Detection of very high energy cosmic rays by the UTR-2 decameter telescope

A. D. Filonenko*)
East Ukrainian State University 348034 Lugansk, Ukraine

(Submitted 27 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 639–641 (25 November 1999)

The author believes that the method of radio detection of 10^{22}–10^{23} eV cosmic rays can be revived. The proof will be the detection of electromagnetic pulses produced by the passage of particles with such energy through the surface of the moon. It is shown that the amplitude of the radio signal attesting to this event will be two orders of magnitude greater than the gallactic noise, if the antenna of the UTR-2 decameter radio telescope (Khar'kov) is used for this purpose. © *1999 American Institute of Physics.* [S0021-3640(99)00122-X]

PACS numbers: 95.55.Vj, 95.55.Jz, 95.85.Ry

Radio detection of cosmic rays, which actually has never been implemented, has a decisive advantage over conventional methods of detection in the energy range 10^{22} eV and above. This conclusion follows from the discussion below and, specifically, from the author's previous work. The main content of this method is due to the possibility of monitoring relatively simply an area greater than 10^6 km^2. For example, for the energy range 10^{22}–10^{23} eV considered here, the working area of the detector with the minimum required frequency of detection of events (~ 10 yr^{-1}) should be of the order of 10^7 km^2. Evidently, under winter conditions such a design based on conventional methods of detection appears to be essentially unrealizable.

In this connection, the idea of using the moon's surface ($S_m \sim 10^7$ km^2) as the working area of a detector has stimulated the development of the method of radio detection of very high energy cosmic rays. This possibility was examined in Ref. 1, and later more detailed assessments were made in Ref. 2. However, this principle of detection has substantial drawbacks. In the first place, the characteristic features of the radio-emission mechanism limit by at least an order of magnitude the working area of the moon's surface, i.e., only a narrow annular zone is used effectively (see Ref. 2). In the second place (and this is most important), the frequency range for which Cherenkov radiation remains coherent is several orders of magnitude too high because the scattering of the shower electrons by neutral atoms in the medium is neglected (see, for example, Ref. 3) and the breakdown of coherence with longitudinal development of a shower is also neglected. These drawbacks (especially taken together) decrease the efficiency of such a detector by several orders of magnitude.

In the present letter a detection method that is free of the drawbacks listed above is examined. Specifically, this is due to the fact that the radio-emission component from a

shower that is not associated with Cherenkov radiation is selected for particle detection. As follows from Refs. 4–6, this component is due to the bremsstrahlung of δ electrons from the electron–photon shower in gaseous or condensed media. According to Refs. 4–6, the coherence range in this case lies in the wavelength range $2L<\lambda<\infty$, where L is the effective longitudinal size of a shower, and the spatial distribution of the emission is essentially identical to the directional pattern of a half-wave dipole. The latter property is one of the decisive factors in selecting a principle of radio detection of cosmic rays.

Since the effective length of an electron–photon shower in lunar soil is 5–10 m, the maximum of the spectral intensity of the radio emission lies, as follows from the remarks made above, in the range 30–50 MHz.[7] In this connection, the present author believes that it is in principle possible to use the wide-band antenna of the UTR-2 decameter radio telescope (Khar'kov) to detect cosmic rays with energy $W_0 \sim 10^{22}-10^{23}$ eV. The effective area of the antenna at 30–32 MHz is of the order of 10^5 m^2, and the direction of the principal lobe of the spatial pattern of the radiation varies smoothly in a $\pm 7°$ range from the vertical direction. These characteristics satisfy the conditions of a real arrangement of an experiment as well as is possible.

It is known that at ~ 30 MHz gallactic radio noise is the main source of noise. Atmospheric noise is substantial only during a local thunderstorm, so that, for example, it is completely absent at night during the winter because the ionosphere is transparent for electromagnetic waves in this range. Let us compare, in order of magnitude, the spectral powers of the useful signal and the gallactic radio noise. For this, we estimate the field intensity in the frequency range $2L \approx \lambda$, for example, using the expression (6) from Ref. 8 or (10) from Ref. 4:

$$E_\omega = \frac{LN_m q\nu \sin\varphi_1}{4\pi\epsilon_0 c^2 R_0 f(\beta)} \ln\frac{\gamma_1}{\gamma_m} (\text{V/mHz}), \tag{1}$$

where $L=5$ m is the longitudinal size of the shower, $N_m = 10^{14}$ is the number of electrons at the maximum of the shower for initial particle energy $W_0 = 10^{23}$ eV, $\nu = 3\times 10^7$ Hz, $\varphi_1 = \pi/2$ is the emission angle, $(4\pi\epsilon_0)^{-1} = 0.9\times 10^{10}$ m/F; $c = 3\times 10^8$ m/s, $R_0 = 0.4 \times 10^9$ m is the distance to the moon, $\gamma_1/\gamma_m \approx 50$ is the maximum to minimum ratio of the Lorentz factor, $f(\beta) = 30$ (see Ref. 8) and $q = 1.6\times 10^{-19}$ C. Calculations give approximately $E_\omega = 0.5\times 10^{-12}$ V/mHz.

Then the ratio of the power $J_s = s_0 c E_\omega^2 \pi/2\tau$ of the useful signal arriving in a unit frequency band per 1 m^2 to the intensity of gallactic radio noise $J_n = 3\times 10^{-40} T\nu^2 \Delta\Omega$ (W$\cdot$m$^{-2}\cdot$Hz^{-1}) will have the form

$$\frac{J_s}{J_n} = \frac{\epsilon_0 c E_\omega^2}{3\times 10^{-40} T\nu^2 \tau \Delta\Omega} = 1.3\times 10^5, \tag{2}$$

where $T = 2\times 10^4$ is the characteristic temperature of the celestial sphere at 30 MHz, $\tau = 2\times 10^{-8}$ s is the period of the oscillations, and $\Delta\Omega = 0.8\times 10^{-4}$ sr is the solid angle of the main lobe of the directional pattern of the antenna. The ratio (2) shows that the maximum signal amplitude U_s (provided that the arrival direction of the particles is favorable)

$$U_s = (s_0 c E_\omega^2 \Delta\nu A R_A/\tau)^{1/2}, \tag{3}$$

where $A=10^5$ m^2 is the effective area of the antenna and $R_A=75\ \Omega$ is the matched resistance of the antenna load, is approximately two orders of magnitude greater than the rms amplitude of the radio noise. Substituting the corresponding quantities from Eq. (1) into Eq. (3) gives $U_s\approx 0.7\times 10^{-3}$ V.

Besides radio noise, local noise of a technical character could also be important. However, this is not fundamental and can be eliminated by means of well-known methods.

In summary, estimates show that the radio signal caused by the passage of a 10^{23} eV cosmic particle through the moon's surface can be successfully detected on the earth's surface. Moreover, the estimates (2) and (3) give grounds for believing that 10^{22} eV particles can be detected. In this case the frequency of the events increases by 1.5–2 orders of magnitude.

In conclusion, it should be noted that the successful completion of such an experiment (or a similar experiment[9]) not only opens up new prospects for the method of radio detection using artificial lunar satellites.[10] The experiment does not involve large expenditures on equipment, but even this experiment can give estimates of the cosmic wave flux intensity in the range above 10^{22} eV. In addition, since information arrives only from one side of the celestial sphere, the anisotropy of cosmic rays in this energy range can also be estimated on the basis of the variation of the yearly intensity. Finally, the observation of 10^{22}–10^{23} eV particles will in itself pose a new problem associated with the hypothesis of the relic cutoff of the energy spectrum of cosmic rays.

*)e-mail: uni@vugu.lugansk.ua

[1] G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **41**, 616 (1961) [Sov. Phys. JETP **14**, 441 (1962)].
[2] R. D. Dagkesamanskiĭ and I. M. Zheleznykh, JETP Lett. **50**, 259 (1989).
[3] S. V. Belen'kiĭ, *Shower Processes in Cosmic Rays* [Gostekhizdat, Moscow, 1948], p. 243.
[4] P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20**(12), 57 (1994) [Tech. Phys. Lett. **20**, 499 (1994)].
[5] P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20**(23), 59 (1994) [Tech. Phys. Lett. **20**, 960 (1994)].
[6] P. I. Golubnichiĭ, A. D. Filonenko, and V. I. Yakovlev, Izv. Ross. Akad. Nauk **58**, 115 (1994).
[7] A. D. Filonenko, Akad. Nuak, Ser. Fiz. **61**, No. 3 (1997).
[8] P. I. Golubnichiĭ and A. D. Filonenko, Ukr. Fiz. Zh. (Russ. Ed.) **41**, 696 (1996).
[9] A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **24**(24), 65 (1998) [Tech. Phys. Lett. **24**, 975 (1998)].
[10] A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **23**(10), 57 (1997) [Tech. Phys. Lett. **23**, 399 (1998)].

