{\sqrt{3}}{6}|j_e,1/2\rangle\langle j_g,-1/2|\right)\right]. \quad (3)$$

Here $\Omega=-dE_0/\hbar$ is the Rabi frequency and d is the reduced dipole moment.

The force acting on the atom can be expressed, by definition, in terms of the spatial gradient of the interaction operator $\hat{H}_{af}$:

$$f=\langle -d\hat{H}_{af}/dz\rangle, \quad (4)$$

where $\langle\ldots\rangle$ denotes quantum-mechanical averaging over the internal degrees of freedom. In the present work we shall consider weak fields, such that the saturation parameter

$$S=|\Omega|^2/(\gamma^2/4+\delta^2)\ll 1, \quad (5)$$

where γ is the rate of spontaneous relaxation of the excited state and $\delta=(\omega-\omega_0)$ is the detuning from the atomic transition frequency ω_0. Moreover, we shall confine our attention to slow atoms, so that an atom moving with velocity v traverses in the spontaneous decay time γ^{-1} a distance that is small compared with the wavelength:

$$kv\ll\gamma. \quad (6)$$

When the conditions (5) and (6) are satisfied, the optical coherences can be excluded adiabatically (see, for example, Ref. 3), and the force (4) for any transitions $j_g\to j_e$ can be expressed in terms of the Wigner density matrix $\hat{\rho}^{gg}$ of the ground state:

$$f=-\frac{\delta}{\hbar(\delta^2+\gamma^2/4)}Tr\left\{\left(\frac{\partial}{\partial z}\hat{V}^\dagger\right)\hat{V}\hat{\rho}^{gg}+\hat{V}^\dagger\left(\frac{\partial}{\partial z}\hat{V}\right)\hat{\rho}^{gg}\right\}$$
$$+\frac{i\gamma/2}{\hbar(\delta^2+\gamma^2/4)}Tr\left\{\left(\frac{\partial}{\partial z}\hat{V}^\dagger\right)\hat{V}\hat{\rho}^{gg}-\hat{V}^\dagger\left(\frac{\partial}{\partial z}\hat{V}\right)\hat{\rho}^{gg}\right\}. \quad (7)$$

The first term (proportional to δ), by its very nature, is the force due to the induced light pressure, and the second term (proportional to γ) is the force due to the spontaneous light pressure. Since the friction mechanism which we have discovered is related with the spontaneous light pressure force, in the present work we shall analyze the second term, which for the transition $j_g=1/2\to j_e=3/2$ in a field with configuration ε–*angle*–$\overline{\varepsilon}$ has the form

$$f_{sp}=F_+\rho_++F_-\rho_-=\frac{\hbar k\gamma S}{6}\sin(2\varepsilon)\Pi, \quad (8)$$

where $F_+=-F_-=\gamma S\sin(2\varepsilon)/6$ are the partial spontaneous light pressure forces, acting on an atom in the ground state with the angular momentum projection $m=1/2$ or $m=-1/2$, respectively; $\Pi=(\rho_+-\rho_-)$ is the difference of the populations of these states or the optical orientation of an atom. The physical meaning of expression (8) is obvious: The force f_{sp} is the result of averaging of the two forces F_+ and F_- which have opposite signs and act on an atom with the corresponding probability weights ρ_+ and ρ_-.

The stationary population difference in the limit (5) satisfies the equation[5]

$$v\,\frac{\partial \Pi}{\partial z}=(R_+-R_-)-(R_++R_-)\Pi, \tag{9}$$

where $R_\pm=|a_\pm|^2\gamma S/18$. The quantity $R=R_++R_-$ characterizes the rate with which Π relaxes. For the ε–*angle*–$\overline{\varepsilon}$ configuration we have

$$R=\gamma S(1+\cos(2\varepsilon)\cos(\psi)\cos(2kz))/9. \tag{10}$$

In the general case, where $\varepsilon\neq\pm\pi/4$ and $\psi\neq\pm\pi/2$, the relaxation rate R depends on the coordinate z.

3. The force (8), acting on an atom moving with velocity v satisfying the slowness condition (6) can be found using the exact solution of Eq. (9). For clarity, we confine ourselves to the friction force in first order in the velocity, i.e., we shall assume the stronger condition

$$kv\ll\gamma S,$$

which means that the orientation of the moving atom adiabatically follows the variation of the field. In this approximation

$$\Pi(z,v)\approx\Pi^{(0)}(z)-vR^{-1}\frac{\partial}{\partial z}\Pi^{(0)}(z)\approx\Pi^{(0)}(z-vR^{-1}), \tag{11}$$

where $\Pi^{(0)}(z)=(R_+-R_-)/R$ is the stationary orientation of the atom at rest. Substituting expression (11) into Eq. (8), we obtain the force up to first order in the velocity:

$$f_{sp}(z,v)=f_{sp}^{(0)}(z)+\xi(z)v. \tag{12}$$

The first term is the gradient of the force due to the spontaneous light pressure

$$f_{sp}^{(0)}=-\frac{\hbar k\gamma S}{12}\,\frac{\sin(4\varepsilon)\sin(\psi)\sin(2kz)}{(1+\cos(2\varepsilon)\cos(\psi)\cos(2kz))}. \tag{13}$$

This force is present because in a light field with a gradient of the polarization a spatially nonuniform orientation of the atom arises as a result of optical pumping processes, and this results in an imbalance of the spontaneous light pressure forces due to the counterpropagating waves. Thus, besides the well-known optical potential associated with the induced light pressure forces, there arises an optical potential due to the spontaneous light pressure forces. The depth of this potential depends on the ellipticity parameter ε and the relative orientation angle ψ between the counterpropagating waves.

The second term in Eq. (12), proportional to the velocity, is the friction force, where

$$\xi(z)=\frac{3\hbar k^2}{2}\,\frac{\sin(4\varepsilon)\sin(\psi)(\cos(2kz)+\cos(2\varepsilon)\cos(\psi))}{(1+\cos(2\varepsilon)\cos(\psi)\cos(2kz))^3} \tag{14}$$

is the coefficient of friction.

It is easy to show that the average (over the period of the field) orientation $\overline{\Pi^{(0)}}$ of an atom at rest is zero and therefore the average (over the period of the field) gradient of the force vanishes. The average coefficient of friction is determined by the expression

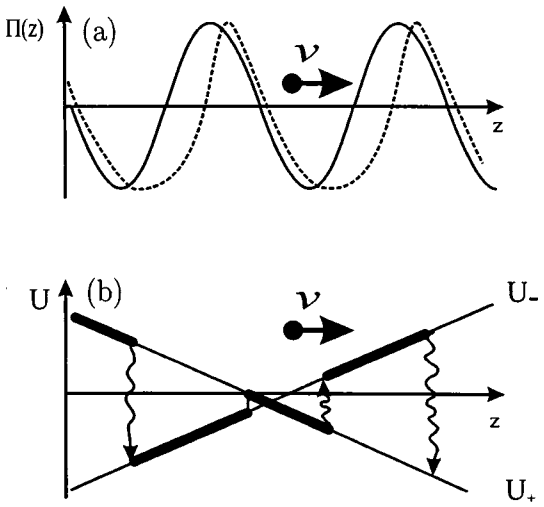

FIG. 2. Sisyphean friction mechanism for spontaneous light pressure forces: a) Orientation of the atom $\Pi(z,v)$: solid line — orientation of the atom at rest, dashed line — orientation of the moving atom. b) Potentials U_+ and U_-, corresponding to the forces F_+ and F_-. The thick line with the arrows marks the most probable trajectory of an atom.

$$\bar{\xi} = -\frac{3\hbar k^2}{8}\frac{\sin(4\varepsilon)\cos(2\varepsilon)\sin(2\psi)}{(1-\cos^2(2\varepsilon)\cos^2(\psi))^{3/2}}, \tag{15}$$

whence one can see that there exists a friction in zeroth order in S, i.e., sub-Doppler cooling is possible.[3] Moreover, even more surprisingly, the coefficient of friction does not depend on the detuning δ of the field, and its sign and therefore the direction of the kinetic process are determined by the configuration of the field (the sign of the product $\sin(2\varepsilon)\sin(2\psi)$).

The velocity dependence of the force (12) is due to the delay effect, which is due to the fact that the relaxation of the orientation Π occurs in a finite time R^{-1} (9). Then, as one can see from Eq. (11), for an atom moving with a low velocity v, the orientation $\Pi(z,v)$ of the atom at the point z is approximately the same as the stationary orientation $\Pi^{(0)}(z-vR^{-1})$ at a point shifted by the amount vR^{-1}. As shown in Ref. 3, the delay due to the motion of an atom in two different potentials $U_+(z)$ and $U_-(z)$ results in Sisyphean friction.

The specific feature of our case is that the forces F_+ and F_- [see Eq. (8)] do not depend on the coordinate z and therefore the spatially uniform delay effect does not result in friction. Indeed, when R does not depend on the coordinate ($\varepsilon = \pm\pi/4$ or $\psi = \pm\pi/2$), the population difference of the moving atom is determined by the spatially uniform shift $\Pi^{(0)}(z-vR^{-1})$ of the function $\Pi^{(0)}(z)$ and therefore its average over the period of the field is zero. Thus, the main reason for the friction in our case is the nonuniformity of the delay effect [the spatial dependence $R(z)$ (10)]. The orientation of the moving atom is not simply shifted, but it is also deformed (Fig. 2a), acquiring a nonzero value on the average, and therefore the atom is subjected to the forces F_+ and

F_- with unequal probability. This mechanism is depicted qualitatively in Fig. 2b, where it is shown that the moving atom is located predominantly in one of the potentials U_+ or U_-.

4. In conclusion, we note that the presence of a contribution in the friction force that does not depend on the detuning of the field is consistent with the general symmetry relation with respect to time reversal. Applying the time reversal operation in the equations for the density matrix, we obtain for the average (over the spatial period) force the relation[6]

$$\overline{f(v,\delta,\{\mathbf{e}(z)\})} = \overline{f(-v,-\delta,\hat{T}\{\mathbf{e}(z)\})}. \tag{16}$$

Here $\{\mathbf{e}(z)\}$ is the spatial configuration of the field, and $\hat{T}\{\mathbf{e}(z)\}$ is the time-reversed configuration. As one can see from Eq. (16), if the configuration of the time-reversed field $\hat{T}\{\mathbf{e}(z)\}$ is equivalent to the initial configuration $\{\mathbf{e}(z)\}$, then the part of the average force (friction) that is linear in the velocity is an odd function of δ. Here equivalence means that $\hat{T}\{\mathbf{e}(z)\}$ can be reduced to $\{\mathbf{e}(z)\}$ by any spatial transformations that leave the Hamiltonian of the atom unchanged (shift, inversion, rotation around the z axis). However, if the field configurations $\{\mathbf{e}(z)\}$ and $\hat{T}\{\mathbf{e}(z)\}$ are nonequivalent, then the symmetry relation (16) does not prevent the friction coefficient from having a contribution that is even in the detuning and, specifically, does not depend on it.

Analyzing an ε–*angle*–$\overline{\varepsilon}$ field directly, it can be shown that under the time reversal operation, in the general case, this configuration becomes nonequivalent to the initial configuration. Therefore the fact that the friction force which we have discovered has not been previously investigated seems to be explained by the fact that, ordinarily, very particular variants of field configurations were studied, for which the coefficient of friction is odd in the detuning even from analysis of the symmetry relation (16). Indeed, for $\varepsilon = 0, \pm\pi/4$ (see, for example, Refs. 3 and 5), after the time reversal operation the field remains equivalent to the initial field.[6]

We note that the ε–*angle*–$\overline{\varepsilon}$ configuration of the field is not the only case where such a friction force arises. In studying the general one-dimensional configuration, where the oppositely propagating waves have arbitrary ellipticities ε_1 and ε_2, we found that the friction force with zero detuning does not vanish when $\varepsilon_1 \neq \varepsilon_2$, $\varepsilon_{1,2} \neq \pm\pi/4$, and $\psi \neq 0, \pm\pi/2$ simultaneously.

Moreover, we have investigated the transition $j_g = 1/2 \rightarrow j_e = 1/2$, and we have found that the new friction force, due to spontaneous light pressure, also exists in this case. The friction coefficient is identical to expression (14) taken with the opposite sign.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 98-02-17794).

*)e-mail: llf@admin.nsu.ru

[1] V. G. Minogin and V. S. Letokhov, *Laser Light Pressure on Atoms* (Gordon and Breach, New York, 1987) [Russian original, Nauka, Moscow, 1986].
[2] A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *Mechanical Effect of Light on Atoms* [in Russian], (Nauka, Moscow, 1991).

[3] J. Dalibard and C. Cohen-Tannoudji, J. Opt. Soc. Am. **6**, 2023 (1989).
[4] O. N. Prudnikov, A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **115**, 791 (1999) [JETP **88**, 433 (1999)].
[5] V. Finkelstein, P. R. Berman, and J. Guo, Phys. Rev. A **45**, 1829 (1992).
[6] G. Nienhuis, P. van der Straten, and S-Q. Shang, Phys. Rev. A **44**, 462 (1991).

Translated by M. E. Alferieff

On self-induced transparency in laser–plasma interactions

V. V. Goloviznin
FOM-Institute for Plasma Physics Rijnhuizen, 3430 BE Nieuwegein, The Netherlands; Kurchatov Institute Russian Research Center, 123182 Moscow, Russia

T. J. Schep
FOM-Institute for Plasma Physics Rijnhuizen, 3430 BE Nieuwegein, The Netherlands

(Submitted 15 July 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 445–450 (10 October 1999)

We study fully relativistic nonlinear one-dimensional equations describing steady-state solutions for an electromagnetic wave interacting with a plasma in the self-induced transparency regime. In addition to the well-known solution that corresponds to the transmission of the electromagnetic wave into plasma, another steady-state solution is shown to exist in a certain range of amplitudes of the wave. The latter solution corresponds to total reflection of the incident wave. The coexistence of the two solutions indicates the possibility of hysteretic behavior in the self-induced transparency. © *1999 American Institute of Physics.* [S0021-3640(99)00619-2]

PACS numbers: 52.40.Nk, 42.50.Md

The fast progress of short-pulse laser technology in the last decade has made it possible to study laser–plasma interactions at relativistic intensities of the optical wave.[1,2] Numerous interesting physical phenomena have been predicted and observed in this new regime. One of the basic effects of this kind is the possibility that a strong electromagnetic wave can propagate in overdense plasmas due to the relativistic increase of the electron mass[3–5] — the so-called self-induced transparency. It still remains a theoretical concept that has not been verified experimentally.

A considerable number of analytical and numerical results on self-induced transparency have been published by now (see Refs. 6–10 and references therein). In the present paper we would like to point out that laser–plasma interactions in the self-induced transparency regime can exhibit even more complicated behavior than one is used to thinking. Namely, we have found that instead of only one characteristic value of the amplitude of the optical wave, which defines the onset of self-induced transparency, there actually exist two such thresholds and a transition region between them, where the plasma may assume either the transparent or opaque form at the same intensity of the incident wave. Although our results have been obtained under steady-state assumptions, they can be of relevance for the propagation of a finite-length optical pulse through a plasma, provided that its typical length scale is much larger than the relativistic skin depth of the plasma.

Thus we consider a strong electromagnetic wave interacting at normal incidence

with a semi-infinite plasma layer with a sharp boundary. In the present paper we are specifically interested in the steady-state field and density distribution, which is considered to be the result of an infinitely slow transition from zero amplitude of the incident wave to its current nonzero value. In the one-dimensional (1D) approximation, all parameters of the problem depend only on the coordinate z and time t. To be specific, we assume the plasma to occupy the half space $z>0$; the electromagnetic wave propagates along the z axis from $-\infty$, and its amplitude is constant in time. The ions are assumed immobile, and for the sake of simplicity we consider a steplike ion density distribution: $n_i=n_0\Theta(z)$, where $\Theta(z)$ is the Heaviside step function.

Following Chen and Sudan,[11] we can write the governing set of equations describing the propagation of an electromagnetic wave through ''cold electron fluid'' as (note the use of relativistic units $\hbar=c=1$ throughout this paper):

$$\left(\Delta-\frac{\partial^2}{\partial t^2}\right)\mathbf{a}=\nabla\frac{\partial\phi}{\partial t}+\frac{\omega_p^2 n}{n_0\gamma}(\mathbf{a}-\nabla\psi), \tag{1}$$

$$\partial\psi/\partial t=-\phi-\gamma+1, \tag{2}$$

with $n=n_i-n_0\omega_p^{-2}\Delta\phi$ and $\gamma^2=1+(\mathbf{a}-\nabla\psi)^2$. Here n and γ denote the local density and Lorentz factor of the moving ''electron fluid,'' and $\mathbf{a}$ and ϕ are the dimensionless vector and scalar potentials of the electromagnetic field, respectively. The Coulomb gauge $\nabla\cdot\mathbf{a}=0$ is used. The plasma frequency ω_p is defined in the usual way as $\omega_p^2=4\pi e^2 n_0/m$.