Translated by M. E. Alferieff

Cosmological γ-ray bursts from a neutron star collapse induced by a primordial black hole

E. V. Derishev,*) V. V. Kocharovsky,†) and Vl. V. Kocharovsky‡)
Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia

(Submitted 23 March 1999; resubmitted 30 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 642–647 (25 November 1999)

A detailed analysis is presented for a novel scenario in which gamma-ray bursts are of intergalactic origin and arise from the induced collapse of an isolated neutron star triggered by a primordial black hole. The energy released from the phase transition of accreted nucleon matter into a quark–gluon plasma is transferred by degenerate neutrinos to the star's surface, where neutrinos annihilate into an electron–positron plasma and produce an inverted temperature layer that preserves a fireball from undue baryonic pollution. Possible observational tests include the absence of apparent cosmological time dilation, the location of γ-ray bursts primarily outside of galaxies, a specific shape of the $\log N$–$\log S$ curve, with a large peak near red shift $z\sim 10$, the emission of $\sim 10^{-3}$ of the total energy in the form of 100-GeV photons, a bimodal distribution of durations, a very weak accompanying pulse of gravitational radiation, etc. © *1999 American Institute of Physics.* [S0021-3640(99)00222-4]

PACS numbers: 97.60.Jd, 97.60.Lf, 95.30.Cq, 98.70.Rz

INTRODUCTION

The problem of γ-ray bursts (GRBs) (see, e.g., Refs. 1 and 2) remains a challenging and controversial issue of modern astrophysics, although a general agreement that they are of cosmological origin seems to be emerging. If so, the burst rate for the whole Universe is at least 800 events per year (only one-third of them is observed). Each GRB lasts for 0.1–100 s ($\sim$10 s on average) and has a broad spectrum with the maximum energy emitted in the 200 keV region.

Many suggestions have been made to explain the subsecond release of a huge amount of energy $\gtrsim 10^{51}$ ergs in the proper spectral range (e.g., Refs. 3 and 4). It should be noted that for two bursts, GRB971214 and GRB990123, the estimated value of the energy release approaches 10^{54} ergs, assuming isotropic emission. Since it is very divergent from the average value ($10^{51}-10^{52}$ ergs) inferred from the analysis of the $\log N$–$\log S$ curve,[5] these bursts may be of different origin.

At present, most theories involve stellar progenitors for GRBs, and there are intrinsic difficulties in scenarios of this kind.[6] The most subtle one is the inevitable baryonic pollution of a hot electron–positron (e^-e^+) relativistic wind (fireball), whereas the lower

limit on the final Lorentz factor of the plasma outflow (equal to the inverse initial fraction of the baryons' rest energy to the energy of the thermal plasma) is $\Gamma \gtrsim 100$.[7]

In the standard fireball models there is another problem, related to the presence of a powerful neutron stream.[8] The latter makes it impossible to have bursts shorter than the apparent lifetime of neutrons, ~ 5 s in the usual models. To eliminate this problem one needs a fireball with a very high Lorentz factor $\sim 10^3$, allowing velocity decoupling between the proton and neutron flows.

Stellar cosmological models place bursters inside galaxies, and this engenders two more problems: few bright galaxies are observed in the error boxes of bright GRBs,[9] and a null result has been obtained in the search for cosmological time dilation in the subsets of bright and dim bursts.[10] Both problems could be avoided if the sources are predominantly located in intergalactic space. In this case the expansion of the Universe affects not only the observed GRB duration but also the shock deceleration time in the local frame (due to the time dependence of the intergalactic gas density). The two effects follow opposing laws inversely proportional to each other, so that no apparent time dilation is expected.

In this paper we propose a novel scenario for cosmological GRBs, which can resolve these difficulties. It is based on the collapse of an isolated neutron star (NS) interacting with a primordial black hole (PBH). PBHs, another great enigma of contemporary cosmology, are relics of the early Universe, and their existence is related to large metric and density perturbations at that stage. The initial mass distribution of PBHs depends on the spectrum of primordial fluctuations,[11] and in the scale-invariant Harrison–Zeldovich case $dn/dm \propto m^{-5/2}$. Those black holes which were lighter than $m_* \simeq 5 \times 10^{14}$g became extinct due to quantum evaporation. Therefore PBHs having masses around m_* are the most abundant now, but their total number is unknown; the best upper limit on the average density, derived from studying the cosmic ray background,[12] is $n_0 < 10^4\ \mathrm{pc}^{-3}$. The MHD plasma regime of PBH evaporation leads to GRBs, but they constitute a small fraction of all bursts.[13]

The new scenario not only explains GRBs but also provides a link between them and PBHs. We expect that it will become a sensitive indicator for the existence of PBHs, 3–5 orders of magnitude more powerful than the present-day tests.

Although we have investigated the model thoroughly, it cannot be presented in detail in a short letter. We therefore give a short description of the whole scenario and focus on the most essential physical point — the neutrino transfer from the quark–nucleon interface. In conclusion we propose accessible observational tests that are capable of distinguishing an actual GRB model.

THE INDUCED COLLAPSE SCENERIO

During contraction of a protostellar cloud, some PBHs along with other dark matter species are gravitationally captured in the cloud, and an appreciable number of PBHs reside inside a newly born star. After a supernova explosion, PBHs with orbits close to the collapsed core remain bound to it.[6] The lifetime of such orbits is limited by gravitational emission and tidal deceleration, and we find that there are enough PBHs that will fall inside NSs within the age of the Universe.

The efficiency of the capturing process depends crucially on the dispersion of PBH

velocities,[14] which took the smallest value just before galaxies were formed. Given the protostellar ($\rho_* \sim 10^{-23}\,\mathrm{g/cm^3}$) and presupernova core ($\rho_c \sim 10^8\,\mathrm{g/cm^3}$) densities, we estimate the probability that a NS becomes a GRB source:

$$P_{\mathrm{GRB}} \simeq 0.3 \left(\frac{R_i}{R_0}\right)\left(\frac{R_i}{R_*}\right)^3 \frac{\Omega_{\mathrm{PBH}}}{\Omega_b}\,\frac{M_{\mathrm{NS}}}{m_*}. \tag{1}$$

Here $R_i = R_c(\rho_c/\rho_*)^{1/4}$ is the radius of a PBH orbit just after trapping in the protostellar cloud, $R_c \simeq 3\times 10^8$ cm is the radius of the core, R_* is the radius of the part of the protostellar cloud with $M \simeq 1.5 M_\odot$, and Ω_b and Ω_{PBH} are the mass fractions of baryons and PBHs. The probability must be a few percent to explain the present-day GRB rate (assuming that $\sim 30\%$ of all bursts are actually observed), so we estimate $\Omega_{\mathrm{PBH}} \sim (2-5)\times 10^{-8}\Omega_b \sim 10^{-9}$, an order of magnitude less than the current upper limit. Small-scale clustering of dark matter prior to the formation of the first stars[15] may significantly increase the probability of PBH capture.

Thus almost all the progenitors should belong to the pregalactic population of stars.[16] These early Population III stars have been captured later by forming galaxies, but NSs are born with a kick velocity of several hundred km/s, which in many cases is sufficient to escape the gravitation of a nearby protogalaxy. Therefore, GRBs have predominantly intergalactic origin.

The fall-down time of a PBH plotted as a function of its periastron distance rises steeply by orders of magnitude when this distance changes by less than 10% in crossing the radius of a NS. As a result, PBHs which induce GRBs are spread over fall-down times in such a manner that there is a peak which traces the formation of Population III stars. PBHs on orbits passing through NSs were swallowed almost immediately at a red shift $z \sim 10$ and gave rise to a distinct population of GRBs. More than 90% of all the PBH–NS binaries ever formed have not produced GRBs yet. They will give rise to future bursts at a steadily decreasing rate.

Decelerated by tidal friction, a PBH comes to rest at the center of the NS, establishing an almost steady infall of nucleons. A moderate compression applied to dense nucleon matter causes its transition to a quark–gluon plasma at a density $\sim(1-2)\times 10^{15}\,\mathrm{cm^{-3}}$ (Ref. 17). The latter occurs at a distance $10-100 r_g$ (r_g is the radius of the event horizon for a central black hole), depending on the actual conditions in the NS core. Close to the event horizon the accretion flow consists of ultrarelativistic particles only (quarks, gluons, photons, etc.) and has a velocity of sound equal to $c/\sqrt{3}$, so that the mass growth rate of black hole is $\dot{m} \simeq 6\sqrt{3}\pi\rho_n r_g^2 c$, where ρ_n is the undisturbed density of nucleon matter. According to this explosive instability equation, the time before the final implosion, $\tau_{\mathrm{im}} \simeq 4\times 10^6 (m_*/m)$ years, is small in comparison with the age of the Universe for $m \sim m_*$. During the whole latent period of PBH-triggered GRB, excluding the last 30 s, no more than 0.1 percent of the total energy is released.

Nucleons destroyed under the extreme pressure in the accretion flow produce a primary plasma consisting of *u* quarks and approximately twice as many *d* quarks, while the equilibrium composition is made of roughly equal numbers of *u*, *d*, and *s* quarks. It is the energy released in the $d \to s$ transition that heats the quark–gluon plasma and powers the burst through neutrino emission. The plasma beneath the phase transition boundary is heated to $\simeq 50$ MeV, as can be found by equating its thermal energy to the energy

difference between the initial and final compositions. The outer part of the star remains cold, and neutrinos and antineutrinos form a degenerate gas with a low effective scattering cross section. The reactions $\nu\bar{\nu}\to\pi^0\gamma$ and $\nu\bar{\nu}\to 3\gamma$ do not take place, because of the high energy threshold in the first case and the small cross section in the second case. Thus neutrinos do not annihilate until they reach the surface layer of the NS, where the electron Fermi level is low enough to permit the process $\nu\bar{\nu}\to e^+e^-$. The role of three-particle processes is small in comparison with elastic scattering. The latter alone deposits inside the NS an amount of energy which would be sufficient to heat the entire star up to 10 MeV, if this thermal energy were not advected by the accretion flow into the black hole. Thanks to the effect of advection, the hot nucleon matter is confined in a thin shell ($\lesssim 0.3$ km) around the quark core.

The main annihilation process $\nu\bar{\nu}\to e^+e^-$ begins at a depth 0.5–1 km, where the neutrino gas is degenerate, and it takes some time — and distance — to heat the flow to the temperature corresponding to the conversion of $\simeq 50\%$ of the neutrinos' energy to energy of the thermal e^+–e^- plasma. As a result, an inverted temperature layer is formed that overlaps the region where the e^-e^+ wind begins. To get into this wind, baryons have to pass along a positive temperature gradient ∇T, where the radiation energy density increases as T^4 while the mass density decreases. Under such conditions the velocity of baryon flow at the sonic point, $[(\partial p/\partial r)/(\partial \rho/\partial r)]^{1/2}$, is many times smaller than the adiabatic velocity of sound for the case when radiation pressure dominates. Since the density at the sonic point is completely determined by the neutrino luminosity, the mass ejection rate in the presence of the inverted temperature layer is much smaller, reducing considerably the baryon pollution of the fireball. The Lorentz factor of the outflow in this case may reach several thousand and is limited by the total mass of nucleons lying above the sonic point.

After this stage our scenario joins the standard fireball models in assuming the formation of a shock in the surrounding medium, but with the principal modification dictated by the ultrarelativistic neutron flow, which carries as much energy as the proton–electron flow usually considered.[8] Depending on the parameters, the neutron flow may or may not decouple from the proton flow, so that the cosmological GRBs have a bimodal distribution of durations. Decoupling may lead to multipeaked light curves[18] and causes pion production in inelastic proton–neutron collisions. The latter results in the emission of $\sim 10^{-3}$ of the total GRB energy in the form of energetic photons originating in the decay of neutral pions and blue-shifted to ~ 100 GeV.

NEUTRINO TRANSFER

Now let us detail the energy transfer from the hot quark–gluon plasma in the NS core to the surface of the star. In the GRB case the transfer time is much less than the induced collapse time. The neutrino flow is therefore quasisteady, and the pressure gradient is equal to the rate of momentum density loss. In the case of spherical symmetry this gives

$$\frac{d\mu}{dR} = -\frac{3}{16\pi}\sigma_0 N_n \Sigma \lambda_c^2 h f, \tag{2}$$

where μ is the neutrino chemical potential, $\sigma_0 \simeq 1.8\times 10^{-44}\mathrm{cm}^2$, N_n is the neutron density, $\lambda_c = h/m_e c$ is the Compton wavelength of the electron, h is Planck's constant, and f is the neutrino flux density. The effective cross section for scattering of degenerate neutrinos $\Sigma = (2m_e c^2/\mu)^2 \sigma_{\mathrm{eff}}/\sigma_0$ has the thermal component[19]

$$\Sigma_T = \frac{1+2g_a^2}{4}\frac{\chi}{\mu}\frac{m_n c}{p_F^n}, \quad g_a \simeq 1.254, \tag{3}$$

where $\chi = \pi^2 T^2/4\varepsilon_F^n$ is thermal energy per neutron, m_n is the neutron mass, ε_F^n and p_F^n are the neutron Fermi energy and momentum ($T \ll \mu, \varepsilon_F^n$), and the nonthermal component is

$$\Sigma_N = 0.24\left(\frac{m_n c}{p_F^n}\right)^2 \frac{\mu}{p_F^n c}\left(\frac{\delta}{\mu}\right)^2, \quad \delta \ll \mu, \tag{4}$$

as may be easily calculated in the limit of heavy neutrons and zero temperature. Here δ is the dipole moment of the deviation of the neutrino Fermi surface $\varepsilon_F^\nu = \mu + \delta\cos\theta$ from the equilibrium level due to the nonzero flux. In the case of interest both the thermal and nonthermal parts are small, ($\sigma_{\mathrm{eff}} \ll \sigma_0$), and the total value Σ is simply their sum.

Equation (2) with the values (3) and (4) must be accompanied by two more continuity equations for the functions f and χ. For simplicity we assume $N_n = \mathrm{const}$ and a constant ratio of the radius of the quark–gluon core to the radius of the central black hole, $q = R_{\mathrm{in}}/r_g$. Next, we invoke the condition of zero neutrino losses, taking into account that their total flux differs at different radii due to the nonzero value of the derivative $\partial\mu/\partial t$, i.e., $dR_{\mathrm{in}}/dt \neq 0$. This provides an equation for the flux f, the same for all six types of particles (ν_e, ν_μ, ν_τ and antiparticles) insofar as they are considered massless.

On the way out, the neutrinos remain degenerate with a temperature equal to that of the local nucleon matter, since any deviation from equilibrium causes rather strong scattering and energy transfer from neutrinos to neutrons. Balancing the thermal neutron energy flux $v_n N_n \chi$ and all neutrino energy fluxes $6\mu f$, where v_n is the fall-down velocity of nucleon matter, yields the last equation for $\chi(R)$.

In both equations we replace the time derivative $\partial/\partial t$ by $-Rr_g^{-1}(q\,dr_g/dt)\partial/\partial R$, which implies a self-similar scaling of the profiles of μ and χ with increasing $R_{\mathrm{in}}(t)$. This approximation, which does not affect the numerical results significantly, reduces the problem to a system of three ordinary differential equations. The boundary conditions are as follows: $\mu = \mu_{\mathrm{in}}$ at $R = R_{\mathrm{in}}$, $\chi = 0$ and $f = k_{\mathrm{tr}}(4\pi/3)(\mu/hc)^3 c/4$ at the NS surface $R \simeq R_{NS}$. Here the factor $k_{\mathrm{tr}} < 1$ includes all possible opacities as compared with the free-streaming case $k_{\mathrm{tr}} = 1$. For definiteness we take $\mu_{\mathrm{in}} = 80$ MeV, although the results are not very sensitive to this value.