In the 1D approximation the gauge condition reads $\partial a_z/\partial z=0$, which implies that the longitudinal component of the vector potential is irrelevant. The first of the above equations then splits into two:

$$\left(\frac{\partial^2}{\partial z^2}-\frac{\partial^2}{\partial t^2}-\frac{\omega_p^2 n}{n_0\gamma}\right)\mathbf{a}_\perp=0 \tag{3}$$

and

$$\frac{\partial\psi}{\partial z}=\frac{n_0\gamma}{\omega_p^2 n}\frac{\partial^2\phi}{\partial t\partial z}. \tag{4}$$

As we have said, we are searching for possible steady-state solutions when the plasma density does not depend on time: $\partial n/\partial t=0$. It immediately follows from the above equations that the amplitude of the optical field $a_\perp^2$ cannot depend on t, either. One must conclude that a steady-state solution of this type is only possible for circular polarization of the incident wave, when the direction of the optical field varies in time but its amplitude does not. In the following we always assume this to be the case.

As a result, inside the plasma layer ($z\geqslant 0$) the above set of governing equations reduces to just one nonlinear equation for the transverse component of the vector potential:

$$\left(\frac{\partial^2}{\partial z^2}+\omega_0^2\right)\mathbf{a}_\perp=\frac{\mathbf{a}_\perp}{(1+a_\perp^2)^{1/2}}\left(\omega_p^2+\frac{\partial^2(1+a_\perp^2)^{1/2}}{\partial z^2}\right), \tag{5}$$

where all time dependence is assumed to be given by the factor $\exp(i\omega_0 t)$, with ω_0 being the frequency of the incident optical wave.

If a solution $\mathbf{a}_\perp(z)$ is known, the corresponding particle density distribution can be found as

$$n(z)=n_i+\frac{m}{4\pi e^2}\frac{\partial^2(1+a_\perp^2)^{1/2}}{\partial z^2}. \tag{6}$$

Any physically meaningful electron density distribution must conserve the total charge, which means that

$$\int_{-\infty}^{\infty}(n(z)-n_i)dz=0.$$

Substituting here Eq. (6), one sees that the boundary of the electron cloud in the steady state must be shifted with respect to the edge of the ion density profile by some distance z_0, where z_0 satisfies the following equation (under the assumption that the z-derivative of the amplitude of the optical field vanishes at $+\infty$):

$$z_0=-\frac{1}{\omega_p^2}\frac{\partial(1+a_\perp^2)^{1/2}}{\partial z}\bigg|_{z=z_0}. \tag{7}$$

All of our considerations imply that the plasma electrons always remain inside the plasma layer (that is, to the right of $z=0$), which is quite natural because the light pressure of the incoming wave is applied from the left. An additional consistency condition is then $z_0\geqslant 0$, which means, in turn, that

$$\frac{\partial a_\perp^2}{\partial z}\bigg|_{z=z_0}\leqslant 0. \tag{8}$$

Note that $z=0$ is no longer a physical boundary for the transverse electromagnetic wave; the actual boundary is $z=z_0$.

For further analysis, it is convenient to write the circularly polarized field $\mathbf{a}_\perp$ in terms of its amplitude $T(z)$ and phase $\theta(z)$ so that $a_x=T(z)\cos(\theta(z)+\omega_0 t)$, $a_y=T(z)\sin(\theta(z)+\omega_0 t)$. Equation (5) then splits into two nonlinear scalar equations. This set of equations can be easily shown to have two constants of motion, C_1 and C_2, which fully determine possible steady-state distributions of the electromagnetic field and the particle density inside the plasma. In terms of the constants of motion, the equation for the field amplitude can be conveniently written as

$$(\partial T/\partial z)^2=(1+T^2)[C_2-C_1^2T^{-2}-\omega_0^2T^2+2\omega_p^2(1+T^2)^{1/2}]. \tag{9}$$

It now looks similar to the equation of motion of an anharmonic oscillator, where z plays the role of ''time.'' One sees that, in general, there are two turning points $T_{\min}$ and $T_{\max}$, where $\partial T/\partial z$ vanishes, so that the amplitude of the optical field T oscillates between these limits. From a physical point of view, such oscillations correspond to the interference pattern of two counterpropagating plane waves. They must be of relevance for a problem with two boundaries, when part of the energy can be reflected from the right boundary

and interfere with the wave coming from the left. In the present paper we restrict our consideration to monotonic solutions, which is the most natural choice for a semi-infinite layer.

Under the additional condition that the amplitude of the wave is monotonic at $+\infty$, we still have two two possibilities:

A. $T(z)=\text{const}$, and $\partial T/\partial z=0$ everywhere inside the plasma. This, in turn, implies that $\partial\theta/\partial z=-k_0=\text{const}$, where

$$k_0^2=\omega_0^2-\omega_p^2(1+T^2)^{-1/2}; \tag{10}$$

B. $\partial T/\partial z$ vanishes at some point T_∞ as $(\partial T/\partial z)^2\sim(T-T_\infty)^\alpha$ with $\alpha\geqslant 2$. This is the case for $T_\infty=0$ (which means no transmitted wave at $+\infty$), and $C_1=0$, $C_2=-2\omega_p^2$. Equation (9) then reduces to

$$(\partial T/\partial z)^2=(1+T^2)[2\omega_p^2(\sqrt{1+T^2}-1)-\omega_0^2T^2]. \tag{11}$$

The first of the above alternatives is just the well-known self-induced transparency regime of propagation.[3] The necessary condition for the existence of such a solution is $k_0^2\geqslant 0$, that is, $\omega_0^2\geqslant\omega_p^2(1+T^2)^{-1/2}$, which is in agreement with earlier results on self-induced transparency. One can also rewrite the above expression as a condition on the amplitude of the wave:

$$T^2\geqslant(\omega_p^4-\omega_0^4)/\omega_0^4. \tag{12}$$

The second option is more complicated. Equation (11) is seen to have solutions under the condition $\omega_p^2\geqslant\omega_0^2$ (one may note that in the limit of $\omega_p\gg\omega_0$ the corresponding solution has been found in Ref. 12). For each choice of the ratio of ω_p/ω_0, the field T that satisfies Eq. (11) cannot exceed some limiting value defined by the right-hand side of Eq. (11):

$$T^2\leqslant 4\omega_p^2(\omega_p^2-\omega_0^2)/\omega_0^4. \tag{13}$$

From the above analysis, one can conclude that, surprisingly, the two solutions **A** and **B** may coexist. Namely, if one switches to new dimensionless variables ν and τ such that

$$\nu=\frac{n_0}{n_{\text{cr}}},\quad \tau=\left(\frac{T}{2}\right)^2\frac{n_{\text{cr}}}{|n_0-n_{\text{cr}}|},$$

there can be found four distinct regions in the (ν,τ) phase space (see Fig. 1):

I. $\nu<1$. This is the usual transparency regime in underdense plasma;

II. $\nu>1$, $\tau<(\nu+1)/4$. Overdense plasma: no transmitted wave, total reflection of the incident energy;

III. $\nu>1$, $\tau>\nu$. Only self-induced transparency solutions in overdense plasma;

IV. $\nu>1$, $(\nu+1)/4<\tau<\nu$. Both transparent and opaque solutions coexist for the same initial plasma density and the same amplitude of the incident wave.

Thus, between the purely reflecting and purely self-induced-transparency regimes there exists a transition region IV in which both solutions are equally possible, and the system has to choose which of the two steady-state solutions will be realized. It is beyond

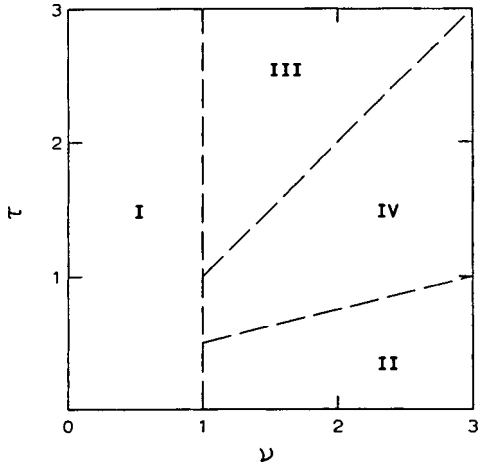

FIG. 1. Four distinct regions in the (ν, τ) phase space: I — the usual transparency regime in underdense plasma, II — the total reflection of a low-intensity wave in overdense plasma, III — the self-induced transparency regime, IV — the transition region.

the scope of the present paper to show how this choice is made (note that the boundary conditions are not sufficient to fix the situation). We just emphasize that the situation is not unique, and that systems that can assume several stationary states usually demonstrate hysteretic behavior, when the current state of the system depends on its history. It still remains to be investigated whether this is the case for self-induced transparency. One must also note that in terms of the (ν, τ) phase-space diagram, an increase of the amplitude of the incident wave corresponds to an increase of τ along a line of constant ν. In realistic situations, there is thus no way to avoid the controversial region IV.

As an example, Fig. 2 shows the vector potential and the electron density distribu-

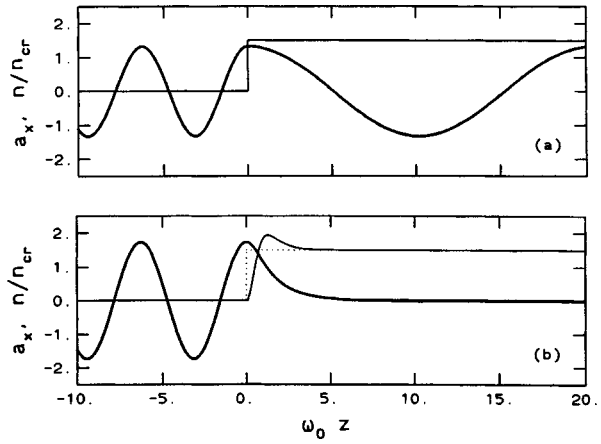

FIG. 2. Dimensionless vector potential a_x (thick line) and the relative electron density $n/n_{\rm cr}$ (thin line) in the vicinity of the plasma boundary for two different regimes of propagation: self-induced transparency (a), and total reflection (b). The input parameters are the same in both cases: $n_0 = 1.5\ n_{\rm cr}$, $a_{\rm in}^2 = 0.75$.

tion in the vicinity of the plasma boundary. The instant in time is chosen in such a way that the x component of the vector potential at $z=0$ is at its maximum. Figure 2a corresponds to the self-induced transparency regime of propagation, and Fig. 2b — to the opaque solution. The amplitude of the incoming wave $a_{\rm in}$ and the plasma density n_0 are the same in both cases. Both parameters are suitably chosen so as to encounter the region where the two solutions coexist: $n_0/n_{\rm cr}=1.5$, $a_{\rm in}^2=0.75$. The field at the left boundary of the plasma layer is found as a numerical solution of the (nonlinear) boundary conditions. For our choice of parameters, the shift of the boundary of the electron cloud appears to vanish: $z_0=0$.

The most noticeable feature of the totally reflecting case is the presence of a strong light pressure which influences the medium. The rearrangement of the plasma electrons under the action of the incident wave is clearly seen in Fig. 2b. The electrons form a smooth maximum situated at about $z\simeq 1.3\omega_0^{-1}$ ($\sim 0.2\,\mu$m for an optical wavelength of $1\,\mu$m). This ''deformation'' of the electron cloud has a dramatic effect on the propagation of light: in Fig. 2a about 72% of the incident energy is transmitted into the plasma, while in Fig. 2b the incoming light is totally reflected. Note that in our case self-induced transparency is already possible for $a_{\rm in}^2=0.3125$. In fact, one sees that even at an irradiance that is 2.4 times higher, the plasma may still ''resist'' transmitting the incoming light.

To conclude, we have considered interaction of a strong optical wave with overdense plasma in the self-induced transparency regime. We have shown that in a certain range of the amplitude of the incident wave two possible steady-state solutions of the corresponding nonlinear equations coexist. One of them describes the well-known self-induced transparency, while the other corresponds to the total reflection of the incident wave. The coexistence of these two solutions indicates that overdense plasmas under the action of a strong laser wave may exhibit hysteretic behavior, when the current state of the system depends on its history.

This work was performed as part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie (FOM) and was supported by the Nederlandse Organizatie voor Wetenschappelijk Onderzoek (NWO).

[1] G. Mourou and D. Umstadter, Phys. Fluids B **4**, 2315 (1992).
[2] M. D. Perry and G. Mourou, Science **264**, 917 (1994).
[3] A. I. Akhiezer and R. V. Polovin, Zh. Éksp. Teor. Fiz. **30**, 915 (1956) [Sov. Phys. JETP **3**, 696 (1956)].
[4] W. Lünow, Plasma Phys. **10**, 879 (1968).
[5] P. K. Kaw and J. Dawson, Phys. Rev. Lett. **13**, 472 (1970).
[6] S. Guerin *et al.*, Phys. Plasmas **3**, 2693 (1996).
[7] S. Guerin *et al.*, Phys. Plasmas **2**, 2807 (1995).
[8] W. B. Mori *et al.*, Phys. Rev. Lett. **72**, 1482 (1994).
[9] C. B. Darrow *et al.*, Phys. Rev. Lett. **69**, 442 (1992).
[10] A. S. Sakharov and V. I. Kirsanov, Phys. Rev. E **49**, 3274 (1994).
[11] X. L. Chen and R. N. Sudan, Phys. Fluids B **5**, 1336 (1993).
[12] R. N. Sudan, Phys. Rev. Lett. **70**, 3075 (1993).

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

Contribution to the theory of the ferromagnetism of metals with a hexagonal close-packed lattice

R. O. Zaĭtsev
Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

(Submitted 15 June 1999; resubmitted 13 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 451–456 (10 October 1999)

The possibility of the appearance of ferromagnetic instability in a hexagonal close-packed system is established on the basis of the idea of a strong interaction in the same unit cell. The conditions for the appearance of spin and orbital instability as a function of the degree of filling of the $2xy$, x^2-y^2 shell are obtained. The physical picture obtained has a direct relationship to ferromagnetism of hexagonal and cubic cobalt, which possesses the highest Curie temperature, 1396 K. © *1999 American Institute of Physics*. [S0021-3640(99)00719-7]

PACS numbers: 75.50.Cc, 75.10.Jm

In the present letter Hubbard's basic idea[1] concerning the possibility of exactly taking into account strong intra-atomic repulsion, which is assumed below to be infinite, is used. Hops between nearest-neighbor atoms occupying lattice sites are taken into account in the next approximation. Our problem is to find the self-consistent equations for finding the occupation numbers as a function of temperature, electron density, and external magnetic field.

According to the electric-neutrality condition, the difference between the number n_h of holes in a four-fold degenerate $2xy$, x^2-y^2 shell and the number n_s of electrons in a $4s$ shell is 1 for cobalt: $n_h=n_s+1$. Various estimates give for the observed magnetic moment $(1.6-1.75)\mu_B$,[2] so that $n_h<2$.

We neglect the hybridization of the *s* and *d* states, so that the total range of variation of the number of holes $0<n_h<2$ will be considered below. This will enable us to study the possibility of the existence of ferromagnetism for the hypothetical model of hexagonal nickel, for which $n_h=n_s<2$.

To study cobalt and nickel with an fcc unit cell, it is necessary to consider the four-fold degenerate $3z^2-r^2$, x^2-y^2 shell, where there is spin but no orbital degeneracy. For this reason, all general results associated with spin instability of hexagonal cobalt and nickel that are obtained below are also valid for the corresponding fcc elements.

Let us assume that hops occur between neighboring atoms, each of which possesses a wave function that is proportional to one of the components $2xy$ or x^2-y^2. To study the magnetic properties associated with orbital splitting, it is necessary to use linear combinations of atomic wave functions that diagonalize the atomic Hamiltonian. Two complex conjugate wave functions $(x\pm iy)^2$, which have different energy in a fixed magnetic field $(=\pm 2H)$, the magnetic field being measured in energy units, satisfy this

condition. The matrix $t_n^m(\mathbf{r})$ of hops between nearest-neighbor ions does not depend on the spin index σ, but does depend strongly on the orbital indices m and n.