We have solved the neutrino transfer equations numerically and found that hot nucleon matter with $T \sim 10$ MeV is localized within several hundred meters above the emitting quark surface, while the rest of the NS remains cold, $T < 1$ MeV, allowing the neutrinos to escape almost freely; at such a temperature $\Sigma_N > \Sigma_T$. The final neutrino pulse has a duration of ~ 1 ms and typical neutrino energies 25–35 MeV. Altogether, neutrinos provide $\sim 10^{51}$ erg in the form of relativistic electron–positron outflow. It should be noted that due to gravitational lensing some bursts may have apparent energies tens and hundreds of times larger; this problem will be considered elsewhere.

CONCLUSION

We have introduced a new kind of astrophysical object, the PBH–NS binary, which arises naturally from Population III stars in the pregalactic epoch. The evolution of such a binary leads to the induced collapse of the NS, which we suggest as a source for cosmological GRBs.

This novel scenario can cure many of the problems typical to cosmological models of GRBs, including the absence of time dilation, the lack of large host galaxies, and the small baryonic pollution of the fireball. The latter is possible thanks to the extraordinary physical conditions in a collapsing NS, where the hot neutrino-emitting quark core and the region of efficient neutrino annihilation are separated by a layer of cold and hence transparent neutron matter. Under these conditions, an inverted temperature layer is formed in the e^-e^+ plasma and serves as a barrier for baryons.

The induced collapse scenario gives several specific predictions. The simplest one to observe is a large peak near $z\sim 10$, corresponding to the first population of bursts and containing about half of all the GRBs which have taken place so far. Then a bimodal distribution of measured durations, with a crossover at the value ~ 5 s, should exist due to decoupling of the neutron and proton flows. We also suggest that ground-based observations of GRBs in the 100-GeV range could afford an opportunity to reveal physical conditions directly in a fireball, providing us with a new standard candle. Such observations are already within the sensitivity of modern instruments, but require at least a 30° effective field of view to ensure one detection per year, with allowance for a 10% duty cycle for such telescopes. Finally, a very weak pulse of gravitational radiation is expected in our scenario, contrary to the model of a coalescing NS binary.

This work was supported in part by the Russian Fund for Fundamental Research (Grant # 99-02-18244).

*)e-mail: derishev@appl.sci-nnov.ru
†)e-mail: kochar@appl.sci-nnov.ru
‡)e-mail: kochar@appl.sci-nnov.ru

[1] D. H. Hartmann and S. E. Woosley, Adv. Space Res. **15**(5), 143 (1995).
[2] I. G. Mitrofanov *et al.*, Astrophys. J. **459**, 570 (1996).
[3] B. Pacsyński, Astrophys. J. **308**, L43 (1986).
[4] S. E. Woosley, Astrophys. J. **405**, 273 (1993).
[5] D. E. Reichart and P. Mészáros, Astrophys. J. **483**, 597 (1997).
[6] E. V. Derishev, V. V. Kocharovsky, and Vl. V. Kocharovsky, Radiophys. Quantum Electron. **41**, 7 (1998).
[7] P. Mészáros and M. J. Rees, Mon. Notes Astron. Soc. South Afr. **269**, L41 (1994).
[8] E. V. Derishev, V. V. Kocharovsky, and Vl. V. Kocharovsky, Astrophys. J. **521**, 640 (1999).
[9] E. Woods and A. Loeb, Astrophys. J. **453**, 583 (1995).
[10] I. G. Mitrofanov, M. L. Litvak, and D. A. Ushakov, Astrophys. J. **490**, 509 (1997).
[11] J. H. MacGibbon and B. J. Carr, Astrophys. J. **371**, 447 (1991).
[12] D. N. Page and S. W. Hawking, Astrophys. J. **206**, 1 (1976).
[13] A. A. Belyanin, V. V. Kocharovsky, and Vl. V. Kocharovsky, Mon. Notes Astron. Soc. South Afr. **283**, 626 (1996).
[14] E. V. Derishev and A. A. Belyanin, Astrophys. J. **343**, 1 (1999).
[15] A. V. Gurevich, K. P. Zybin, and V. A. Sirota, Usp. Fiz. Nauk **167**, 913 (1997) [Phys. Usp. **40**, 869 (1997)].
[16] J. P. Ostriker and N. Y. Gnedin, Astrophys. J. **472**, L63 (1996).

[17] Ch. Kettner *et al.*, Phys. Rev. D **51**, 1440 (1995).
[18] E. V. Derishev, V. V. Kocharovsky, and Vl. V. Kocharovsky, Astron. Astrophys. **345**, L51 (1999).
[19] A. Burrows and J. M. Lattimer, Astrophys. J. **307**, 178 (1986).

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

Conformal blocks and correlators in a bosonized WZNW model

K. A. Saraĭkin*)

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia; Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

(Submitted 12 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 648–653 (25 November 1999)

The solutions of the Knizhnik–Zamolodchikov equations as conformal blocks of the Wess–Zumino–Novikov–Witten $SU(2)$ model on a sphere are examined. An action that permits finding the N-point correlators of the model, which are constructed in a natural manner from the conformal blocks, is proposed. This is an action of three free fields perturbed by a special marginal operator. The construction described should extend to the case of other groups and surfaces of higher genus.
 [S0021-3640(99)00322-9]

PACS numbers: 11.25.Hf, 02.20.Sv

1. The WZNW model[1,2] is a two-dimensional conformal field theory with an additional Kac–Moody algebraic symmetry. This model has the distinction that all known rational conformal field theories can be obtained from it by a coset construction[3] or by Drinfel'd–Sokolov reduction.[4] The primary fields $\Phi^{m_i\bar{m}_i}_{\Delta_i,\bar{\Delta}_i}$ of the WZNW model with conformal dimensions $(\Delta_i,\bar{\Delta}_i)$ belong to a tensor product of two finite-dimensional representations of a semi-simple Lie group. The indices m_i and $\bar{m}_i$ correspond to the ''left-hand'' and ''right-hand'' respresentations, respectively. The correlators of the primary fields, in accordance with the basic ideas of conformal field theory,[5] possess the holomorphic factorization property

$$\langle \Phi^{m_1\bar{m}_1}_{\Delta_1,\bar{\Delta}_1}(z_1,\bar{z}_1)\ldots\Phi^{m_N\bar{m}_N}_{\Delta_N,\bar{\Delta}_N}(z_N,\bar{z}_N)\rangle_{\mathrm{WZNW}}$$
$$=\sum_{a,b} C^{ab}\mathscr{F}_a^{m_1\ldots m_N}(z_1,\ldots,z_N)\overline{\mathscr{F}_b^{\bar{m}_1\ldots\bar{m}_N}(z_1,\ldots,z_N)}. \qquad (1)$$

For this reason the main objects of the theory become the so-called conformal blocks $\mathscr{F}_a^{m_1\ldots m_N}(z_1,\ldots z_N)$. They can be found by the bosonization method[6,7] based on the choice of a definite parameterization of a group element. For an appropriate parameterization, the action becomes quadratic in the fields, and the limitations arising in the process on the region of integration in the functional integral can be described as an insertion of so-called screening operators (integrals of certain 1-forms over noncontractable contours) into the correlators. However, such a construction does not give a

prescription for obtaining correlators from conformal blocks (the constants C^{ab} in Eq. (1) are not known in a general form). Moreover, many important properties of the conformal blocks remain veiled (the fact that they must satisfy a special differential[2] and algebraic[8,9] equations which follow from the internal symmetry of the model).

In the present letter an action that solves these problems is proposed. It makes it possible to find the complete correlators of the WZNW $SU(2)$ model. In so doing, the constituent conformal blocks immediately satisfy the required conditions. This is the action of three free fields perturbed by a special marginal operator.

2. Let us consider the $sl(2)$ algebra. It is given by the following commutation relations between the Chevalley generators $e, f,$ and h:

$$[h,e]=2e, \quad [h,f]=-2f, \quad [e,f]=h. \tag{2}$$

Then the Knizhnik–Zamolodchikov equations[2] which the conformal blocks must satisfy assume the form

$$\nabla_{\mathrm{KZ}}\mathscr{F}(\mathbf{z})\equiv\kappa\frac{\partial}{\partial z_i}\mathscr{F}(\mathbf{z})-\sum_{j\neq i}\frac{\Omega_{ij}}{z_i-z_j}\mathscr{F}(\mathbf{z})=0, \tag{3}$$

where $\Omega_{ij}\equiv 1/2h_ih_j+e_if_j+f_ie_j$, $\mathbf{z}\equiv(z_1,\ldots,z_N)$, $\kappa\equiv k+2$, and $k\in\mathbf{N}$. These are equations for a function of N variables $\mathscr{F}(z_1,\ldots,z_N)$ with values in the tensor product of N representations of the algebra $sl(2)$. The index i for the operator h_i (e_i,f_i) means that it operates as $h(e,f)$ on the i-th component in the tensor product and as a unit operator on all other components. We shall use the representation with the highest weight j:

$$h|j,0\rangle=-j|j,0\rangle, \quad f|j,0\rangle=0, \quad e^m|j,0\rangle=|j,m\rangle, \quad e^{j+1}|j,0\rangle=0.$$

For convenience, in what follows we denote by a $\mathbf{v}$ the ''vacuum vector'' of the tensor product of N representations:

$$\mathbf{v}\equiv|j_1,0\rangle\otimes\ldots\otimes|j_N,0\rangle.$$

Let us consider a realization of the Kac–Moody algebra $sl(2)_k$ using free fields:[10]

$$J^+=\beta, \quad H=:\beta\gamma:-\frac{iq}{\sqrt{2}}\partial\phi, \quad J^-=:\beta\gamma^2:-i\sqrt{2}q\gamma\partial\phi-k\partial\gamma. \tag{4}$$

Here $q=\sqrt{\kappa}$, ϕ is a scalar field with values on a circle, and β and γ are bosonic 1- and 0-differentials with the operator expansions

$$\phi(z)\phi(w)=-\log(z-w)+O(1), \quad \beta(z)\gamma(w)=\frac{1}{z-w}+O(1). \tag{5}$$

The holomorphic parts of the primary fields $\Phi^{m_i\bar{m}_i}_{\Delta_i,\bar{\Delta}_i}$ in such a formulation correspond to the vertex operators $V_{j,m}(z)$ forming the $sl(2)_k$ representation with the highest weight j (see Refs. 7 and 11):

$$V_{j,m}(z)=\gamma^m(z):\exp\left(i\frac{j}{\sqrt{2}q}\phi(z)\right): \tag{6}$$

All operators from this representation possess the same conformal dimension

$$\Delta_j = \frac{j/2(j/2+1)}{k+2}. \tag{7}$$

To reproduce the correct central charge of the theory, it is necessary to add to the action for the field ϕ a term with curvature, which with a special choice of the metric is equivalent to calculating the correlators with the insertion of the so-called vacuum charge[12]

$$V_{\rm vac}(R) = \gamma(R) \ :\exp\left(i\frac{\sqrt{2}}{q}\phi(R)\right):.$$

In addition, the theory contains a special operator with zero conformal dimension

$$\oint dtS(t) = \oint dt:\exp\left(-i\frac{\sqrt{2}}{q}\phi(t)\right):\beta(t) \tag{8}$$

— the so-called screening operator (the Feĭgin–Fuks operator[13]). It commutes with the algebra (4), and it does not change the conformal properties of the correlator.

3. Following Ref. 8, it is convenient to combine the vertex operators (6) into a single multiplet:

$$\tilde{V}_j = \sum_{m=0}^{\infty} c_m V_{j,m} e^m \tag{9}$$

(terms with $m>j$ in this sum are immaterial, since because $e^{i+j}|j,0\rangle = 0$ they vanish under the action of $\tilde{V}_j$ on **v**). Now, calculating the correlators of the operators $\tilde{V}_{j_1}, \ldots, \tilde{V}_{j_N}$, we obtain all possible N-point conformal blocks — they are given by the coefficients of $e_1^{m_1} \ldots e_N^{m_N}$. The constants c_m in Eq. (9) are uniquely determined from the well-known operator expansion,[2] which the primary fields with currents must have:

$$J^a(z)\tilde{V}_j(w) = \frac{t^a}{z-w}\tilde{V}_j(w) + O(1) \tag{10}$$

(here $a=\{1,2,3\}$, $J^1 \equiv J^+$, $J^2 \equiv J^-$, $J^3 \equiv H$, $t^1 \equiv e$, $t^2 \equiv f$, $t^3 \equiv h$). From Eq. (10) we find[1)]

$$c_m = \frac{1}{m!},$$

after which the series (9) sums to an exponential — the vertex operator is ''dressed''

$$\tilde{V}_j(z) \equiv :\exp\left(i\frac{j}{\sqrt{2}q}\phi(z)\right):\exp(\gamma(z)e), \tag{11}$$

The conformal blocks are the holomorphic parts of the correlators of the ''dressed'' vertex operators (11) with the insertion of n screening operators:[2)]

$$\mathcal{F}_a(\mathbf{z}) = \oint_a dt_1 \ldots dt_n \langle S(t_1)\ldots S(t_n)\tilde{V}_{j_1}(z_1)..\tilde{V}_{j_N}(z_N)\rangle \mathbf{v}, \tag{12}$$

where the index a enumerates the various collections of integration contours. The number n in this formula is dictated by the law of conservation of charge:

$$n=\frac{1}{2}\sum_{i=1}^{N} j_i+1, \tag{13}$$

where the number 1 corresponds to the contribution of the vacuum charge. Indeed, the functional integral over the field ϕ[3)] in Eq. (12) gives the general factor

$$\prod_{p<q}(t_p-t_q)^{\frac{2}{\kappa}}\prod_{i<l}(z_i-z_l)^{\frac{j_i j_l}{2\kappa}}\prod_{p=1}^{n}\prod_{l=1}^{N}(t_p-z_l)^{-\frac{j_l}{\kappa}}, \tag{14}$$

and the integral over the (β,γ) fields is

$$\sum_{\Sigma m_i=n}\frac{1}{m_1!\ldots m_N!}\langle\beta(t_1)\ldots\beta(t_n)\gamma^{m_1}(z_1)\ldots\gamma^{m_N}(z_N)\rangle e_1^{m_1}\ldots e_N^{m_N}\mathbf{v}$$

$$=\sum_{\Sigma m_i=n}\sum_{\text{perm}}\frac{e_{\sigma(1)}}{t_1-z_{\sigma(1)}}\ldots\frac{e_{\sigma(m)}}{t_m-z_{\sigma(m)}}\mathbf{v}=\prod_{p=1}^{n}\sum_{i=1}^{N}\frac{e_i}{t_p-z_i}\mathbf{v}, \tag{15}$$

where the symbol Σ_{perm} signifies summation over all permutations of the numbers $\{\sigma(1),\ldots,\sigma(N)\}$ such that among them the number i $(i=1,\ldots,N)$ is encountered exactly m_i times. Following Ref. 14, it is convenient to introduce the function

$$Y(\mathbf{z},t)=\prod_{l=1}^{N}(t-z_l)^{-\frac{j_l}{\kappa}}\sum_{i=1}^{N}\frac{e_i}{t-z_i}. \tag{16}$$

Then, collecting together Eqs. (14) and (15), we obtain the expression

$$\mathcal{F}_a(\mathbf{z})=\prod_{i<l}(z_i-z_l)^{\frac{j_i j_l}{2\kappa}}\oint_a dt_1\ldots dt_n\prod_{p<q}(t_p-t_q)^{\frac{2}{\kappa}}Y(\mathbf{z},t_1)\ldots Y(\mathbf{z},t_n)\mathbf{v}, \tag{17}$$

which is identical to well-known Shekhtman–Varchenko solution[14–16] of the Knizhnik–Zamolodchikov equations.