Introducing the creation $\hat{a}^+_{(m,\sigma)}(\mathbf{r})$ and annihilation $\hat{a}_{(n,\sigma)}(\mathbf{r})$ operators for hole states in a cell with coordinates $\mathbf{r}$, we write the interaction Hamiltonian in terms of the transition matrix between neighboring atoms:

$$\hat{H}=\sum_{m,n,\sigma,\mathbf{r},\mathbf{r}'} t_n^m(\mathbf{r}-\mathbf{r}')\hat{a}^+_{(m,\sigma)}(\mathbf{r})\hat{a}_{(n,\sigma)}(\mathbf{r}')$$

$$-\sum_{m=\pm,\sigma,\mathbf{r}}(\mu+\sigma H+2mH)\hat{a}^+_{(m,\sigma)}(\mathbf{r})\hat{a}_{(m,\sigma)}(\mathbf{r}). \tag{1}$$

After the zeroth-order Hamiltonian corresponding to nonoverlapping atomic states is diagonalized, the creation and annihilation operators are represented as an expansion in all possible transitions between N and $N\pm1$ particle states (see Ref. 1):

$$\hat{a}^+_{(m,\sigma)}(\mathbf{r})=\sum_\alpha g_\alpha^{m,\sigma}\hat{X}^\alpha_{\mathbf{r}}, \quad \hat{a}_{(n,\sigma)}(\mathbf{r})=\sum_\beta g_\beta^{n,\sigma}\hat{X}^\beta_{\mathbf{r}}. \tag{2}$$

Here the indices α, β correspond to the reciprocal transitions $s\to p$, i.e., $\beta(p,s)=-\alpha(p,s)$. The quantities $g_\alpha^{m,\sigma}$ are called genealogical coefficients and are calculated below. In the simplest case, when we are dealing with transitions between a completely filled (no holes) and nine-electron (one-hole) shells, the genealogical coefficients equal 1, and the inverse matrix of the virtual Green's function has the form

$$\begin{array}{cc} & \begin{array}{cccc} A_+, & A_-, & B_+, & B_- \end{array} \\ \hat{G}^{-1}= \begin{array}{c} A_+ \\ A_- \end{array} & \left(\begin{array}{cccc} i\omega-\epsilon_a^+-f_a^{(+)}\Upsilon_{\mathbf{p}} & -f_a^{(+)}\beta & A & B \\ -f_a^{(-)}\gamma & i\omega-\epsilon_a^--f_a^{(-)}\Upsilon_{\mathbf{p}} & C & D \end{array}\right). \end{array} \tag{3}$$

Here the external field H is introduced, and $\epsilon^m_{a,b}=\epsilon_{a,b}-2mH-\sigma H$, and $2m=\pm2$ is the projection of the orbital angular momentum. The so-called terminal factors $f^{(\sigma,\pm)}_{a,b}$ can be expressed in terms of the average occupation numbers n_0 of empty and $n^{(\sigma,\pm)}_{a,b}$ single-particle states:

$$f^{(\sigma,\pm)}_{a,b}=n_0+n^{(\sigma,\pm)}_{a,b}. \tag{4}$$

The coefficients β and γ can be expressed in terms of the rotation angle of the corners of the unit cell $\varphi=\pi/3$:

$$\Upsilon_{\mathbf{p}}=\cos(p_x)+\cos\left(\frac{1}{2}p_x+\frac{\sqrt{3}}{2}p_y\right)+\cos\left(\frac{1}{2}p_x-\frac{\sqrt{3}}{2}p_y\right),$$

$$\beta=\gamma^*=\cos(p_x)+e^{4i\varphi}\cos\left(\frac{1}{2}p_x+\frac{\sqrt{3}}{2}p_y\right)+e^{-4i\varphi}\cos\left(\frac{1}{2}p_x-\frac{\sqrt{3}}{2}p_y\right), \tag{5}$$

and are the same as in the hopping matrix for the tight-binding approximation. In this case the matrix elements are calculated with respect to the atomic wave functions, belonging to orbital angular momentum $l=2$ and differing by the sign of its projection $l_z=\pm2$.

The transition matrix elements A, B, C, and D between layers in an hcp lattice are proportional to the fourth power of the sine of the angle between the c axis and the vector connecting the nearest-neighbor atoms between the layers. This quantity is close to 1/9, so that for the wave functions $(x \pm iy)^2$ employed the transitions between layers can be neglected and the analysis below can be conducted only in the xy plane.

To establish the conditions of ferromagnetic instability the equation of state must be written taking account of the dependence on the spin ($\sigma = \pm$) and orbital ($m = \pm$) indices. Calculating the diagonal matrix elements of the matrix (3), we obtain a closed system of equations for four occupation numbers:

$$n_m^{(\sigma)} = f_m^{(\sigma)} \sum_{\mathbf{p},\lambda=\pm} A_{\mathbf{p}}^{(sgn(m\lambda))} n_F(\xi_{\mathbf{p}}^{\lambda}); \quad f_m^{(\sigma)} = n_0 + n_m^{(\sigma)}; \quad n_0 + \sum_{\sigma,m=\pm} n_m^{(\sigma)} = 1. \tag{6}$$

The latter relation is a normalization condition, signifying that we are taking account of transitions only between the zero-hole (0) and single-hole states. The normal coordinates and the excitation spectrum are expressed only in terms of the external field and the terminal factors:

$$A_{\mathbf{p}}^{(\pm)} = \frac{1}{2}\left\{1 \pm \frac{r_{\mathbf{p}}}{\sqrt{r_{\mathbf{p}}^2 + 4\Phi_+^{\sigma}\Phi_-^{\sigma}\Delta_{\mathbf{p}}}}\right\}; \quad r_{\mathbf{p}} = -4H + \Upsilon_{\mathbf{p}}(\Phi_+^{\sigma} - \Phi_-^{\sigma}); \tag{7}$$

$$\xi_{\mathbf{p}}^{(\pm)} = \frac{1}{2}\Upsilon_{\mathbf{p}}(\Phi_+^{(\sigma)} + \Phi_-^{(\sigma)}) \pm \sqrt{\frac{r_{\mathbf{p}}^2}{4} + \Phi_+^{(\sigma)}\Phi_-^{(\sigma)}\Delta_{\mathbf{p}}} - \sigma H - \mu;$$

$$\Delta_{\mathbf{p}} = \Upsilon_{\mathbf{p}}^2 - 3\Pi_{\mathbf{p}}, \quad p_1 = p_x, \quad p_2 = \frac{p_x}{2} + \frac{p_y\sqrt{3}}{2}, \quad p_3 = \frac{p_x}{2} - \frac{p_y\sqrt{3}}{2};$$

$$\Phi_{\pm}^{(\sigma)} = f_{\pm}^{(\sigma)} = n_0 + n_{I,\pm}^{(\sigma)}; \quad \Upsilon_{\mathbf{p}} = \sum_{k=1}^{3} \cos p_k; \quad \Pi_{\mathbf{p}} = \sum_{k,n;k>n} \cos p_k \cos p_n. \tag{8}$$

The equations obtained solve the problem of calculating all occupation numbers assuming an infinite Hubbard energy and in the zeroth self-consistent field approximation.

The situation is somewhat more complicated when the system resonates between four single-hole and three high-spin two-hole states with spin 1 and occupation numbers n_{II}^0 and n_{II}^{σ}.[3] In this case the creation and annihilation operators are expressed in terms of the X operators using four genealogical coefficients, $g_{(1,m)} = \mathrm{sgn}(m)$ and $g_{(2,m)} = \mathrm{sgn}(m)/\sqrt{2}$, and the right-hand side of the self-consistency equations depends on a single combination:

$$\Phi_m^{(\sigma)} = g_{(1,m)}^2 f_{(1,m)}^{(\sigma)} + g_{(2,m)}^2 f_{(2,m)}^{(\sigma)}, \quad f_{(1,m)}^{(\sigma)} = n_{\mathrm{II}}^{(\sigma)} + n_{(\mathrm{I},-m)}^{(\sigma)}; \quad f_{(2,m)}^{(\sigma)} = n_{\mathrm{II}}^{(0)} + n_{(\mathrm{I},-m)}^{(-\sigma)}. \tag{9}$$

Correspondingly, the equation of state is

$$g_{(1,m)}^2 n_{\mathrm{II}}^{(\sigma)} + g_{(2,m)}^2 n_{\mathrm{II}}^{(0)} = \Phi_m^{(\sigma)} R_m^{(\sigma)} = \Phi_m^{(\sigma)} \sum_{\mathbf{p},\lambda=\pm} A_{\mathbf{p}}^{(sgn(m\lambda))} n_F(\xi_{\mathbf{p}}^{\lambda}). \tag{10}$$

We write the missing equations using two pairs of auxiliary coefficients $\gamma_{(1,m)}$ and $\gamma_{(2,k)}$ satisfying orthogonality conditions

TABLE I.

Interval	R	f_h	β_h	γ_s	γ_l	g^2
$0<n_h<1$	$n_h/(4-3n_h)$	$1-3n_h/4$	4	0	0	1
$1<n_h<2$	$4(n_h-1)/(2+n_h)$	$(2+n_h)/12$	3	1/3	-1	3/2
$2<n_h<3$	$3(n_h-2)/(6-n_h)$	$(6-n_h)/12$	4	$-4/3$	0	3/2
$3<n_d<4$	$4(n_h-3)/(3n_h-8)$	$(3n_h-8)/4$	1	-1	-1	1

$$g_{(1,m)}\gamma_{(1,m)}+g_{(2,m)}\gamma_{(2,m)}=0. \tag{11}$$

The right-hand side of these equations contains the same function $R_m^{(\sigma)}$ as in Eqs. (10). As a result we obtain the relation

$$g_{(1,m)}\gamma_{(1,m)}n_{\mathrm{II}}^{(\sigma)}+g_{(2,m)}\gamma_{(2,m)}n_{\mathrm{II}}^{(0)}=[g_{(1,m)}\gamma_{(1,m)}f_{1,m}^{(\sigma)}+g_{(2,m)}\gamma_{(2,m)}f_{2,m}^{(0)}]R_m^{(\sigma)}. \tag{12}$$

It is evident that Eqs. (12), taking account of the orthogonality condition (11) and the natural condition for variation of the occupation numbers for states with zero spin projection $\delta n_{\mathrm{II}}^{(0)}=0$, lead to relations which in the linear approximation in δH do not depend explicitly on the applied field:

$$\delta n_{\mathrm{II}}^{(\sigma)}[1-R_m^{(\sigma)}]=[\delta n_{\mathrm{I},-m}^{(\sigma)}-\delta n_{\mathrm{I},-m}^{(-\sigma)}]R_m^{(\sigma)}. \tag{13}$$

Thus, the variation of the two-hole occupation numbers is antisymmetric in the spin index and can be eliminated from the left-hand side of the equation of state (10):

$$g_{(1,m)}^2\delta n_{\mathrm{II}}^{(\sigma)}=\frac{1}{2}(\delta\Phi_m^{(\sigma)}-\delta\Phi_{-m}^{(\sigma)})L_m^{(\sigma)}+\frac{1}{2}(\delta\Phi_m^{(\sigma)}+\delta\Phi_{-m}^{(\sigma)})S_m^{(\sigma)}+\delta\Phi_m^{\sigma}R_m^{(\sigma)}. \tag{14}$$

Here three coefficients, which are independent of the spin and orbital indices in the limit of zero external field, are determined:

$$R_m^{(\sigma)}=R=\frac{1}{2}\sum_{\lambda,\mathbf{p}}n_F(\xi^{\lambda}(\mathbf{p})); \quad S_m^{(\sigma)}=Q_s=\frac{1}{2}\sum_{\lambda,\mathbf{p}}[\xi^{\lambda}(\mathbf{p}))+\mu]n_F'(\xi^{\lambda}(\mathbf{p})); \tag{15}$$

$$L_m^{(\sigma)}=Q_l=-\sum_{\mathbf{p}}\frac{\Upsilon_{\mathbf{p}}}{2\sqrt{\Upsilon_{\mathbf{p}}^2-3\Pi_{\mathbf{p}}}}\{n_F[\xi_{\mathbf{p}}^{(-)}]-n_F[\xi_{\mathbf{p}}^{(+)}]\}. \tag{16}$$

To these relations we must add the equation of state and the form of the excitation energy for $H=0$:

$$n_h=[n_h]+\beta_h f_h R; \quad \xi_{(\mathbf{p})}^{(\pm)}=g^2 f_h(\Upsilon_{\mathbf{p}}\pm\sqrt{\Upsilon_{\mathbf{p}}^2-3\Pi_{\mathbf{p}}})-\mu. \tag{17}$$

Dividing the equations into even and odd with respect to the spin index, which lead to even and odd solutions with respect to the sign of the projection of the orbital angular momentum, we find the solvability conditions in the following general form:

$$R(1-R)=Q_s(\gamma_s+R), \quad R(1-R)=Q_l(\gamma_l+R). \tag{18}$$

The function R and the coefficients f_h, β_h, $\gamma_{s,l}$, and g^2 were calculated for each integer range of variation of the number n_h of holes and are presented in Table I.

At $T=0$ the condition for the appearance of ferromagnetism is determined in terms of the density of states at the Fermi surface, with $\overline{\mu}=\mu/f_h$,

$$R(1-R)=-\mu\frac{1}{2}\sum_{\lambda,\mathbf{p}}\delta(\xi_{\mathbf{p}}^{(\lambda)})(\gamma_s+R)=-\overline{\mu}\frac{1}{2}\sum_{\lambda}\rho^{(\lambda)}(\epsilon=\overline{\mu})(\gamma_s+R). \tag{19}$$

All general relations (15)–(19) referring to spin instability are also applicable to the fcc elements Co and Ni. It is sufficient to set $Q_l=0$ and calculate the spectrum of elementary excitations for $H=0$:

$$\xi_{\mathbf{p}}^{(\pm)}=f_h(\Upsilon_{\mathbf{p}}\pm\sqrt{\Upsilon_{\mathbf{p}}^2-3\Pi_{\mathbf{p}}})-\mu,$$

where

$$\Upsilon_{\mathbf{p}}=\cos p_x\cos p_y+\cos p_y\cos p_z+\cos p_z\cos p_x;$$

$$\Pi_{\mathbf{p}}=\cos p_x\cos p_y\cos p_z\sum_{k=x,y,z}\cos p_k. \tag{20}$$

To study the conditions under which spin and orbital instabilities arise for a hexagonal system, it is necessary to calculate two densities of states:

$$\rho^{(\pm)}(\epsilon)=\sum_{\mathbf{p}}\delta(\epsilon-\Upsilon_{\mathbf{p}}\mp\sqrt{\Upsilon_{\mathbf{p}}^2-3\Pi_{\mathbf{p}}}).$$

Substituting here the definition of the functions $\Upsilon_{\mathbf{p}}$ and $\Pi_{\mathbf{p}}$ from Eq. (8), we determine a system of singular points inside the region $(0<p_{2,3}<2\pi)$.

When the number of holes is small $(n_h\ll 1)$ and their energy is close to the minimum value -3, the system tends toward ferromagnetism. However, neither spin nor orbital instability arises here. The reason is that the product of the density of states by the minimum energy at this point is small compared to 1 $(\approx 1/3)$. As the energy increases to the level of six equivalent saddle points $(\approx -2\frac{1}{16})$, the density of states increases sharply, which makes it possible to satisfy the spin instability condition (19). Therefore orbital instability does not arise in the region $n_h<1$, while spin instability should occur in a narrow concentration range where the Fermi surface passes near six saddle points.

A completely different physical situation arises for the region $1<n_h<2$. Here the system resonates between two groups of magnetic states. One group possesses spin 1 and the other corresponds to spin 1/2, so that for a positive scattering amplitude for excitations with opposite spins the system passes into the ferromagnetic state (see Fig. 1).

When the number of holes differs little from one $(n_h-1)\ll 1$, the right-hand side of Eq. (19) is $\approx 1/9$, while the left-hand side vanishes according to a linear law $\approx(n_h-1)$, so that in this region the system is ferromagnetic at low temperature. As the hole density increases further, the right-hand side of Eq. (19) increases rapidly, which occurs because the Fermi surface approaches the six singular saddle points.

Correspondingly, when the number of holes $n_h>1$, spin ferromagnetism exists in a quite wide density range from 1 up to a critical value as long as the average hole energy at the Fermi surface is negative. For positive excitation energies, the orbital part of the susceptibility increases, but orbital ordering does not arise under any conditions.

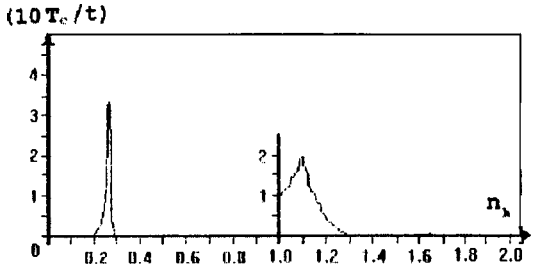

FIG. 1. Dependence of the ferromagnetic transition temperature on the number n_h of holes per unit cell (schematic representation).

The physical picture obtained has a direct relationship to ferromagnetism in hexagonal cobalt. The observed magnetic moment for cobalt is $(1.6-1.75)\mu_B$. According to our estimates, the Fermi level corresponding to the characteristic van Hove singularities corresponds to $n_h \approx 4/3$. The missing difference, $\approx 0.27-0.42$, must be attributed to a correction due to orbital splitting.

It should also be noted that hexagonal samples of nickel are not ferromagnets. This phenomenon likewise has a qualitative explanation, since for $n_h < 1$ ferromagnetic ordering arises only if the Fermi surface passes near saddle-type van Hove singularities.

A qualitatively equivalent situation arises for the phase diagram of fcc cobalt. For a small number of excitations $(n_h - 1 \ll 1)$, the three-dimensional Fermi surface is almost spherical, so that the density of states is proportional the first power of the Fermi momentum. The left-hand side of Eq. (19) is proportional to the first power of the number of excitations $(n_h - 1 \ll 1)$, so that it vanishes much more rapidly than the right-hand side, which for $\gamma_s = 1/3$ and $T = 0$ is proportional to the density of states. Therefore at $T = 0$ the right-hand side (19) is greater than the left-hand side, which corresponds to ferromagnetic instability. As the excitation energy increases, the density of states corresponding to the spectrum (20) increases and then passes through a maximum, so that the Curie temperature depends on the density qualitatively just as for a hexagonal lattice.

In summary, the assumption that the x^2-y^2 and $2xy$ shells of hexagonal cobalt are filled simultaneously leads to a magnetic phase diagram that is qualitatively equivalent to that of fcc cobalt, where the x^2-y^2 and $3z^2-r^2$ shells are filled simultaneously. The existence of van Hove points belonging to the region where the conditions for ferromagnetic instability are satisfied results in an appreciable increase of the ferromagnetic transition temperature for cubic cobalt. This general result agrees with experiment (see Ref. 4).