It is well known that, generally speaking, not all solutions of the KZ equations are conformal blocks of the WZNW model, since the latter correspond to integrable representations of the Kac–Moody algebra and must satisfy additional algebraic equations[8,9] guaranteeing splitting off of the null vectors. Feĭgin, Shekhtman, and Varchenko[18] have proved that the solution (17) satisfies these equations.[4)]

4. Let us now construct the correlators from the conformal blocks. We shall confine ourselves to the case of spinless primary fields, for which $\Delta=\overline{\Delta}$. Let us construct for the primary fields the analogs of multiplets (9):

$$\Phi_{\Delta_j}(z,\overline{z})=\sum_{m,\overline{m}=0}^{j}c_{m\overline{m}}\Phi_{\Delta_j}^{m\overline{m}}(z,\overline{z})\;e_L^m\otimes e_R^{\overline{m}}. \tag{18}$$

Given the weights $\{j_1,\ldots,j_N\}$ or, equivalently, the conformal dimensions $\{\Delta_{j_1},\ldots,\Delta_{j_N}\}$, we obtain from the formula (12) a set of conformal blocks $\mathcal{F}_a(\mathbf{z})$, differing from one another by the choice of integration contours. The constants C^{ab} in Eq. (1) must be determined by the requirement of single-valuedness as some points in the

correlator are carried around other points. By analogy with the Dotsenko–Fateev representation for minimal models,[12] it is natural to propose for the desired correlator the expression

$$\langle \Phi_{\Delta_{j_1}}(z_1,\overline{z_1})\ldots\Phi_{\Delta_{j_N}}(z_N,\overline{z_N})\rangle_{\mathrm{WZNW}}$$

$$\sim \int_\Sigma d^2t_1\ldots d^2t_n|\langle S(t_1)\ldots S(t_n)\tilde{V}_{j_1}(z_1)\ldots\tilde{V}_{j_N}(z_N)\rangle|^2\mathbf{v}\otimes\mathbf{v}, \tag{19}$$

where the symbol $\int_\Sigma$ denotes integration over the entire surface. The requirement of single valuedness evidently is satisfied. It is noteworthy that it is then possible to ensure the appearance of screening operators by simply adding to the free action

$$S_0=\frac{1}{4\pi}\int_\Sigma d^2t\left(\frac{1}{2}\partial\phi(t,\overline{t})\overline{\partial}\phi(t,\overline{t})-\beta(t)\overline{\partial}\gamma(t)-\tilde{\beta}(\overline{t})\partial\tilde{\gamma}(\overline{t})+i\frac{\sqrt{2}}{q}\mathscr{R}\phi(t,\overline{t})\right) \tag{20}$$

a term with an interaction S_{int} of the form

$$S_{\mathrm{int}}=\alpha\int_\Sigma d^2t\ \ \beta(t)\tilde{\beta}(\overline{t}):\exp\left(-i\frac{\sqrt{2}}{q}\phi(t,\overline{t})\right): \tag{21}$$

with a constant α and introducing the ''dressed'' vertex operator[5)]

$$\tilde{V}_j(z,\overline{z})\equiv:\exp\left(i\frac{j}{\sqrt{2}q}\phi(z,\overline{z})\right):\ \exp(\gamma(z)e_L+\tilde{\gamma}(\overline{z})e_R). \tag{22}$$

Then the WZNW correlator will be given by the functional integral

$$\langle \Phi_{\Delta_{j_1}}(z_1,\overline{z_1})\ldots\Phi_{\Delta_{j_N}}(z_N,\overline{z_N})\rangle_{\mathrm{WZNW}}$$

$$=\int \mathscr{D}[\phi,\beta,\gamma,\tilde{\beta},\tilde{\gamma}]\exp(-S_0-S_{\mathrm{int}})\tilde{V}_{j_1}(z_1,\overline{z_1})\ldots\tilde{V}_{j_N}(z_N,\overline{z_N})\mathbf{v}\otimes\mathbf{v}. \tag{23}$$

Indeed, expanding $\exp(-S_{\mathrm{int}})$ in a series we obtain correlators with the insertion of screening operators, and because of the condition (13) only a term of order n remains and we arrive at the expression (19).

The integrals over a Riemann sphere, which appear on the right-hand side and can contain a divergence, must be understood in the sense of analytic continuation with respect to the parameters, using formulas of the type

$$\int d^2z|z|^{2a}|1-z|^{2b}=\pi\frac{\Gamma(1+a)\Gamma(1+b)\Gamma(-1-a-b)}{\Gamma(-a)\Gamma(-b)\Gamma(2+a+b)}. \tag{24}$$

It should be noted that the formulas (14), (15), and (19) are similar to the expressions (19) and (20) from Ref. 19 for the scalar product of the quantum states of the Chern–Simons theory — it is obtained by making the substitution

$$(k+2)\to-(k+2).$$

Thus, we have obtained the following prescription for constructing the conformal blocks and correlators of the WZNW $SU(2)$ model: *An N-point conformal block is the holomorphic part of the correlator of N ''dressed'' vertices and n screening operators in the*

theory of free (ϕ,β,γ) fields. In addition, the number n is dictated by charge conservation and the N-point WZNW correlator is the correlator of N ''dressed'' vertex operators in the theory of (ϕ,β,γ) fields with an interaction of the form (12).

The construction described aboved extends to other groups for which the conformal blocks are well known.[7] The extension to a surface of higher genus is most tempting. As a consequence we should obtain the solutions of the Knizhnik–Zamolodchikov–Bernard equations[20,21] that were previously known only for a theory on a torus. Moreover, it is certainly of interest to study the (ϕ,β,γ) system with the action $S_0+S_{\rm int}$. The renewed interest shown in recent years in the bosonization of the WZNW model in connection with the variants of compactification, studied in string theory, with the participation of the spaces AdS_3 and AdS_5 should also be noted.[22]

I thank A. Yu. Morozov for stimulating discussions and his interest in this work, I. V. Polyubin, A. D. Mironov, and A. V. Marshakov for critical remarks, and A. S. Losev for valuable discussions and encouragement at the final stage of this work.

This work was supported in part by a grant from the Russian Fund for Fundamental Research (No. 98-02-16575) and a grant from the President of the Russian Federation (No. 96-15-96939).

*)e-mail: saraikin@itp.ac.ru

1)This is most easily seen from the operator expansion with J^+.

2)Here and below the expression $\langle\ldots\rangle$ signifies the holomorphic part of the corresponding correlator from the theory of free fields, while $\langle\ldots\rangle_{\rm WZNW}$ is the complete correlator of the WZNW model.

3)We recall that $\langle\Pi_{i=1}^N :\exp(i\alpha_i\phi(z_i)):\rangle=\Pi_{i<l}(z_i-z_l)^{\alpha_i\alpha_l}\delta_{\Sigma\alpha_i,0}$, where the Kronecker δ function arises because of the integration over the zeroth mode. In what follows, the condition $\Sigma\alpha_i=0$ is always assumed to hold and the Kronecker symbol is dropped. Likewise, for brevity, we do not write out the vacuum charge explicitly. Wick's theorem holds for the correlators of the (β,γ) fields. In addition, in this prescription all correlators in which the number of β fields is different from the number of γ fields are zero.

4)We wish to make a remark concerning the choice of the integration contours in Eq. (17). Substituting the integrand in Eq. (17) into the algebraic equations and the KZ equations, the total derivatives with respect to the variable t_i of expressions of the type (14) (exact forms) are obtained. For this reason, the contours must be chosen so that such exact forms would vanish on them. In specific calculations, one can use, for example, Fel'der's[17] or the Dotsenko–Fateev[12] prescription.

5)The $\tilde\beta,\tilde\gamma$ fields correspond to an antiholomorphic sector of the model, and e_L and e_R act on the left- and right-hand representations, respectively.

[1] E. Witten, Commun. Math. Phys. **92**, 455 (1984).
[2] V. Knizhnik and A. Zamolodchikov, Nucl. Phys. B **247**, 83 (1984).
[3] P. Goddard, A. Kent, and D. Olive, Phys. Lett. **152B**, 88 (1985).
[4] V. Drinfel'd and V. Sokolov, *Current Problems of Mathematics* (VINITI SSSR, 1984) p. 24.
[5] A. Belavin, A. Polyakov, and A. Zamolodchikov, Nucl. Phys. B **241**, 333 (1984).
[6] B. Feĭgin and E. Frenkel', Usp. Mat. Nauk **43**, 227 (1988); B. Feigin and E. Frenkel, Lett. Math. Phys. **19**, 307 (1990).
[7] A. Gerasimov, A. Marshakov, A. Morozov *et al.*, Int. J. Mod. Phys. A **5**, 2495 (1990).
[8] A. Zamolodchikov and V. Fateev, Yad. Fiz. **43**, 1031 (1986) [Sov. J. Nucl. Phys. **43**, 657 (1986)].
[9] D. Gepner and E. Witten, Nucl. Phys. B **278**, 493 (1986).
[10] M. Wakimoto, Commun. Math. Phys. **104**, 605 (1986).
[11] V. Dotsenko, Nucl. Phys. B **338**, 747 (1990).
[12] V. Dotsenko and V. Fateev, Nucl. Phys. B **240**, 312 (1984); V. Dotsenko and V. Fateev, Nucl. Phys. B **251**, 691 (1985).
[13] B. Feigin and D. Fuks, Funct. Anal. Appl. **17**, 241 (1983).

[14] P. Etingof, I. Frenkel, and A. Kirillov, *Lectures on Representation Theory and the Knizhnik-Zamolodchikov Equations.*
[15] V. Schechtman and A. Varchenko, Lett. Math. Phys. **20**, 279 (1990); Invent. Math. **106**, 139 (1991).
[16] H. Awata, Suppl. Prog. Theor. Phys. **110**, 303 (1992), hep-th/9202032.
[17] G. Felder, preprint Zurich (1988); G. Felder and R. Silvotti, Phys. Lett. B **231**, 411 (1989).
[18] B. Feigin, V. Schechtman, and A. Varchenko, Commun. Math. Phys. **163**, 173 (1994).
[19] F. Falceto, K. Gawedzki, and A. Kupiainen, Phys. Lett. B **260**, 101 (1991).
[20] D. Bernard, Nucl. Phys. B **303**, 77 (1988), D. Bernard, Nucl. Phys. B **309**, 145 (1988).
[21] D. Ivanov, Int. J. Mod. Phys. A **10**, 2507 (1995); hep-th/9410091; D. Ivanov and A. Losev, ITEP-TH-79/97..
[22] O. Andreev, hep-th/9909222; M. Banados and A. Ritz, hep-th/9906191; J. de Boer and S. Shatashvili, hep-th/9905032; D. Kutasov and N. Seiberg, hep-th/9903219; N. Berkovits, C. Vafa, and E. Witten, hep-th/9902098; J. de Boer, H. Ooguri, H. Robins, and J. Tannenhauser, hep-th/9812046; A. Giveon, D. Kutasov, and N. Seiberg, hep-th/9806194.

Translated by M. E. Alferieff

Exact solution of the problem of laser focusing of an atomic beam

V. V. Klimov*)
P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

V. S. Letokhov
Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

(Submitted 28 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 654–659 (25 November 1999)

The exact solution (Green's function) of a parabolic equation describing the motion of neutral atoms in the field of a hollow TEM_{01}^{*} mode of laser radiation is found. The positions of the focus and the width of the focal spot are found on the basis of this solution for various configurations of the laser atomic beam. © *1999 American Institute of Physics.* [S0021-3640(99)00422-3]

PACS numbers: 39.10.+j, 42.62.−b

Techniques and methods for controlling coherent and incoherent atomic beams are now being actively developed. The main control elements are atomic lenses and atomic mirrors.[1–3] Atomic lenses and atomic mirrors based on an optical gradient force have been investigated in detail theoretically and experimentally for sufficiently fast atomic beam.[4–8] Because of the complexity of the potential of the gradient force, the characteristics of the atomic lenses have been found either asymptotically (weak effect of the field)[5,6] or on the basis of quite rough approximations (ray optics)[7] or numerically for fixed values of the parameters.[6,8] The lack of exact analytical solutions makes it difficult to analyze the properties of atomic lenses, especially in the strong focusing region. In the present letter an atomic lens on a TEM_{01}^{*} mode for focusing atomic beams with arbitrary velocities is studied, and it is shown that the equation in the parabolic approximation has an exact analytic solution. The geometry of the problem is shown in Fig. 1.

The time-independent wave function of an atom transmitted through an atomic lens is described by the Schrödinger equation

$$\frac{\hbar^2 k_{dB}^2}{2m}\Psi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_{\mathrm{opt}}(\mathbf{r})\right]\Psi(\mathbf{r},t), \tag{1}$$

where V_{opt} is the focusing potential of the gradient force, $k_{dB}=2\pi/\lambda_{dB}$ is the wave number of the atom, and m is the mass of the atom. For the focusing potential we shall use a focused TEM_{01}^{*} mode[4,9]

$$V_{\mathrm{opt}}(r,z)=\frac{\hbar\Omega}{2}\ln[1+p(r,z)], \quad p(r,z)=\frac{I(r,z)}{I_s}\frac{\gamma^2}{\gamma^2+4\Omega^2},$$

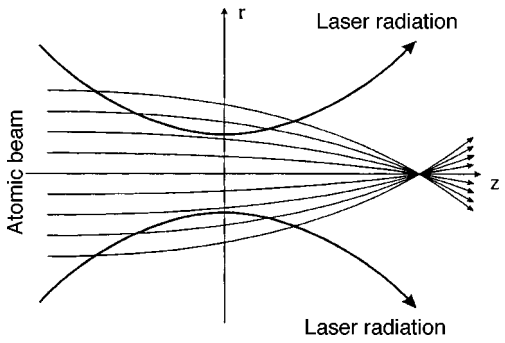

FIG. 1. Geometry of the focusing problem.

$$I(r,z)=4I_0\frac{2w_0^2r^2}{w^4(z)}\exp\left(-\frac{2r^2}{w^2(z)}\right). \tag{2}$$

In Eq. (2) $\Omega=\omega-\omega_0$ is the detuning of the radiation frequency ω relative to the frequency ω_0 of the atomic transition, γ and I_s are the line width and saturation intensity of the atomic transition, w_0 is the radius of the waist of the laser beam in the $z=0$ plane, $w^2(z)=w_0^2(1+z^2/z_R^2)$ is the size of the waist in the z plane, $z_R=(\pi/\lambda)w_0^2$ is the Rayleigh length, and λ is the wavelength of the optical radiation. The potential of the gradient force (2) is shown in Fig. 2a.

It is difficult to solve Eq. (1) with the potential (2) analytically. For this reason, further simplifications of Eq. (1) are required. The first simplification is a parabolic approximation, based on the fact that the de Broglie wavelength λ_{dB} of an atom is small compared with the characteristic longitudinal gradients. The solution is sought in the form[10]

$$\Psi=e^{ik_{dB}z}\psi, \tag{3}$$

where ψ is a slowly varying function in the direction of propagation of the beam. Neglecting the second derivatives, we obtain along the beam axis the standard parabolic equation[10]

$$ik_{dB}\frac{\hbar^2}{m}\frac{\partial}{\partial z}\psi(r,z)=\left[-\frac{\hbar^2}{2m}\Delta^{\perp}+V_{\rm opt}(r,z)\right]\psi(r,z), \tag{4}$$

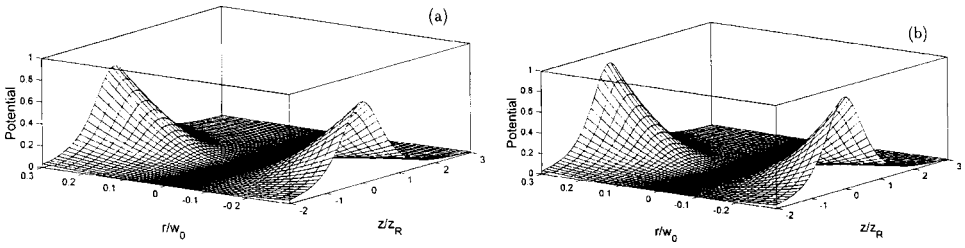

FIG. 2. Form of the focusing potential (a — exact, b — paraxial approximation).

where $\Delta^{\perp}=\partial^2/\partial x^2+\partial^2/\partial y^2$.

The equation (4) likewise cannot be solved analytically. However, for sufficiently narrow beams the potential (2) can be replaced by only the first term in its expansion in terms of the transverse coordinates x,y:

$$V_{\rm opt}(r,t)\approx\hbar\gamma\frac{C(r/w_0)^2}{2(1+(z/z_R)^2)^2},\quad C=\frac{I_0}{I_s}\frac{8(\Omega/\gamma)}{1+4(\Omega/\gamma)^2}. \tag{5}$$

The potential (5) is shown in Fig. 2b.

Using the dimensionless parameters $\tau=z/z_R$ and $\tilde{\mathbf{r}}=\mathbf{r}/w_0$, Eq. (4) with the potential (5) can be written in the form

$$i\frac{\partial}{\partial\tau}\psi(\tilde{\mathbf{r}},\tau)=p_1\left[-\frac{1}{4}\tilde{\Delta}^{\perp}+p_2\frac{C}{2}\frac{\tilde{r}^2}{(1+\tau^2)^2}\right]\psi(\tilde{\mathbf{r}},\tau), \tag{6}$$

where $p_1=\lambda_{dB}/\lambda$ and $p_2=\gamma m w_0^2/2\hbar$.