This study was supported by the Russian Fund for Fundamental Research, Project 98-02-17388.

[1] J. Hubbard, Proc. R. Soc. London, Ser. A **277**, 237 (1964).
[2] J. Goodenough, *Magnetism and the Chemical Bond* (New York, 1963) [Russian translation, Metallurgiya, Moscow, 1988].
[3] R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **112**, 2223 (1997) [JETP **85**, 1218 (1997)].
[4] E. O. Wollan, Phys. Rev. **117**, 387 (1960).

Translated by M. E. Alferieff

Persistent currents and magnetic flux trapping in fragments of carbon deposits containing multiwalled nanotubes

V. I. Tsebro*) and O. E. Omel'yanovskiĭ

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia; International Laboratory of High Magnetic Fields and Low Temperatures, 53421 Wroclaw, Poland

A. P. Moravskiĭ

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

(Submitted 30 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 457–462 (10 October 1999)

It is found that the magnetization curves of samples of fragments of cathode carbon deposits with a high content of multiwalled nanotubes exhibit a pronounced irreversible character, attesting to the induction of persistent currents in the samples and to magnetic flux trapping, as happens in a multiply connected superconducting structure. A decrease of the trapped flux in time could not be observed at low (helium) temperatures with a measurement time of about 20 h. For intermediate ($\sim$30 K) and room temperatures the trapped magnetic flux decays slowly with characteristic relaxation times of the order of 150 and 15 h, respectively. © *1999 American Institute of Physics.* [S0021-3640(99)00819-1]

PACS numbers: 81.05.Ys, 73.23.Ra

1. The electronic properties of carbon nanotubes are the subject of great interest and intensive investigations.[1,2] Notable among recent works are experimental and theoretical works devoted to coherent electron transport in single-walled nanotubes[3–7] and theoretical works[8–10] examining the associated question of circulating, persistent currents in closed toroidal nanotubes.

Transport spectroscopy data[3,4] show that coherent electron transport occurs in single-walled nanotubes at very low temperatures, and it occurs over very large distances, estimated in Ref. 3 to be right up to the total nanotube lengths of several microns. According to the theoretical results obtained in Ref. 5, the conduction electrons in a single-walled nanotube are affected by the disorder averaged over the circumference of the nanotube, and this results in an increase of the electron mean-free path length with increasing nanotube diameter and, in consequence, exceptional ballistic transport properties over unprecedentedly long distances of the order of 10 μm and larger, which is what explains the experimental results. We note that for understandable reasons most theoretical and experimental works on coherent transport concern single-walled nanotubes. For

this reason, the result obtained in Ref. 11 is very noteworthy. It was shown there that the conductance of multiwalled carbon nanotubes, ranging in diameter from 5 to 25 nm and up to 10 μm long, measured at room temperature, is quantized in the sense that it does not depend on the nanotube length or diameter, being equal to $G_0 = 2e^2/h = (12.9\,\mathrm{k}\Omega)^{-1}$. According to Ref. 11, multiwalled carbon nanotubes are capable of carrying at room temperature a current density above 10^7 A/cm^2, which indicates that high-temperature electric transport in such nanotubes is ballistic and occurs without the release of heat.

On this basis, a very important question is the experimental observation of persistent currents in closed nanotube structures. In the present letter we report the results of measurements of magnetization curves of samples of fragments of cathodic carbon deposits, formed during arc synthesis of multiwalled nanotubes. It follows from the data obtained that the carbon medium of such deposits, which consist of contiguous components with different morphology, the main one being the multiwalled nanotubes, is capable of carrying persistent currents at low (liquid-helium) temperatures or very weakly decaying currents even at high (room) temperatures, and the magnetic flux in such a nanotube is trapped, resulting in hysteresis of the magnetization curves, as happens in a multiply connected superconducting structure.

2. Our experimental samples consisted of small fragments, extracted from the central part of the carbon deposits formed on the cathode during arc synthesis of multiwalled nanotubes in the narrow traditional technology used to fabricate such tubes (see, for example, Ref. 1). Ordinarily, such cathode deposits are subjected to special intense treatment (ultrasonic dispersing followed by treatment with strong oxidizers) in order to remove nanoparticles and other carbon formations and to obtain material consisting of essentially multiwalled nanotubes only. We used for our investigations fragments of deposits that were not subjected, after completion of the arc discharge, to any special treatment that destroys the structure of the material.

The qualitative composition and internal structure of the deposits prepared in different technological regimes, including the samples which we prepared for magnetic measurements, have been investigated in detail in Ref. 12. As a rule, their central part possesses a columnar structure, oriented along the growth axis of the deposit. According to Ref. 12, the carbon columns of such a structure consist of three basic components: multiwalled nanotubes ranging in diameter from 5 to 45 nm (the most likely value is $\sim$15 nm), multiwalled polyhedral particles ranging in size from 20 to 90 nm, and curved graphitized formations. Their relative amounts and characteristic sizes are determined by the parameters of the arc process. While all three components are present inside the columns, the outer shell of the columns consists predominantly of only intertwined multiwalled nanotubes. Multiwalled nanotubes in the form of chaotic braids are also present in the space between the columns. Nanotubes from different parts of a deposit are oriented predominantly at obtuse angles with respect to its growth axis.

The samples used for the measurements of the magnetization curves consisted of either (1) carbon columns (average column diameter $\sim$50 μm) obtained from the center of the deposit, assembled and held together with a negligible amount of nonmagnetic glue, and oriented along the z_c axis or (2) small, $\sim$2.5 mm in diameter, bulk cylinders cut from the center of the deposit along the growth axis z_d. In the first variant the samples ranged from one to several milligrams in mass, and they were used for measurements of

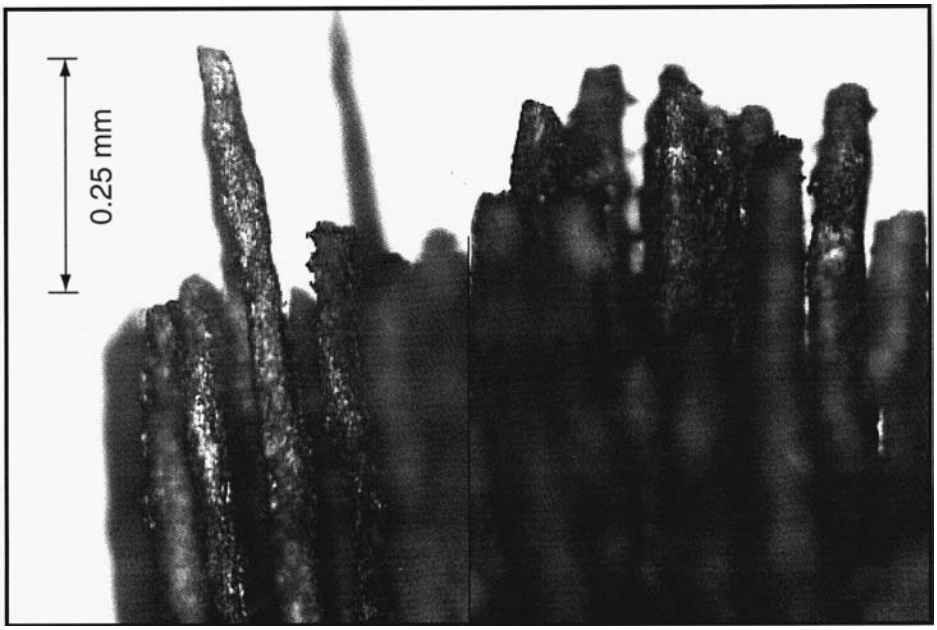

FIG. 1. Optical image of the profile of the end of sample No. 196-1s, consisting of carbon columns assembled together.

the magnetization curves in weak magnetic fields (<500 Oe) using a SQUID magnetometer with a sensitivity of the order of 5×10^{-9} emu with respect to the magnetic moment. Figure 1 shows an optical image of the profile of the end of a sample (No. 196-1s), consisting of carbon columns assembled together. In the second variant the cylindrical samples were of the order of several tens of milligrams in mass, and they were used for magnetic measurements in strong magnetic fields using a self-compensated magnetometer with a capacitance sensor.[13]

3. *Weak magnetic fields.* Figure 2 shows the results of measurements performed on a SQUID magnetometer at $T=4.2$ K of the magnetization curve of sample No. 196-1s

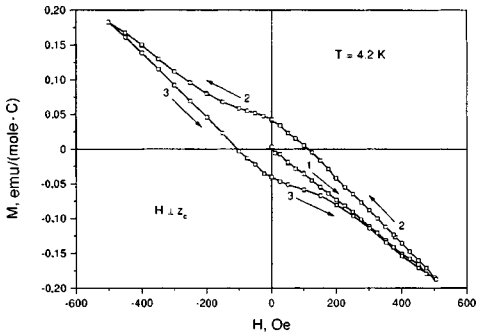

FIG. 2. Hysteresis loop of the magnetization curve of sample No. 196-1s at $T=4.2$ K. The sample mass is 1.65 mg. The magnetic field is directed perpendicular to the axis of the carbon columns.

(sample mass 1.65 mg, $H \perp z_c$), held for a long time at room temperature in a zero (the Earth's) magnetic field. The initial increase of the field in this case always results in a virtually linear dependence $M(H)$ (curve *1*) with slope (magnetic susceptibility) for this sample equal to $\chi = -3.8 \times 10^{-4}$ emu/(mole·C). As the magnetic field decreases (curve *2*), trapped magnetic flux, corresponding in this case to a paramagnetic moment $M_r \approx 0.04$ emu/(mole·C), remains in the sample, and with further cycling of the magnetic field from -500 to $+500$ Oe (curves *2* and *3*) a characteristic hysteresis loop is observed.

Long-time (up to 20 h) observations of M_r at liquid-helium temperature did not show, within the limits of measurement accuracy ($\sim 1\%$), any appreciable decrease of the moment; this indicates that the currents induced in the sample are persistent at low (liquid-helium) temperatures. Measurements of M_r is a function of time and temperature at higher temperatures showed that up to room temperatures M_r depends mainly on the time and not the temperature. Thus, when the sample is heated to intermediate temperatures (~ 20 K), M_r does not change but it already shows appreciable, exponential, relaxation with characteristic relaxation time $\tau_0 \sim 150$ h. When the sample is heated relatively rapidly (in order to eliminate the time factor) up to room temperature, M_r decreases by several percent, and measurements of the time dependences $M_r(t)$ showed that for such high temperatures the relaxation time τ_0 remains quite long, of the order of 15 h. Therefore, in order to return the sample essentially into the initial state (which means, for example, M_r is decreased to a level $<1\%$ of the initial value), the sample must be held at room temperature for approximately three days.

A repeated check of different samples prepared from different parts of the same deposit and from different deposits showed that hysteresis of the magnetization curves occurs in virtually all cases — only its magnitude changes, and in very wide limits from sample to sample. For a number of samples the contribution of the irreversible part of the magnetization was very small, and the hysteresis properties of the $M(H)$ curves could be illustrated satisfactorily only by presenting the difference curves $M(H) - \chi_0 H$ (where χ_0 is the static magnetic susceptibility at the extreme points of the hysteresis loop). In this sense the data presented in Fig. 2 are better, in terms of the magnitude of the effect, than the data that we obtained in weak magnetic fields for samples consisting of carbon columns.

We also note that the magnitude of the hysteresis of the magnetization curves of samples consisting of carbon columns depends very strongly on the direction of the magnetic field relative to the axis z_c of the columns. The magnetic susceptibility itself or the magnetization of the sample is also anisotropic. But the magnetization anisotropy in weak magnetic fields is very small — of the order of several percent — and the magnetic susceptibility is, as a rule, larger for $H \| z_c$, whereas the hysteresis of the $M(H)$ curves changes severalfold as a function of the orientation of the sample, the effect being maximum for $H \perp z_c$. As an example, Fig. 3 shows for one of the samples (No. 196-11) the complete hysteresis loop of the form $M(H) - \chi_0 H$ for $H \perp z_c$ and the initial curves $M(H) - \chi_0 H$ with increasing and decreasing magnetic field with the orientation $H \| z_c$. It is evident that the trapped flux M_r for $H \| z_c$ is approximately four times smaller than for $H \perp z_c$.

4. *Strong magnetic fields.* The measurements were performed at $T = 4.2$ K in the nonuniform field of a superconducting solenoid using a self-balancing magnetometer with a capacitance sensor[13] on cylindrical samples cut from the central part of the deposit

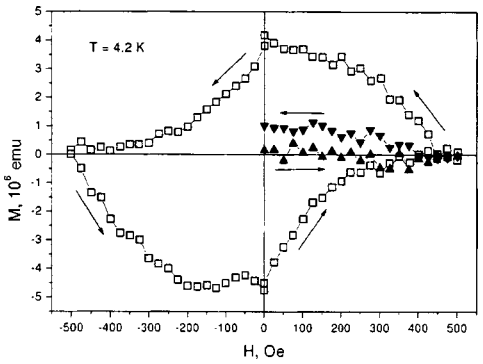

FIG. 3. Hysteresis anisotropy of the magnetization curves for sample No. 196-11 at $T=4.2$ K. The absolute values of the magnetic moment minus $\chi_0 H$ are presented (see text): □ — complete hysteresis loop for $H\perp z_c$, initially increasing (▲) and decreasing (▼) magnetic field for $H\|z_c$.

along the growth axis z_d of the deposit. The magnetic field was not switched, i.e., the $M(H)$ curves were measured with increasing and decreasing fields. Figure 4 shows magnetization curves for one of the samples (No. 140) with increasing and decreasing magnetic field for two orientations: $H\|z_d$ (curve *1*) and $H\perp z_d$ (curve *2*). These data illustrate the general pattern of the results obtained in strong magnetic fields. In the first place, the magnetization curves are strongly nonlinear, which indicates the complicated character of the magnetic interactions in the system. We note that in the process there is a strongly nonlinear field dependence of the magnetization anisotropy, where rapid growth of the ratio $M_\parallel/M_\perp$ in the field range 0–20 kOe is replaced by slow monotonic growth of the ratio up to $M/M_\perp \sim 1.5$ for $H\sim 100$ Oe. In the second place, the magnetization curves are irreversible even in strong magnetic fields. Just as in weak magnetic fields and for carbon column samples, the hysteresis of the $M(H)$ curves is appreciably anisotropic, but the hysteresis is greater for $H\|z_d$ (see Fig. 4). We also note that when the field completely leaves the region of strong magnetic fields, for some samples the residual moment M_r reaches values ~ 2 emu/(mole·C) for $H\|z_d$.

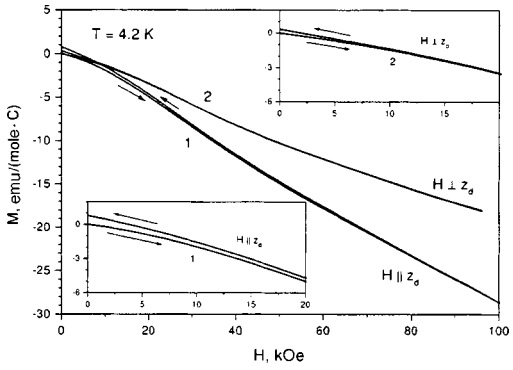

FIG. 4. Magnetization curves for sample No. 140 in strong, increasing and decreasing magnetic fields with orientation along (*1*) and perpendicular (*2*) to the growth axis of the deposit. $T=4.2$ K. The sample mass is 66.6 mg. Inserts: Initial sections of the curves for different orientations of the magnetic field.

5. In summary, we have observed that samples of fragments of cathode carbon deposits, which were not damaged by special treatment in order to remove the multiwalled nanotubes contained in them, can carry persistent magnetic-field induced currents at low temperatures (liquid-helium) or very weakly decaying currents at high (room) temperatures. This property is observed for magnetization curves that show a pronounced irreversible character, i.e., magnetic flux is trapped in the samples, as happens in a multiply connected superconducting structure.

At present one can only surmise how the system of paths conducting persistent or weakly decaying currents in the carbon medium of such samples is organized. It is possible that the structure is similar to a so-called ''Mendelssohn sponge''[14] (a multiply connected system of thin superconductor strands in a normal matrix), and the character of the irreversible behavior of its magnetization corresponds to the well-known critical-state model.[15] If this is so, then the virtually linear dependence $M(H)$ with the field increasing initially (see curve *1* in Fig. 2) indicates that the critical current of the filaments is very high, and the corresponding field-dependent penetration depth of the magnetic field in such a sponge is small. What comprises such a sponge structure is also unclear, but in this case there is a system of interconnected nanotubes in which electric transport is loss-free or the losses are negligibly small. According to electron-microscope data,[12] the structure of the outer shell of the carbon columns of the deposit matches such a sponge structure; such a shell consists of a quite dense network of intertwined and interconnected nanotubes. Since the planes of the cells of such a network are perpendicular to the axis of the columns, the large anisotropy of the hysteresis of the magnetization curves and the fact that the trapped flux is greater in the case $H \perp z_c$ become understandable. A more sparse but also micron-size network of nanotubes is also present in the space between the columns along their entire length.[12] This reticular structure, where the planes of the cells are perpendicular to the growth axis of the deposit, likewise seems to trap magnetic flux well, and its effect becomes determining in bulk samples cut from a deposit, the trapped flux being maximum for $H \perp z_d$. Since, as noted above, the hysteresis of the $M(H)$ curves varies very strongly from sample to sample, it can be inferred that it is the quality of the intertube connections that determines the wholeness of the sponge and the corresponding trapping of the magnetic flux, leading to hysteresis of the magnetization curves. As a rule, appreciable hysteresis of the magnetization curves has always been observed in samples whose magnetic susceptibility at low temperatures was much greater in absolute magnitude than the value adopted for nanotubes $\chi \sim -3 \times 10^{-4}$ emu/(mole·C).[16,17] Apparently, persistent currents already make a large contribution to the diamagnetic response of such samples to an external magnetic field. It has not been ruled out that the proposed current-carrying sponge does not work in the manner that has been supposed, as a whole, in which case it must be assumed that magnetic flux is trapped in individual cells of the nanotube network which are unconnected or weakly connected with one another. Further investigations are required to clarify the nature of the observed persistent currents and the corresponding current-carrying structure.

We thank the State Scientific and Technical Program ''Topical Problems in Condensed-Matter Physics'' for support (''Fullerenes and Atomic Clusters'' No. 2-5-99).

*)e-mail: tsebro@sci.lebedev.ru

[1] T. W. Ebbesen, Phys. Today **49**, No. 6, 26–32 (June 1996).
[2] R. E. Smalley, Rev. Mod. Phys. **69**, 723 (1997).
[3] S. J. Tans, M. H. Devoret, H. Dai *et al.*, Nature (London) **386**, 474 (1997).
[4] M. Bockrath, D. H. Cobden, P. L. McEuen *et al.*, Science **275**, 1922 (1997).
[5] C. T. White and T. N. Todorov, Nature (London) **393**, 240 (1998).
[6] R. Egger and A. O. Gogolin, http://xxx.lanl.gov/abs/cond-mat/9803128.
[7] M. Bockrath, D. H. Cobden, J. Lu *et al.*, Nature (London) **397**, 598 (1999).
[8] R. C. Haddon, Nature (London) **388**, 31 (1997).
[9] M. F. Lin and D. S. Chuu, Phys. Rev. B **57**, 6731 (1998).
[10] A. A. Odintsov, W. Smit, and H. Yoshioka, http://xxx.lanl.gov/abs/cond-mat/9805164.
[11] S. Frank, P. Poncharal, Z. L. Wang, and W. A. de Heer, Science **280**, 1744 (1998).
[12] N. A. Kiselev, A. P. Moravsky, A. B. Ormont, and D. N. Zakharov, Carbon **37**, 1093 (1999).
[13] *Null-Balance Magnetometer with Capacitance Sensor, Internal Specification*, International Laboratory of High Magnetic Fields and Low Temperatures, Wroclaw, Poland, 1998.
[14] K. Mendelssohn, Proc. R. Soc. London, Ser. A **152**, 34 (1935).
[15] C. P. Bean, Phys. Rev. Lett. **8**, 250 (1962).
[16] R. C. Haddon, Nature (London) **378**, 249 (1995).
[17] M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes* (Academic Press, San Diego, California, 1996), p. 836.

Translated by M. E. Alferieff

Optical phonons in nanosize GaAs and AlAs clusters in an InAs matrix

D. A. Ténné,*) V. A. Gaĭsler, A. K. Bakarov, A. I. Toropov, and A. K. Gutakovskiĭ

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

A. P. Shebanin

Joint Institute of Geology, Geophysics, and Mineralogy, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

D. R. T. Zahn

Institut für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

(Submitted 31 August 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 463–467 (10 October 1999)

A Raman scattering method is used to investigate structures containing nanosize GaAs and AlAs clusters, which were grown by molecular-beam epitaxy on InAs substrates by the mechanism of self-organized growth under mechanical stress. A large shift of the phonon lines of GaAs and AlAs clusters with respect to the phonon frequencies in the bulk materials (36 and 24 cm^{-1} for GaAs *LO* and *TO* phonons and 55 and 28 cm^{-1} for AlAs *LO* and *TO* phonons, respectively) is observed in the spectra. This fact is explained by the presence of strong mechanical stresses in the GaAs and AlAs clusters. A comparison of the experimental data with the computed strain dependences of the phonon frequencies shows that the GaAs and AlAs clusters are pseudomorphic, i.e., they do not contain dislocations, which lead to relaxation of the mechanical stresses. In the interval between the InAs *TO* and *LO* phonon frequencies, the Raman scattering spectra contain features associated with interfacial phonons. The position of these features also attests to the formation of three-dimensional GaAs and AlAs islands and are described well by a continuum dielectric model. © *1999 American Institute of Physics.* [S0021-3640(99)00919-6]

PACS numbers: 81.05.Ys, 61.46.+w, 78.30.Fs

Semiconductor structures with quantum dots, obtained by molecular-beam epitaxy in strained heteroepitaxial systems using self-organized growth effects, have been under intensive study in the last few years. The self-organization of nanostructures consists in the fact that when a material is deposited on a lattice-mismatched substrate, under certain conditions spontaneous formation of three-dimensional islands of the deposited material (quantum dots) occurs and the total elastic energy of the heterosystem decreases (Stranskiĭ–Krastanov growth mechanism). Many works have now been published on the

theoretical and experimental investigation of mechanisms of the formation of and the structural, electronic, and optical properties of self-organized quantum dots. The best studied system is a system of InAs quantum dots grown on a GaAs surface (a review is given in Ref. 1). The vibrational properties of structures with quantum dots have been much less studied. Data on the optical phonon spectrum, which were obtained by resonance photoluminescence,[1] and theoretical calculations taking account of the distributions of the mechanical stresses are available for InAs quantum dots in a GaAs matrix.[2] There are also a few works reporting on the study of Raman scattering (RS) of light by phonons in quantum dots, grown in other heteroepitaxial systems (Ge on Si,[3] (In,Ga,Al)Sb on GaAs,[4] InSb on InP,[5] and InAs on InP[6]).

In the present work we investigated the phonon spectrum of nanosize GaAs and AlAs clusters in an InAs matrix. A characteristic feature of a system of GaAs and AlAs clusters on an InAs surface, as compared with other strained heteroepitaxial systems studied previously (InAs on GaAs,[1] Ga on Si,[3,7] (In,Ga,Al)Sb on GaAs,[4] and CdSe on ZnSe[8]), is the lower value of the lattice constant of the cluster material relative to the substrate, as a result of which the elastic strain has the opposite sign (dilitation of the clusters). As far as we know, such nanostructures have not yet been investigated.

Samples of structures with self-organized GaAs/InAs and AlAs/InAs clusters were grown by molecular-beam epitaxy in a Riber-32P system. Growth was carried out on the (001)InAs surface at substrate temperatures of 430–445 °C. The InAs, GaAs, and AlAs growth rates were 0.25, 0.06, and 0.04 monolayers (ML) per second, respectively. The growth process was monitored by RHEED. According to the RHEED data, after deposition of 1.8 ML of GaAs (AlAs) on an InAs surface a transition was observed from layerwise (two-dimensional) to three-dimensional growth, i.e., the onset of island formation. Structures consisting of five layers of GaAs clusters, separated by 10 nm thick InAs layers, were grown. The nominal thicknesses of the GaAs layers were 5, 3.5, and 2 ML for samples *A*, *B*, and *C*, respectively. The top layer of GaAs clusters was covered with a thin (2 nm) layer of InAs. A sample *D*, completely analogous to sample *C* but containing AlAs instead of GaAs clusters, was also grown.

The formation of three-dimensional GaAs and AlAs islands in the grown structures was confirmed by high-resolution transmission electron microscopy. Figure 1 shows an image of sample *D* with atomic resolution. The AlAs islands have average sizes of 3–4 nm in the plane of the layers and 1.5–2 nm in a perpendicular direction. According to the electron microscopy data for a sample similar to sample *B* but containing only one layer of GaAs clusters, the GaAs islands are characterized by larger sizes (greater than 10 nm in the plane of the layers).

The RS spectroscopy method was used to investigate the phonon spectrum of structures with quantum dots experimentally. Light scattering was studied at temperatures 300 and 80 K using 514.5 nm Ar^+-laser light for excitation. The RS spectra were recorded on Jobin Yvon U1000 and Dilor XY800 spectrometers in backscattering geometry.

Figure 2 shows the RS spectra of samples A–D. Together with the InAs LO and TO phonon peaks (242 and 219 cm^{-1}), the spectra of the samples A–C contain the LO and TO phonon lines of GaAs clusters (259 and 247 cm^{-1}). The position of the lines is virtually identical for all samples. No features were observed at the phonon frequencies of bulk (unstrained) GaAs (295 and 271 cm^{-1} for LO and TO phonons). A large shift of

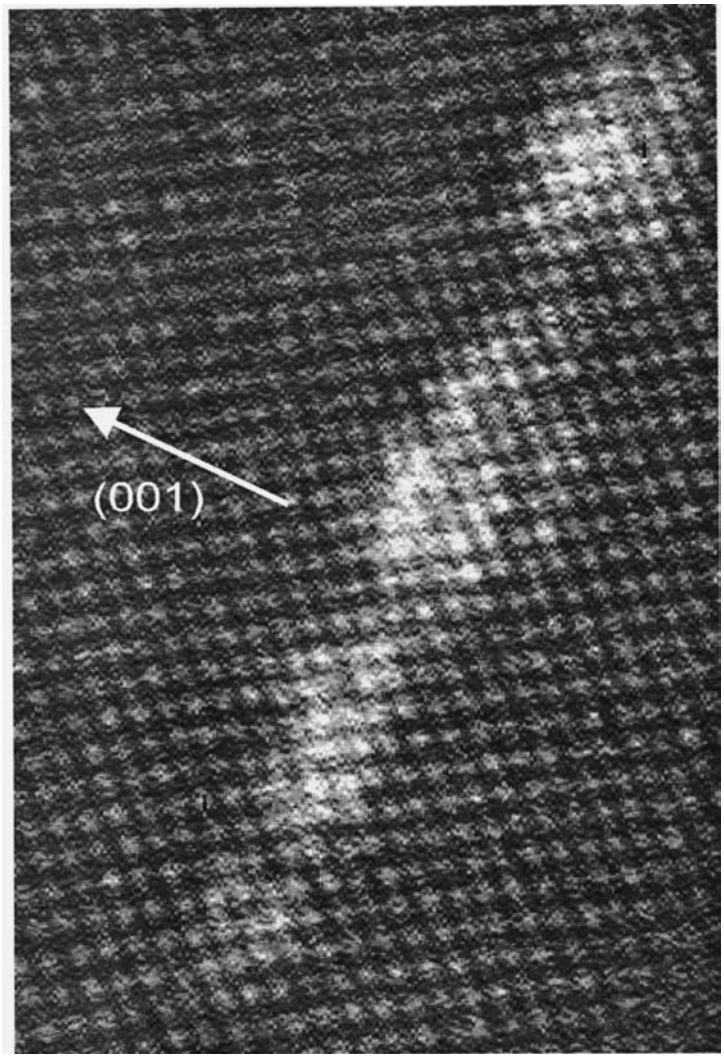

FIG. 1. High-resolution transmission electron microscope image of a transverse section of sample *D*. The light colored regions are AlAs clusters. The arrow indicates the direction of growth of the structure.

the phonon frequencies of AlAs clusters relative to the values for the bulk material also occurs in sample *D*. In the RS spectrum of this sample the *LO* and *TO* phonon lines of the AlAs clusters are observed at 350 and 335 cm^{-1}, while their frequencies in bulk AlAs are 405 and 363 cm^{-1}, respectively. Such large shifts of the phonon frequencies can be explained only by the presence of strong mechanical stresses in the GaAs and AlAs clusters. Since the average size of the GaAs clusters is quite large, localization effects do not strongly influence the phonon frequencies. The AlAs clusters are smaller, but the shift of the phonon frequencies due to localization is small even in this case because of the weak dispersion of the AlAs phonons.[9,10] The shift of the *LO* phonon line observed in the RS spectrum relative to the value for the bulk unstrained AlAs (55 cm^{-1}) is approxi-

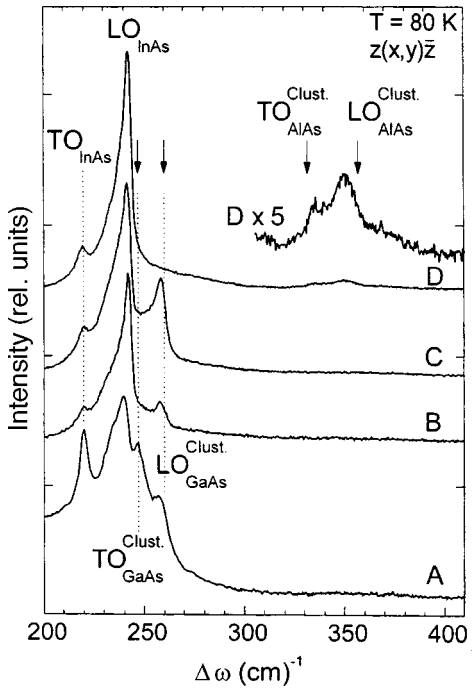

FIG. 2. Raman scattering spectra of samples A–D. The arrows show the phonon frequencies calculated for pseudomorphic GaAs and AlAs layers.

mately 5 times greater than the total dispersion of the AlAs LO phonons in the (001) direction (8–10 cm^{-1}).[9]

The dependence of the frequencies of optical phonons in GaAs on the mechanical stresses is well known. The shifts of the phonon frequencies in strained zinc blende crystals can be calculated by solving the characteristic equation.[11] Specifically, in nanostructures grown in the (001) direction, the phonon frequencies are determined from the equation

$$p\varepsilon_{zz}+q(\varepsilon_{xx}+\varepsilon_{yy})=\omega^2-\omega_0^2 \tag{1}$$

for LO phonons and

$$p\varepsilon_{xx}+q(\varepsilon_{yy}+\varepsilon_{zz})=\omega^2-\omega_0^2 \tag{2}$$

for TO phonons. Here ε_{xx}, ε_{yy}, and ε_{zz} are the diagonal components of the strain tensor, ω and ω_0 are the phonon frequencies in the strained and unstrained material, respectively, and p and q are the phonon deformation potentials, which determine the dependence of the phonon frequencies on the mechanical stresses. For GaAs we used the data of Ref. 12. In Fig. 2 the arrows mark the phonon frequencies of strained GaAs, calculated according to the pseudomorphic layer model described in Refs. 13 and 14: $\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon=0.07$ and $\varepsilon_{zz}=2S_{12}/(S_{11}+S_{12})\varepsilon$, where S_{11} and S_{12} are the elastic compliance constants, taken from Refs. 11 and 15. In real structures the stresses are distributed nonuniformly,[2] but because the phonons are localized over whole clusters, their frequen-

cies are determined by the average values of the stresses. The LO phonon energy, calculated in this manner, for pseudomorphic InAs is 32.3 meV, which is virtually identical to the value obtained in Ref. 2 taking account of the nonuniform stress distribution (32.1 meV). As one can see from Fig. 2, the experimental positions of the LO and TO phonon lines in the RS spectra of GaAs clusters correspond well to the computed frequencies, which indicates that in GaAs there are no dislocations resulting in relaxation of mechanical stresses.

As far as we know, for AlAs at present there are no published data on the phonon deformation potentials. To estimate approximately the positions of the phonon frequencies in the Al and As clusters, we assumed the values to be the same as for GaAs. This assumption is justified, since these parameters can be expressed in terms of the effective interatomic force constants.[11] First-principles theoretical calculations of phonon dispersion[9] show that the force constants of GaAs and AlAs are very close and the phonon spectrum of bulk AlAs is described very well by a calculation based on the force constants of GaAs. The elastic constants of GaAs and AlAs also differ negligibly (by no more than 5%[15]). The values of the phonon frequencies for pseudomorphic AlAs clusters, calculated similarly to the method described above for GaAs, are marked by arrows in Fig. 2. As one can see from the figure, for AlAs good agreement is observed between the computed and experimentally measured phonon frequencies.

In the range of InAs frequencies between the TO and LO phonon lines of bulk InAs, a feature associated with interfacial phonons appears in the RS spectra in the parallel polarization configuration (Fig. 3). In structures with GaAs clusters (samples A–C) this feature lies at 233 cm^{-1}, irrespective of the amount of GaAs deposited, and in sample D, containing AlAs clusters, it is shifted in the high-frequency direction (235.5 cm^{-1}). Interfacial phonons, i.e., vibrational modes localized near the heterointerface and decaying exponentially on both sides away from it, they exist in the frequency ranges of the cluster and the surrounding material. Using the continuum dielectric model,[16] we calculated the frequencies of the first interfacial modes in the InAs range for spherical GaAs and AlAs clusters surrounded by an InAs matrix. The computational results (see Fig. 3) agree well with the experimental data and describe correctly the difference between the position of the lines of the interfacial phonons in structures with GaAs and AlAs clusters. This difference is determined by the difference of the frequencies of the optical phonons and, correspondingly, the dielectric functions of GaAs and AlAs.

It should be noted that in a superlattice with flat faces and a large layer thickness ratio, as in the present samples (10 nm InAs and 0.6–1 nm GaAs or AlAs), the interfacial phonon bands should appear in narrow ranges (1–2 cm^{-1}) near the TO and LO phonon frequencies of the bulk materials.[17,18] For this reason, the appearance of the interfacial phonon lines approximately at the center of the frequency interval between the InAs TO and LO phonons attests to the formation of three-dimensional GaAs (AlAs) clusters.

In summary, in the present work the spectrum of optical phonons of nanosize GaAs and AlAs clusters in an InAs matrix was investigated. According to the calculation of the dependence of the phonon frequencies on the mechanical stresses, the phonon frequency shifts observed in the RS spectra attest to the fact that the GaAs and AlAs clusters do not contain dislocations leading to stress relaxation. In the InAs frequency range, interfacial phonon lines, whose position is described well by the continuum dielectric model, are observed in the spectra.

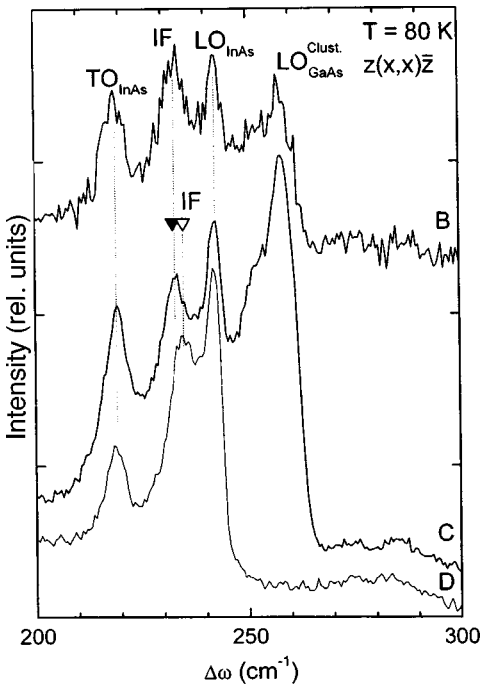

FIG. 3. Raman scattering spectra of samples *B*, *C*, and *D* in parallel polarization geometry. The open and filled triangles show the frequencies of the first interfacial modes of the InAs range for GaAs and AlAs clusters, respectively.

This work was supported in part by the Russian Fund for Fundamental Research (Project No. 98-02-17827) and a Grant from the Siberian Branch of the Russian Academy of Sciences for Young Scientists (D. A. Ténné).

*)e-mail: tenne@thermo.isp.nsc.ru

[1] N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin *et al.*, Fiz. Tekh. Poluprovodn. **32**, 385 (1998) [Semiconductors **32**, 343 (1998)].
[2] M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **42**, 11969 (1995).
[3] J. Groenen, R. Carles, S. Christiansen *et al.*, Appl. Phys. Lett. **71**, 3856 (1997).
[4] B. R. Bennett, B. V. Shanabrook, and R. Magno, Appl. Phys. Lett. **68**, 958 (1996).
[5] G. Armelles, T. Utzmeier, P. A. Postigo *et al.*, J. Appl. Phys. **81**, 6339 (1997).
[6] J. Groenen, A. Mlayah, R. Carles *et al.*, Appl. Phys. Lett. **69**, 943 (1996).
[7] Feng Liu, M. C. Lagally, Surf. Sci. **386**, 169 (1997).
[8] J. L. Merz, S. See, and J. K. Furdyna, J. Cryst. Growth **184/185**, 228 (1998).
[9] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).
[10] V. A. Gaĭsler, D. A. Ténné, N. T. Moshegov *et al.*, JETP Lett. **61**, 376 (1995).
[11] F. Cerdeira, C. J. Buchenauer, F. H. Pollak, and M. Cardona, Phys. Rev. B **5**, 580 (1972).
[12] P. Wickboldt, E. Anastassakis, R. Sauer, and M. Cardona, Phys. Rev. B **35**, 1362 (1987).
[13] E. Anastassakis, in *Light Scattering in Semiconductor Structures and Superlattices*, edited by D. Lockwood and J. F. Young (Plenum, New York, 1991), p. 173.
[14] E. Anastassakis, J. Appl. Phys. **68**, 4561 (1990).
[15] S. Adachi, J. Appl. Phys. **58**, R1-R29 (1985).
[16] P. A. Knipp, and T. L. Reinecke, Phys. Rev. B **46**, 10310 (1992).

[17] B. Jusserand and M. Cardona, in *Light Scattering in Solids V*, edited by M. Cardona and G. Güntherodt (Springer, Heidelberg, 1989), p. 49.
[18] A. K. Sood, J. Menendez, M. Cardona, and K. Ploog, Phys. Rev. Lett. **54**, 2115 (1985).

Translated by M. E. Alferieff

Texture and magnetic anisotropy of carbon nanotubes in cathode deposits obtained by the electric-arc method

A. S. Kotosonov
Scientific-Research Institute of Graphite Materials, 111524 Moscow, Russia

(Submitted 9 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 468–472 (10 October 1999)

The principal values of the diamagnetic susceptibility of nanotubes are found by comparing the texture parameters of carbon multiwalled nanotubes in cathode deposits with their magnetic anisotropy. It is shown that the susceptibility along the nanotubes is essentially identical to the atomic susceptibility of carbon, and the large average diamagnetism of multiwalled nanotubes is due to the susceptibility component normal to the carbon layers, whose value in the temperature range 4–900 K can be completely explained on the basis of the band model of quasi-two-dimensional graphite. © *1999 American Institute of Physics.* [S0021-3640(99)01019-1]

PACS numbers: 81.05.Ys, 61.46.+w

Carbon nanotubes consist of rolled-up ''graphene'' layers and are hollow seamless cylinders with nanometer diameter and up to several microns long.[1] In carbon multiwalled nanotubes (MWNTs) the cylinders are nested coaxially, and because of the steric conditions between neighboring layers there is no three-dimensional order, while the interlayer distance ($\approx$0.342 nm) is greater than for a graphite crystal (0.3354 nm). Ordinarily, the dc electric arc method is used to obtain MWNTs in macroscopic amounts.[2] When cylindrical graphite electrodes and the optimal production conditions are used, the carbon evaporated from the anodes settles in the form of a rod on the cathode; the rod possesses a thin gray outer shell and a dark core, containing primarily MWNTs and a small fraction of ''onion'' and polyhedron type nanoparticles. One characteristic of multiwalled nanotubes which is of practical importance is their high diamagnetic susceptibility (DMS).[3–7] It has been established experimentally[3–6] that the orientationally averaged value of DMS of nanotubes can be much greater than the average susceptibility of graphite, considered until recently the best diamagnet after superconductors. According to the viewpoint commonly expressed in the literature,[3–6,8] the large average diamagnetism of MWNTs is due to the additional contribution of ring currents to the susceptibility, which are induced by the magnetic field component applied along the axis of the nanotube. Moreover, in Refs. 4, 6, and 8–10 it was asserted that the component of the DMS along a nanotube is dominant. At the same time, according to the theoretical estimates,[11,12] the DMS along the axis of the nanotubes should be 1 to 2 orders of magnitude lower than in a direction perpendicular to the axis.

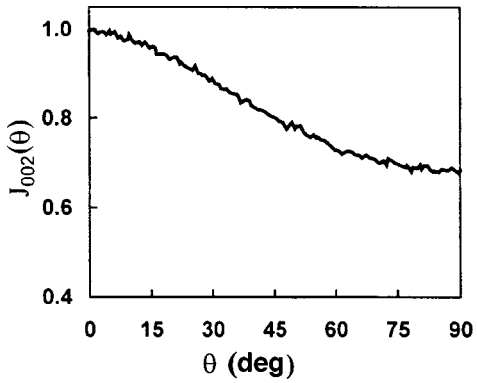

FIG. 1. X-ray textural function.

In our opinion, the confusion concerning the anisotropy of the DMS of nanotubes could be due to the inaccurate estimate of their orientation in macroscopic volumes, on which measurements of DMS were performed. As a rule, the cathode deposits consist of 50–250 μm in diameter columnar structures (clearly observed under an optical microscope) whose axes are oriented along the electrode axis, i.e., along the lines of the electric current. In the literature it is customarily assumed[3–6,13,14] that nanotubes and bunches of nanotubes are oriented predominantly along the columnar formations, i.e., likewise along the axis of the electrodes. In this connection, the susceptibility measurements along ''columns'' were identified with measurements along nanotubes, though quantitative data on the orientation of nanotubes in the deposits usually were not presented. In the present work textural measurements were performed on the deposits obtained in comparison with their magnetic anisotropy, which made it possible to determine more accurately the angular orientation of the tubes in the samples and to find the components of the DMS for individual nanotubes.

The investigations were performed on 2 mm in diameter and 3 mm high rods, cut from the central part of the cathode deposits with an outer diameter of 6–7 mm. According to high-resolution electron microscopy data, the samples selected contained primarily MWNTs with an average diameter of 14 nm and a small fraction of 20–25 nm in diameter polyhedral nanoparticles. The samples possessed axial textural symmetry, and the axis of the texture coincided with the electrode axis. The parameter $\langle \sin^2 \theta \rangle$ was chosen as the characteristic of the texture, where θ is the angle between the z axis of the texture (electrode axis) and the local normals to the carbon layers in MWNTs, while the brackets indicate averaging over the sample. X-ray textural analysis[15] using the most intense line (002) for MWNTs was used to estimate $\langle \sin^2 \theta \rangle$. Measurements of the magnetic susceptibility along and perpendicular to the texture axis were performed in the temperature range 4–900 K with magnetic field intensity 2–9 kOe using electronic microbalances. The measurement error did not exceed 10^{-8} emu/g.

Figure 1 shows the textural curve for the experimental samples. It is known[15] that for such wide curves the orientation function $J_{002}(\theta)$ is essentially independent of the parameters of the (002) line of individual MWNTs, and its intensity is proportional to the mass fraction of the nanotube layers, the normals to which make a local angle θ with the

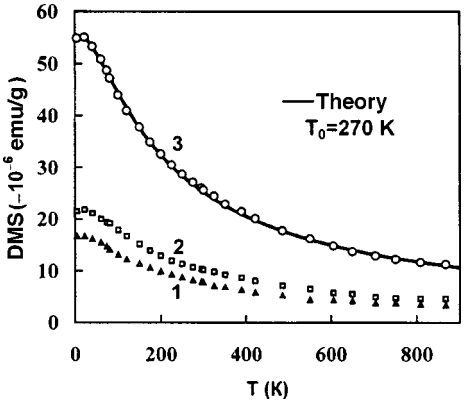

FIG. 2. Temperature dependence of the DMS of a cylinder cut along the axis of the cathode deposit. Dots — experiments; line — calculation in the band model of QTGs with the indicated value of T_0: *1* — $(\chi_y-\chi_a)$, magnetic field perpendicular to the cylinder axis; *2* — $(\chi_z-\chi_a)$, magnetic field parallel to the cylinder axis; *3* — $\chi_{or}=\chi_z+2\chi_y-3\chi_a$.

texture axis. Graphical integration showed that for this sample the average textural parameter $\langle\sin^2\theta\rangle=0.61$. Since the nanotubes are cylinders, the average textural parameter for the axes of the nanotubes themselves is $\langle\sin^2\alpha\rangle=2(1-\langle\sin^2\theta\rangle)=0.78$, where α is the local angle between the texture axis and the axis of an individual nanotube. In the isotropic case $\langle\sin^2\theta\rangle=\langle\sin^2\alpha\rangle=2/3$. Hence it follows that nanotubes in the cathode deposit are oriented predominantly perpendicular to the ''columns'' and the electrode axis, and not along this axis, as usually assumed in the literature. Previously, a similar conclusion about the texture of MWNTs in deposits was drawn on the basis of an analysis of the temperature behavior of their ferromagnetic anisotropy.[16]

The results of susceptibility measurements on a macrosample are presented in Fig. 2. As one can see, χ_z, the DMS along the texture axis, i.e., along columnar formations, in the entire temperature range of measurements is greater in absolute magnitude than χ_y, the DMS in the perpendicular direction. This agrees with measurements performed by other investigators.[3–6] Knowing $\langle\sin^2\theta\rangle$ and the values χ_z and χ_y, the components of the susceptibility along $(\chi_\parallel)$ and perpendicular $(\chi_\perp)$ to the axes of individual nanotubes can be estimated as

$$\chi_\parallel=\chi_y-(\chi_z-\chi_y)\langle\sin^2\theta\rangle/\langle 2-3\sin^2\theta\rangle,\quad \chi_\perp=(\chi_z+2\chi_y-\chi_\parallel)/2. \tag{1}$$

These relations remain even when isotropic nanoparticles of the type ''onions'' and polyhedra are present in the sample, provided that the average values of the susceptibility of the nanoparticles and MWNTs are the same. In order for this condition to be satisfied, additional monitoring of samples with respect to the shape of the ESR signal line was used to select cathodic deposits.[17] This made it possible to choose samples with the most uniform magnetic properties. As a rule, in such samples the average diameter of nanoparticles is 50–70% greater than the diameter of the nanotubes. We note that in this case an increase in the fraction of isotropic particles in the sample is equivalent with respect to their effect on $\langle\sin^2\theta\rangle$, χ_z, and χ_y to an increase in the fraction of MWNTs with equally probable angular orientation.

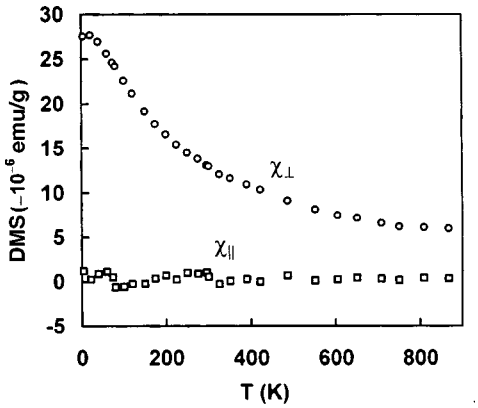

FIG. 3. Temperature dependence of the components $\chi_{\parallel}$ and $\chi_{\perp}$ of the nanotubes.

Figure 3 shows the values of $\chi_{\parallel}$ and $\chi_{\perp}$ obtained for MWNTs using the relations (1) from measurements of χ_z and χ_y on macrovolumes (Fig. 2). It is evident that the DMS along the axis of a nanotube is essentially independent of the temperature, and its value is $\chi_{\parallel}=(-0.3\pm 0.35)\times 10^{-6}$ emu/g, i.e., it is essentially identical to the DMS along the plane of the graphite layers, χ_1, which is close to the atomic susceptibility of carbon $\chi_a=-0.3\times 10^{-6}$ emu/g. The value of the DMS $\chi_{\perp}$ perpendicular to the nanotube axis, depends strongly on temperature and is much greater than $\chi_{\parallel}$. Therefore the large average diamagnetism of multiwalled nanotubes is due to the component $\chi_{\perp}$. At the same time, $\chi_{\perp}=(\chi_3+\chi_1)/2\approx(\chi_3+\chi_a)/2$, where χ_3 is the DMS in a direction perpendicular to the two-dimensional layer. At the same time, the orbital diamagnetism χ_{or} of 2D electrons, which is the interband contribution of current carriers to the observed susceptibility, makes the main contribution to χ_3: $\chi_3=\chi_a+\chi_{\mathrm{or}}$. The quantity χ_{or} determines the high average values of the DMS of a number of carbon materials, including nanotubes.[16] The large value of χ_{or} in graphites and quasi-two-dimensional graphites (QTGs) is due to the characteristic features of their band structure near the contact of filled and empty bands. As it turned out,[16] χ_{or} for MWNTs, just as for QTGs, can be described in a wide temperature range on the basis of the band model of QTGs by the relation

$$\chi_{\mathrm{or}}=\frac{5.45\times 10^{-3}\gamma_0^2}{(T+\delta)(2+\exp(\eta)+\exp(-\eta))}\ \mathrm{emu/g}, \tag{2}$$

where γ_0 is the 2D band parameter (≈ 3 eV); T is the lattice temperature; δ is an additional parameter, which formally takes account of the broadening of the density of states near the Fermi level because of the influence of the potential due to structural defects; $\eta=E_F/k_B(T+\delta)$ is the reduced Fermi level (E_F); and, k_B is Boltzmann's constant. The value of η for any temperature can be found from the electric-neutrality equation on the basis of the same model:

$$F_1(\eta)-F_1(-\eta)=(\eta_0^2/2)\mathrm{sgn}(\eta_0). \tag{3}$$

Here $F_1(\eta)$ is the first-order Fermi integral; $\eta_0=T_0/(T+\delta)$; and, T_0 is the degeneracy temperature of extrinsic current carriers.

The experimental values of χ_{or} for the nanotubes investigated were obtained from measurements of χ_z and χ_y on macrosamples using the relation $\chi_{or}=\chi_3-\chi_a=\chi_z+2\chi_y-3\chi_a$, which is a consequence of the additivity of the DMS. The values obtained for χ_{or} are represented by dots in Fig. 2 (series 3). As one can see, the curve computed using the QTG model fits the experimental data well.[18] The degeneracy temperature T_0 of the extrinsic carriers and the additional parameter δ were used as adjustable parameters. The value of T_0 ($=270$ K) corresponds to the density of extrinsic carriers of the order of 4×10^{10} cm^{-2}, which in QTG is, as a rule, the same as the density of defects in the layers, and in the present case it attests to a quite perfect structure of graphene layers in MWNTs. The best-fit value of δ was found to be $0.5T_0$, which shows that the layered defects are linear in character. For point defects, usually, $\delta\approx0.15T_0$ for QTG.

In summary, by comparing the results of x-ray textural analysis with the magnetic anisotropy of deposits of multiwalled nanotubes it was shown that the large average diamagnetism of nanotubes is due to the component perpendicular to the carbon layers of a nanotube, and it is of an interband nature, just in quasi-two-dimensional graphites. The small value of DMS along a nanotube agrees with the general theoretical estimates in Refs. 9 and 10, and therefore it does not confirm the concept of diamagnetic currents around loops of nanotubes.[3,6–8] The experimental data showing the nanotubes to be oriented predominantly perpendicular to the axis of the deposit are extremely important for clarifying the mechanism of nanotube growth in the arc method for obtaining nanotubes.

I thank A. P. Moravskiĭ, N. A. Kiselev, and D. M. Kheĭker for assisting in the investigations and for a fruitful discussion of the results. This work was partially supported by the Russian Fund for Fundamental Research (Project 99-03-32382) and the program ''Fullerenes and Atomic Clusters'' (Project No. 98030).

[1] S. Iijima, Nature (London) **354**, 56 (1991).
[2] T. W. Ebbesen and P. M. Ajayan, Nature (London) **358**, 220 (1992).
[3] J. Heremans, C. H. Olk, and D. T. Morelli, Phys. Rev. B **49**, 15122 (1994).
[4] X. K. Wang, R. P. H. Chang, A. Patashinski, and J. B. Ketterson, J. Mater. Res. **9**, 1578 (1994).
[5] A. P. Ramirez, R. C. Haddon, O. Zhou *et al.*, Science **265**, 84 (1994).
[6] O. Chauvet, L. Forro, W. Basca *et al.*, Phys. Rev. B **52**, 6963 (1995).
[7] S. Bandow, J. Appl. Phys. **80**, 1020 (1996).
[8] A. A. Ovchinnikov, Phys. Lett. A **195**, 95 (1994).
[9] J. P. Lu, Phys. Rev. Lett. **74**, 1123 (1995).
[10] R. C. Haddon, Nature (London) **378**, 249 (1995).
[11] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **50**, 14698 (1994).
[12] H. Ajiki and T. Ando, J. Phys. Soc. Jpn. **62**, 2470 (1993); **64**, 4382 (1995).
[13] M. Kosaka, T. W. Ebbesen, H. Hiura, and K. Tanigaki, Chem. Phys. Lett. **225**, 161 (1994); **233**, 47 (1995).
[14] T. W. Ebbesen, Phys. Today **49**, No. 6, pp. 26–32 (June 1996).
[15] R. Bacon, in *Chemistry and Physics of Carbon* (Marcel Dekker, New York, 1973), Vol. 9, p. 2.
[16] A. S. Kotosonov and S. V. Kuvshinnikov, Phys. Lett. A **230**, 377 (1997).
[17] A. S. Kotosonov and D. V. Shilo, Carbon **36**, 1649 (1998).
[18] A. S. Kotosonov, JETP Lett. **43**, 37 (1986); Fiz. Tverd. Tela (Leningrad) **33**, 2616 (1991) [Sov. Phys. Solid State **33**, 1477 (1991)].

Translated by M. E. Alferieff

Anomalous temperature behavior of the resistivity in lightly doped manganites around a metal–insulator phase transition

S. Sergeenkov
SUPRAS, Institute of Physics, B5, University of Liège, B-4000 Liège, Belgium; Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

M. Ausloos
SUPRAS, Institute of Physics, B5, University of Liège, B-4000 Liège, Belgium

H. Bougrine
SUPRAS, Institute of Physics, B5, University of Liège, B-4000 Liège, Belgium; SUPRAS, Montefiore Electricity Institute, B28, University of Liège, B-4000 Liège, Belgium

A. Rulmont and R. Cloots
SUPRAS and LCIS, Institute of Chemistry, University of Liège, B-4000 Liège, Belgium

(Submitted 10 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 473–478 (10 October 1999)

An unusual temperature and concentration behavior of the resistivity $\rho(T,x)$ in $La_{0.7}Ca_{0.3}Mn_{1-x}Cu_xO_3$ is observed at slight Cu doping ($0 \leqslant x \leqslant 0.05$). Namely, the introduction of copper results in a splitting of the resistivity maximum around the metal–insulator transition temperature $T_0(x)$ into two differently evolving peaks. Unlike the original Cu-free maximum, which steadily increases with doping, the second (satellite) peak remains virtually unchanged for $x < x_c$, increases for $x \geqslant x_c$, and finally disappears at $x_m \simeq 2x_c$, with $x_c \simeq 0.03$. The observed phenomenon is thought to arise from a competition between substitution-induced strengthening of the potential barriers (which hinder the charge hopping between neighboring Mn sites) and weakening of the kinetic energy of the carriers. The data are well fitted by assuming a nonthermal tunneling conductivity theory with randomly distributed hopping sites. © *1999 American Institute of Physics.* [S0021-3640(99)01119-6]

PACS numbers: 72.60.+g, 75.50.Dd, 75.30.Vn

To clarify the underlying microscopic transport mechanisms in manganites exhibiting colossal magnetoresistance, numerous studies (both experimental and theoretical) have been undertaken during the past few years[1–17] which have revealed a rather intricate correlation of structural, magnetic, and charging properties in these materials based on a crucial role of the Mn^{3+}–O–Mn^{4+} network. In addition to the so-called double-exchange (DE) mechanism (allowing conduction electrons to hop from the singly occu-

pied e_{2g} orbitals of Mn^{3+} ions to empty e_{2g} orbitals of neighboring Mn^{4+} ions), these studies emphasized the important role of the Jahn–Teller (JT) mechanism associated with the distortions of the network's bond angle and length and leading to polaron formation and electron localization in the paramagnetic insulating region. In turn, the onset of ferromagnetism below the Curie point increases the effective bandwidth, simultaneously dissolving the spin polarons into band electrons and rendering the material more metallic. To modify this network, substitution effects on the properties of the most popular $La_{0.7}Ca_{0.3}MnO_3$ manganites have been studied, including the isotopic substitution of oxygen (''giant'' isotope effect[8,9]) and rare-earth (RE)[10–14] and transition-element (TE)[15–17] doping at the Mn site. In particular, an unusually sharp *decrease* of the resistivity $\rho(T)$ in $La_{0.7}Ca_{0.3}\,Mn_{0.96}Cu_{0.04}O_3$ due to just 4% Cu doping has been reported[17] and attributed to the Cu-induced weakening of the carrier kinetic energy $E_0(x)$. On the other hand, the opposite temperature behavior of the resistivity (that is, an *increase* of ρ upon TE doping) can also be expected, based on the deactivation of the DE Zener mechanism. Indeed, this mechanism is effective when electrons can hop (tunnel) between nearest-neighbor TE ions without altering their spin or energy. Hence, the observed[16] lowering of the metal–insulator (M–I) transition temperature and hopping-based conductivity by TE substitution can be ascribed to an inequivalence of the ground-state energies of neighboring Mn and TE ions, resulting in the appearance of a doping-dependent potential barrier $U(x)$. More precisely, this potential energy exceeds the polaron bandwidth (virtually weakening the DE interaction between neighboring TE and Mn ions and thus impeding the possibility of energy-conserving coherent hops) and is defined as the difference between the binding energies of an electron on a TE ion (e.g., Cu) and the Mn ion.

In an attempt to pinpoint the aforementioned potential-energy-controlled hopping mechanism and gain some insight into the doping profile of the barrier, in this letter we present a comparative study of resistivity measurements on Cu-doped polycrystalline manganite samples from the $La_{0.7}Ca_{0.3}\,Mn_{1-x}Cu_xO_3$ family for $0 \leqslant x \leqslant 0.05$ over a wide temperature interval (from 20 K to 300 K). As we shall see, the data are reasonably well fitted (for all T and x) by a unique (nonthermal) tunneling expression for the resistivity which assumes a random (Gaussian) distribution of hopping sites and an explicit form for the temperature- and doping-dependent effective potential $U_{\rm eff}(T,x)=U(x)-E(T,x)$. Besides, the Cu-doping-induced competition between the barrier height profile $U(x)$ and the previously found[17] behavior of the carrier kinetic energy $E_0(x)\equiv E(0,x)$ results in the emergence of a satellite peak in the temperature behavior of the observed resistivity on the insulating side.

The samples examined in this study were prepared by the standard solid-state reaction from stoichiometric amounts of La_2O_3, $CaCO_3$, MnO_2, and CuO powders. The necessary heat treatment was performed in air, in alumina crucibles at 1300 °C for 2 days to preserve the right phase stoichiometry. The powder x-ray diffraction patterns are characteristic of perovskites and show structures that reflect the presence of orthorhombic (or tetragonal) distortions induced by Cu doping. It was confirmed that our data for the undoped samples are compatible with the best results reported by other groups, thus ensuring the quality of our sample processing conditions and procedures.

The electrical resistivity $\rho(T,x)$ was measured using the conventional four-probe method. To avoid the Joule and Peltier effects, a dc current $I=1$ mA was injected (as a

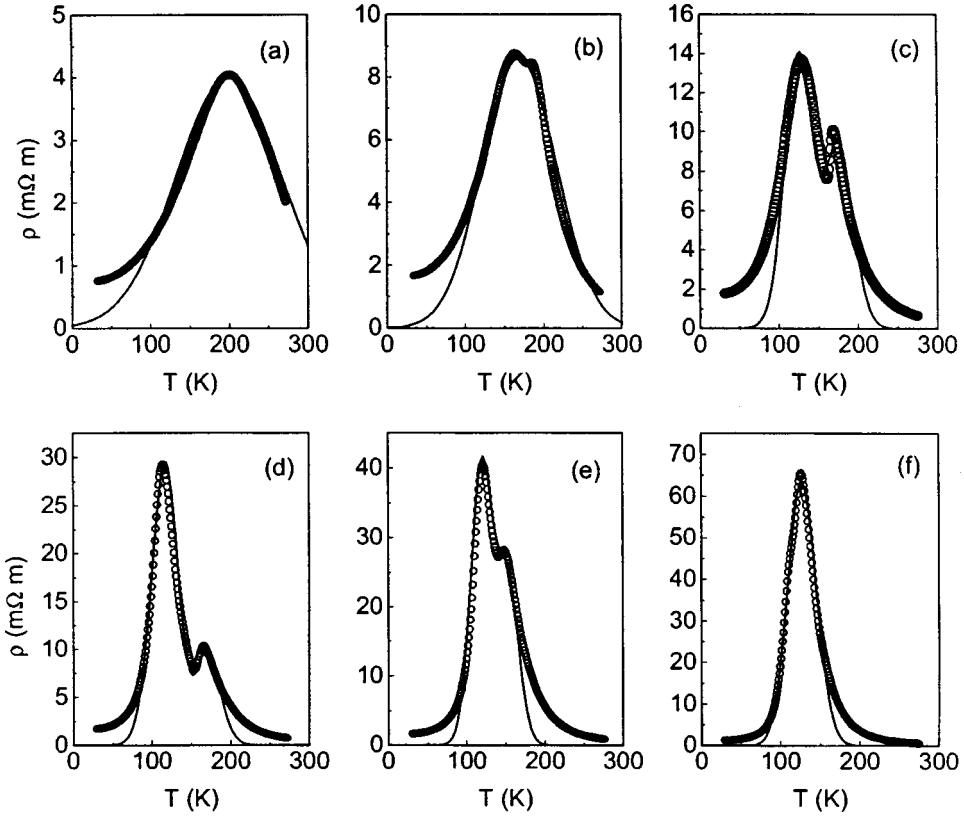

FIG. 1. Temperature behavior of the observed resistivity $\rho(T,x)$ in $La_{0.7}Ca_{0.3}Mn_{1-x}Cu_xO_3$ for different copper content: $x=0$ (a), $x=0.01$ (b), $x=0.02$ (c), $x=0.03$ (d), $x=0.04$ (e), and $x=0.05$ (f). The solid lines are the best fits according to Eqs. (3)–(10).

one-second pulse) successively on both sides of the sample. The voltage drop V across the sample was measured with high accuracy by a KT256 nanovoltmeter. Figure 1 presents the temperature behavior of the resistivity $\rho(T,x)$ for six $La_{0.7}Ca_{0.3}Mn_{1-x}Cu_xO_3$ samples, with $0 \leqslant x \leqslant 0.05$. Notice the rather broad bell-like shape of the resistivity for the undoped sample (Fig. 1a), reaching a maximum at the so-called metal–insulator transition (peak) temperature $T_0(0)=200\,\mathrm{K}$. Upon Cu doping, two markedly different processes occur. First of all, the Cu-free (left) resistivity peak increases and becomes narrower (with $T_0(x)$ shifting towards lower temperatures). Secondly, at a higher temperature another (satellite) peak emerges, splitting off from the original peak. It remains virtually unchanged for small x (up to $x_c \simeq 0.03$) and starts to increase for $x > x_c$ until it finally merges with the main (left) peak at the highest doping level of $x=0.05$.

Due to tangible microstructural changes (observed upon copper doping), the JT mechanism plays a decisive role in the above-described resistivity anomalies by assisting electron localization near the M–I transition temperature. Given the growing experimental evidence[14,15] that polaronic distortions (evident in the paramagnetic state) persist in the ferromagnetic phase as well, we consider the observed resistivity to arise from the

tunneling of small spin polarons through the doping-created potential barriers. According to the conventional picture,[5–7,14,17] the conductivity due to tunneling of a carrier through an effective barrier of height $U_{\rm eff}$ and width R is

$$\sigma=\sigma_h e^{-2R/L}, \tag{1}$$

here $L=h/\sqrt{2mU_{\rm eff}}$ is a characteristic length, where h is Plank's constant and m is an effective carrier mass.

To account for the observed anomalous behavior of the resistivity in our samples, we assume that around the metal–insulator transition, in addition to the Cu-doping-induced slight modification ($x\ll 1$) of the barrier height $U(x)\equiv U_{\rm eff}(T_0,x)\simeq xU_1+(1-x)U_2$ the effective potential $U_{\rm eff}=U(x)-E(T,x)$ will also depend on the temperature via the corresponding dependence of the carrier energy $E(T,x)=h^2/2m\xi^2(T,x)$, with some characteristic length $\xi(T,x)\simeq\xi_0(x)/[1-T/T_0(x)]$ (where $\xi_0(x)\simeq\xi_0(0)/(1-x)^2$)[17] which plays the role of the charge carrier localization length above T_0 (in the insulating phase) and the correlation length below T_0 (in the metallic phase), so that $\xi^{-1}(T_0,x)=0$. Furthermore, given the rather wide temperature dependence of the resistivity for the undoped sample (see Fig. 1a), we adopt the effective-medium approximation scheme and assume a random distribution of hopping distances R with the normalized function $f(R)$, which leads to

$$\rho\equiv\langle\sigma^{-1}\rangle=\frac{1}{Z}\int_0^{R_m}dRf(R)\sigma^{-1}(R), \tag{2}$$

for the resistivity of the effective medium, where $Z=\int_0^{R_m}dRf(R)$, with R_m being the largest hopping distance. In what follows, for simplicity we consider a Gaussian distribution (around a mean value R_0) with the normalized function $f(R)=(2\pi R_0^2)^{-1/2}e^{-R^2/2R_0^2}$ resulting in the following expression for the observed resistivity

$$\rho(T,x)=\rho_h e^{\gamma^2}\left[\frac{\Phi(\gamma)-\Phi(\gamma-\gamma_m)}{\Phi(\gamma_m)}\right], \tag{3}$$

where

$$\gamma(T,x)=\sqrt{\mu(x)-\gamma_0(x)\left[1-\frac{T}{T_0(x)}\right]^2}, \tag{4}$$

with

$$\mu(x)=\frac{2mU(x)R_0^2}{h^2}\equiv\mu(0)+x\Delta\mu, \tag{5}$$

(which measures the substitution-induced potential barriers $U(x)$ hampering the charge hopping between neighboring Mn sites) and

$$\gamma_0(x)=R_0^2/\xi_0^2(x)\simeq\gamma_0(0)(1-x)^4 \tag{6}$$

(which measures the effects due to the carrier kinetic energy $E_0(x)\equiv E(0,x)$, see above). Here $\rho_h=1/\sigma_h$, $\gamma_m=R_m/R_0$, and $\Phi(\gamma)$ is the error function.

Turning to the discussion of the main (left) resistivity profile, we note that the Cu-induced changes of its peak temperature $T_0(x)$ are well fitted by the exponential law

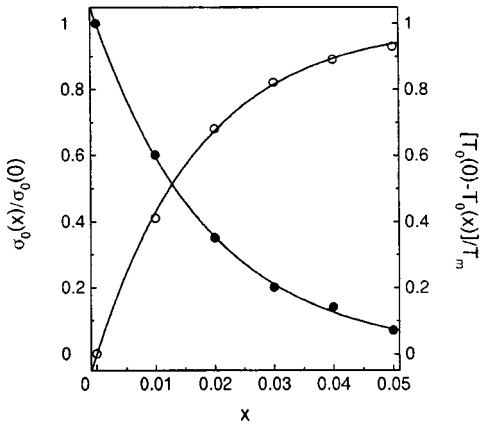

FIG. 2. The dependence of the normalized left peak temperature $[T_0(0)-T_0(x)]/T_m$ (open dots) and conductivity $\sigma_0(x)/\sigma_0(0)$ (solid dots) on copper doping x in $La_{0.7}Ca_{0.3}Mn_{1-x}Cu_xO_3$. The solid lines are the best fits according to Eqs. (7) and (8).

$$T_0(x)=T_0(0)-T_m(1-e^{-x\tau}), \tag{7}$$

with $T_0(0)=200\,\mathrm{K}$, $T_m=73\,\mathrm{K}$ and $\tau=56$. At the same time, according to Eqs. (3)–(7) (and in agreement with the observations, see Fig. 2), the corresponding peak resistivity $\rho_0(x)\equiv\rho(T_0,x)$ increases with x as follows

$$\rho_0(x)=\rho_0(0)e^{x\Delta\mu}, \tag{8}$$

yielding $\rho_0(0)=\rho_h e^{\mu(0)}=4\,\mathrm{m}\Omega\cdot\mathrm{m}$ and $\Delta\mu=54$ for the model parameters and suggesting that $\rho_0(x)\propto 1/T_0(x)$. To further emphasize this similarity, Fig. 2 depicts the extracted doping variation of the normalized quantities, $[T_0(0)-T_0(x)]/T_m$ (open dots) and left peak conductivity $\sigma_0(x)/\sigma_0(0)=\rho_0(0)/\rho_0(x)$ (solid dots) along with the fitting curves (solid lines) according to Eqs. (7) and (8).

A more careful analysis of Eq. (3) shows that in addition to the main peak at $T_0(x)$, the equation $d\rho(T,x)/dT=0$ has two more conjugated extreme points at $T=T_S^{\pm}(x)$, intrinsically linked to the main peak, viz.,

$$T_S^-(x)=T_0(x)\left[1-\sqrt{\frac{\mu(x)-\mu^-}{\gamma_0(x)}}\right] \quad \text{and} \quad T_S^+(x)=T_0(x)\left[1-\sqrt{\frac{\mu^+-\mu(x)}{\gamma_0(x)}}\right]$$

with $\mu^{\pm}=\sqrt{2}(2\pm\gamma_m)$. To attribute these temperatures to the observed satellite (right) peak (see Fig. 1), first of all, we have to satisfy the ''boundary conditions'' at zero ($x=0$) and the highest ($x=x_m=0.05$) doping levels by assuming $T_S^-(0)=T_0(0)$ and $T_S^+(x_m)=T_0(x_m)$, which lead to the following constraints on the model parameters: $\mu^-=\mu(0)$ and $\mu^+=\mu(x_m)$. Secondly, to correctly describe the observed evolution of the satellite peak with copper doping and to introduce a critical concentration parameter x_c into our model, we use the continuity condition $T_S^+(x_c)=T_S^-(x_c)$. As a result, we find that the satellite peak is governed by a unique law over the whole doping interval, with

$$T_S^-(x)=T_0(x)\left[1-\sqrt{\frac{x}{2x_c}}\right], \quad 0\leqslant x<x_c, \tag{9}$$

$$T_S^+(x)=T_0(x)\left[1-\sqrt{1-\frac{x}{2x_c}}\right], \quad x_c\leqslant x\leqslant x_m, \tag{10}$$

where $x_m=2x_c$ with $x_c=\gamma_0(0)/\Delta\mu$. Since according to Eqs. (4)–(10), $\gamma^2(T_S^-)=\mu(0)$ and $\gamma^2(T_S^+)=\mu(0)+2(x-x_c)\Delta\mu$, in good agreement with the observations (see Fig. 1), it follows now from Eq. (3) that indeed the satellite peak shows practically no changes with x (up to $x\simeq x_c$), since $\rho_S^-(x)=\rho_h\exp[\gamma^2(T_S^-)]\simeq\rho_0(0)$, and starts to increase above the threshold (for $x>x_c$) as $\rho_S^+(x)=\rho_0(0)\exp[2(x-x_c)\Delta\mu]$ until it merges completely with the main peak at $x\simeq x_m$. By comparing the above expressions with our experimental data for the resistivity peaks at $x=x_c$ and $x=x_m$, we get $\gamma_0(0)=R_0^2/\xi_0^2(0)\simeq 1.5$, which [along with the value of $\Delta\mu$ extracted above, see Eq. (8)] leads to $x_c=\gamma_0(0)/\Delta\mu=E_0(0)/\Delta U\simeq 0.03$ for the critical concentration of copper, in very good agreement with the observations. As expected, x_c reflects the competition between the carrier kinetic energy and the copper-induced potential barrier. In turn, assuming as usual[5,6,14,17] a value $R_0\simeq 5.5$ Å for the mean value of the hopping distance and using the value of the free-electron mass for m, the above estimates yield $U_2\equiv U(0)\simeq E_0(0)\simeq 0.1$ eV and $U_1\simeq\Delta U\simeq 3$ eV for the barrier height of the undoped and maximally doped samples, respectively.

Finally, given the above explicit dependences for $T_0(x)$ and $T_S^{\pm}(x)$ along with the fixed model parameters, we are able to fit *all* of the resistivity data with a single function $\rho(T,x)$ given by Eq. (3). The solid lines in Fig. 1 are the best fits according to this equation, assuming the nearest-neighbor hopping approximation (with $\gamma_m=1$).

In summary, due to the competition between the copper-modified kinetic energy $E(0,x)$ of the carriers and the potential barriers $U(x)$ between Mn^{3+}–Mn^{4+} dominated hopping sites, a rather unusual ''double-peak'' behavior of the resistivity $\rho(T,x)$ is observed around the metal–insulator transition temperature $T_0(x)$ in $La_{0.7}Ca_{0.3}Mn_{1-x}Cu_xO_3$ at slight Cu doping. The temperature and x dependences of the resistivity are rather well fitted by a coherent (nonthermal) tunneling of charge carriers with heuristic expressions for the effective potential $U_{\rm eff}(T,x)=U(x)-E(T,x)$ and the critical concentration of copper x_c.

Part of this work has been financially supported by the Action de Recherche Concertées (ARC) 94-99/174. S. S. thanks FNRS (Brussels) for some financial support.

[1] H. Y. Hwang, S-W. Cheong, P. G. Radaelli *et al.*, Phys. Rev. Lett. **75**, 914 (1995).
[2] J. Fontcuberta, M. Martinez, A. Seffar *et al.*, Phys. Rev. Lett. **76**, 1122 (1996).
[3] T. T. M. Palstra, A. P. Ramirez, S-W. Cheong *et al.*, Phys. Rev. B **56**, 5104 (1997).
[4] J. Fontcuberta, V. Laukhin, and X. Obradors, Appl. Phys. Lett. **72**, 2607 (1998).
[5] M. Viret, L. Ranno, and J. M. D. Coey, Phys. Rev. B **55**, 8067 (1997).
[6] L. Sheng, D. Y. Xing, D. N. Sheng *et al.*, Phys. Rev. Lett. **79**, 1710 (1997).
[7] L. P. Gor'kov and V. Z. Kresin, JETP Lett. **67**, 985 (1998).
[8] N. A. Babushkina, L. M. Belova, O. Yu. Gorbenko *et al.*, Nature (London) **391**, 159 (1998).
[9] N. A. Babushkina, L. M. Belova, V. I. Ozhogin *et al.*, J. Appl. Phys. **83**, 7369 (1998).
[10] A. M. Balagurov, V. Yu. Pomyakushin, V. L. Aksenov *et al.*, JETP Lett. **67**, 705 (1998).
[11] N. A. Babushkina, L. M. Belova, D. I. Khomskii *et al.*, Phys. Rev. B **59**, 6994 (1999).
[12] A. M. Balagurov, V. Yu. Pomyakushin, D. V. Sheptyakov *et al.*, JETP Lett. **69**, 50 (1999).
[13] A. M. Balagurov, V. Yu. Pomjakushin, D. V. Sheptyakov *et al.*, Phys. Rev. B **60**, 383 (1999).
[14] S. Sergeenkov, H. Bougrine, M. Ausloos *et al.*, JETP Lett. **70**, 141 (1999).

[15] M. Rubinstein, D. J. Gillespie, J. E. Snyder *et al.*, Phys. Rev. B **56**, 5412 (1997).
[16] K. Ghosh, S. B. Ogale, R. Ramesh *et al.*, Phys. Rev. B **59**, 533 (1999).
[17] S. Sergeenkov, H. Bougrine, M. Ausloos *et al.*, JETP Lett. **69**, 858 (1999).

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

Solitons in semiconductor microstructures with a two-dimensional electron gas

A. O. Govorov,*) V. M. Kovalev, and A. V. Chaplik
Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 13 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 7, 479–481 (10 October 1999)

Nonlinear waves in a two-dimensional electronic plasma with metal screening gates are investigated. It is shown that solitons described by the KdV equation exist in such a system. © *1999 American Institute of Physics.* [S0021-3640(99)01219-0]

PACS numbers: 73.40.Kp, 73.20.Dx

There have been many theoretical[1,2] and experimental[3–5] works on linear plasma waves in a two-dimensional (2D) electron plasma. In the nonlinear regime, the spectral characteristics of the waves start to depend on the wave amplitude. In our opinion, in application to a 2D electron plasma this question has not yet been fully elucidated in the literature. In recent work,[6] solitons are investigated in edge magnetoplasmons (i.e., actually in one-dimensional formations), and a model approach based on a local approximation and on the incompressible liquid approximation is used.

In the present letter, we investigate one-dimensional solitons in a 2D plasma with two screening gates, using the hydrodynamic approximation and a realistic model of a microstructure based on a GaAs/AlGaAs system. The model of the system under study is shown in Fig. 1. Waves propagate in the positive direction along the x axis. The fields in the 2D electron plasma are screened by the top and back gates. The top gate is ordinarily made of metal, and the back gate often consists of a strongly doped GaAs layer.[7] A nonlinear perturbation in the form of a solitary wave in a 2D plasma can be produced by applying a voltage pulse to an isolated gate, shown at the top left in the figure. The solution of the Poisson equation for the region bounded by the metal surfaces at $z=-d_1$ and $z=d_2$ gives for the Fourier component of the potential in the plane of a 2D plasma $(z=0)$

$$\phi_k(z=0)=\frac{4\pi e\tilde{n}_k}{|k|\epsilon[\coth(|k|d_1)+\coth(|k|d_2)]}, \tag{1}$$

where $\tilde{n}_k$ is the Fourier component of the perturbation of the charge density, ϵ is the permittivity, and d_1 and d_2 are the barrier widths. The Fourier transform was applied along the x coordinate. In the long-wavelength approximation (kd_1, $kd_2 \ll 1$) Eq. (1) leads to the following expression for the electric field $E_\parallel=-\partial\phi(z=0)/\partial x$ in the plane of the system:

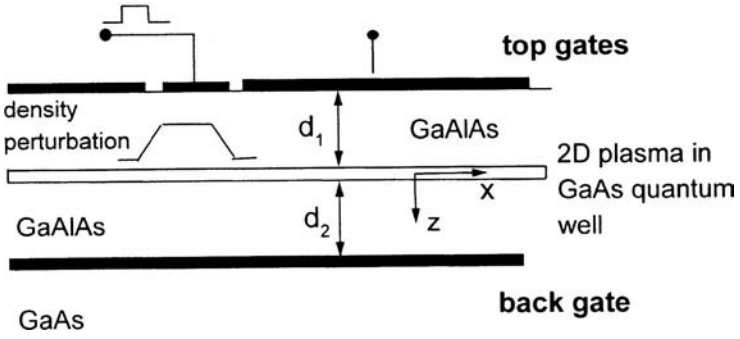

FIG. 1.

$$E_{\|}=-\frac{4\pi e\overline{d}}{\epsilon}\frac{\partial\tilde{n}(x,t)}{\partial x}-\frac{4\pi e d\overline{d^2}}{3\epsilon}\frac{\partial^3\tilde{n}(x,t)}{\partial x^3}. \tag{2}$$

Here $d=d_1+d_2$, $1/\overline{d}=1/d_1+1/d_2$, and t is the time. In Eq. (2) we retained only the first two terms in the expansion in the parameters $d_{1(2)}/L$, where L is the characteristic lateral size of the soliton. In addition, we neglected delay effects. The equations of nonlinear hydrodynamics of a 2D plasma have the form

$$m\frac{\partial v}{\partial t}+mv\frac{\partial v}{\partial x}=eE_{\|}-\frac{m\gamma}{N_0}\frac{\partial n}{\partial x}, \tag{3}$$

$$e\frac{\partial n}{\partial t}+\frac{\partial}{\partial x}env=0, \tag{4}$$

where $v(x,t)$ is the velocity of the plasma, m is the electron mass, $\gamma=v_F^2/2$, and v_F is the Fermi velocity. The electron density is $n=\tilde{n}+N_0$, where N_0 is the equilibrium density.

In the linear approximation ($\tilde{n}\ll N_0$), taking account of only the first term in expression (2), the solution of the system of equations (2)–(4) for a wave traveling in the positive direction along the x axis depends only on the coordinate $\xi=x-v_pt$, where $v_p=\sqrt{4\pi e^2N_0\overline{d}/\epsilon m+v_F^2/2}$. Therefore $v(x,t)=v_p\tilde{n}(x,t)/N_0$ for a solitary wave in this approximation. This follows from Eq. (4) after integration. Now we introduce the new coordinates ξ and $\tau=t$ in Eqs. (2)–(4). We shall assume the τ dependence to be weak, since it arises only when the nonlinearity and dispersion of the plasmon velocity are taken into account. Then combining Eqs. (2)–(4) and integrating once over ξ we arrive at a Korteweg–de Vries (KdV) type of equation[8]

$$\frac{\partial\tilde{n}}{\partial\tau}+\frac{3}{2}\frac{v_p}{N_0}\tilde{n}\frac{\partial\tilde{n}}{\partial\xi}+\frac{v_p\overline{d^2}}{6}\frac{\partial^3\tilde{n}}{\partial\xi^3}=0. \tag{5}$$

This equation was obtained for $\tilde{n}\ll N_0$, $L\gg d_{1(2)}$ and $4\pi e^2N_0\overline{d}/\epsilon m\gg v_F^2$. The equation (5) possesses soliton-like solutions that depend on the variable $\eta=\xi-u\tau$, where $u=v_p(\tilde{n}_{\max}/2N_0)$.[9] The soliton velocity is $v=v_p(1+\tilde{n}_{\max}/2N_0)$. The shape of the soliton is

$$\tilde{n}(\eta)=\tilde{n}_{\max}\cosh^{-2}\left(\frac{\eta}{2\overline{d}}\sqrt{\frac{3\tilde{n}_{\max}}{N_0}}\right) \quad \text{(Ref. 9)}.$$

We now give estimates for the characteristics of the soliton. For typical parameters of the system $N_0=10^{12}$ cm^{-2}, $d_1=d_2=2000$ Å , $\epsilon=12.5$, and $m=0.07m_0$ we obtain the velocity $v_p=1.9\times10^8$ cm/s. The characteristic size of the soliton for $\tilde{n}_{\max}=0.1N_0$ is

$$L\sim\frac{2}{\sqrt{3}}d_1\sqrt{\frac{N_0}{\tilde{n}_{\max}}}\sim 7000\ \text{Å}\,.$$

Formation of a soliton is possible only in plasmas with high mobility, where the soliton formation time is much shorter than the transport relaxation time τ_t of the plasma. The characteristic formation time of a stable soliton shape is determined by the second or third term in Eq. (5). For the parameters presented above, it can be estimated as $\tau_1\sim(N_0/n_{\max})^{3/2}d_1/v_p\sim3\times10^{-12}$ s. The condition $\tau_1\ll\tau_t$ can be easily realized in practice. For example, for a two-dimensional plasma with mobility $\mu\sim3\times10^6$ cm^2/V·s the time $\tau_t\sim10^{-10}$ s. It should be possible to observe a soliton regime in a 2D plasma in time resolved experiments. By recording the transit time of a pulse between the exciting and receiving electrodes it is possible to check the dependence of the velocity of the pulse on $\tilde{n}_{\max}$, i.e., ultimately, on the amplitude of the voltage pulse applied to the exciting electrode (in the linear regime the plasmon velocity does not depend on the wave amplitude). For typical lateral sizes of the samples ($\sim$0.3 cm) the time intervals are of the order of several nanoseconds.

This work was supported by the Russian Fund for Fundamental Research (Projects Nos. 99-02-17019 and 99-02-17127) and the program ''Physics of Solid-State Nanostructures.''

*)e-mail: govor@isp.nsc.ru

[1] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, No. 2, 437 (1982).
[2] A. V. Chaplik and M. V. Krasheninnikov, Surf. Sci. **98**, 533 (1980).
[3] C. C. Grimes and G. Adams, Phys. Rev. Lett. **36**, 145 (1976).
[4] S. J. Allen Jr., D. C. Tsui, and R. A. Logan, Phys. Rev. Lett. **38**, 980 (1977).
[5] T. N. Theis, J. P. Kotthaus, and P. J. Stiles, Solid State Commun. **24**, 273 (1977); Surf. Sci. **73**, 434 (1978).
[6] C. Wexler and A. T. Dorsey, Phys. Rev. Lett. **82**, 620 (1999).
[7] W. R. Frank, A. O. Govorov, J. P. Kotthaus *et al.*, Phys. Rev. B **55**, R1950 (1997).
[8] D. A. Dahl and L. J. Sham, Phys. Rev. B **16**, 651 (1977).
[9] E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Pergamon Press, New York, 1981) [Russian original, Nauka, Moscow, 1979].

Translated by M. E. Alferieff