The equation (6) can be solved analytically. To find this solution we note that Eq. (6) is a time-dependent Schrodinger equation, where the coordinate τ plays the role of time and the potential is of a harmonic, time-dependent character. In this case, the path integral method makes it possible to show[11] that the expression for the Green's function, i.e., the solution of the inhomogeneous Schrödinger equation

$$i\frac{\partial}{\partial\tau}G(\tilde{\mathbf{r}},\tau;\tilde{\mathbf{r}}_1,0)-p_1\left[-\frac{1}{4}\tilde{\Delta}^{\perp}+p_2\frac{C}{2}\frac{\tilde{r}^2}{(1+\tau^2)^2}\right]G(\tilde{\mathbf{r}},\tau)=i\,\delta(\tau)\,\delta(\tilde{\mathbf{r}}-\tilde{\mathbf{r}}_1) \tag{7}$$

can be represented in the form

$$G(\tilde{\mathbf{r}}_2,\tau_2;\tilde{\mathbf{r}}_1,\tau_1)=F(\tau_2,\tau_1)e^{iS_{cl}[\tau_2,\tau_1]}, \tag{8}$$

where S_{cl} is the action along a classical trajectory connecting the points $(\mathbf{r}_1,\tau_1)$ and $(\mathbf{r}_2,\tau_2)$

$$S_{cl}[\tau_2,\tau_1]=\frac{2}{p_1}\int_{\tau_1}^{\tau_2}d\tau L[\mathbf{r}(\tau)],$$

$$L[\tilde{\mathbf{r}}(\tau)]=\left[\frac{1}{2}\left(\frac{d\tilde{\mathbf{r}}(\tau)}{d\tau}\right)^2-\frac{Cp_1^2p_2}{4}\frac{\tilde{\mathbf{r}}(\tau)^2}{(1+\tau^2)^2}\right], \tag{9}$$

and the factor $F(\tau_2,\tau_1)$ does not depend on the transverse coordinates.

To find S_{cl} it is necessary to determine the classical trajectories, whose equations have the form

$$\frac{d^2\mathbf{r}(\tau)}{d\tau^2}+\frac{Cp_1^2p_2}{2}\frac{\mathbf{r}(\tau)}{(1+\tau^2)^2}=0. \tag{10}$$

We note that the ray equation (10) is well known in the theory of magnetic lenses for charged particles.[12,13] The solution of Eq. (10) can be found analytically using the substitution $\tau=\cot(\alpha)$[12,13] and has the form

$$\mathbf{r}(z)=\frac{\mathbf{r}_1 \sin\alpha_1 \sin A(\alpha_2-\alpha)+\mathbf{r}_2 \sin\alpha_2 \sin A(\alpha-\alpha_1)}{\sin\alpha \sin A(\alpha_2-\alpha_1)}, \tag{11}$$

where $A=\sqrt{1+C^2p_1^2p_2/2}$ and $\alpha_{1,2}=\cot\tau_{1,2}$. Substituting the expression (11) into Eq. (9), we obtain the classical action S_{c1}

$$S_{cl}[\tau_2,\tau_1]=\frac{1}{p_1}\left[\begin{array}{l}\tilde{\mathbf{r}}_2^2 \sin^2\alpha_2(-A\cot A(\alpha_2-\alpha_1)+\cot\alpha_2)\\ -\tilde{\mathbf{r}}_1^2 \sin^2\alpha_1(A\cot A(\alpha_2-\alpha_1)+\cot\alpha_1)\\ +2\tilde{\mathbf{r}}_1\tilde{\mathbf{r}}_2 A\dfrac{\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)}\end{array}\right], \tag{12}$$

which is valid for any dimension of the space of transverse coordinates.

To determine the factor $F(\tau_2,\tau_1)$ we used the condition

$$G(\tilde{\mathbf{r}}_2,\tau_2;\tilde{\mathbf{r}}_1,\tau_1)=\int d\tilde{\mathbf{r}}_3 G(\tilde{\mathbf{r}}_2,\tau_2;\tilde{\mathbf{r}}_3,\tau_3)G(\tilde{\mathbf{r}}_3,\tau_3;\tilde{\mathbf{r}}_1,\tau_1). \tag{13}$$

Substituting the expressions (8) and (12) into Eq. (13) we arrive at the functional equation

$$F(\tau_2,\tau_1)=F(\tau_2,\tau_3)F(\tau_3,\tau_1)\frac{\pi p_1}{iA}\frac{\sin A(\alpha_3-\alpha_1)\sin A(\alpha_2-\alpha_3)}{\sin^2\alpha_3 \sin A(\alpha_2-\alpha_1)}, \tag{14}$$

which has the solution

$$F(z_2,z_1)=\frac{iA}{\pi p_1}\frac{\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)}. \tag{15}$$

As a result, the final expression for the Green's function of the parabolic equation (7) assumes the form

$$G(\tilde{\mathbf{r}}_2,\tau_2;\tilde{\mathbf{r}}_1,\tau_1)=\frac{iA}{\pi p_1}\frac{\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)} \times\exp\left[\frac{i}{p_1}\left(\begin{array}{l}\tilde{\mathbf{r}}_2^2 \sin^2\alpha_2(-A\cot A(\alpha_2-\alpha_1)+\cot\alpha_2)\\ -\tilde{\mathbf{r}}_1^2 \sin^2\alpha_1(A\cot A(\alpha_2-\alpha_1)+\cot\alpha_1)\\ +2\tilde{\mathbf{r}}_1\tilde{\mathbf{r}}_2 A\dfrac{\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)}\end{array}\right)\right], \tag{16}$$

where $\alpha_{1,2}=\cot\tau_{1,2}$ and $A=\sqrt{1+Cp_1^2p_2/2}$. The expression (16) for the Green's function is the main result of the present letter and can serve as a basis for finding the solutions of a variety of problems. For example, if the wave function $\varphi_0(\tilde{\mathbf{r}})$ is prescribed at $z=z_1$, then using the formula

$$\psi(\tilde{\mathbf{r}}_2,\tau)=\int d^2\tilde{\mathbf{r}}_1 G(\tilde{\mathbf{r}}_2,\tau;\tilde{\mathbf{r}}_1,\tau_1)\varphi_0(\tilde{\mathbf{r}}_1) \tag{17}$$

an expression can be found for the atomic probability density at any point in space. Since the Green's function is a Gaussian function of the transverse coordinates, the integral in Eq. (17) can be calculated analytically for an entire class of beams. Specifically, the choice of $\varphi_0(\mathbf{r})$ in a Gaussian axisymmetric form

$$\varphi_0(\mathbf{r}) = \exp(-\tilde{\mathbf{r}}^2/\tilde{\sigma}_0^2) \tag{18}$$

is of interest for practical applications. Here $\tilde{\sigma}_0^2$ describes the initial width of the beam. Substituting the expression (18) into Eq. (17) and performing the integration, we obtain for the wave function at any point in space

$$\begin{aligned}
&\psi(\tilde{\mathbf{r}}_2,\tau_2) = \frac{iA\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)(p_1/\tilde{\sigma}_0^2+iD_1)} e^{-\tilde{\mathbf{r}}_2^2/\tilde{\sigma}^2},\\
&\tilde{\sigma}^2 = p_1\frac{p_1/\tilde{\sigma}_0^2+iD_1}{D_2-ip_1/\tilde{\sigma}_0^2 D_3},\\
&D_1 = \sin^2(\alpha_1)[A\cot(A(\alpha_2-\alpha_1))+\cot\alpha_1],\\
&D_2 = \sin(\alpha_1)\sin(\alpha_2)\begin{bmatrix} A^2\sin(\alpha_1)\sin(\alpha_2)+\cos(\alpha_1)\cos(\alpha_2)\\ -A\sin(\alpha_2-\alpha_1)\cot(A(\alpha_2-\alpha_1))\end{bmatrix},\\
&D_3 = \sin^2(\alpha_2)[-A\cot(A(\alpha_2-\alpha_1))+\cot\alpha_2],
\end{aligned} \tag{19}$$

which simplifies substantially on the axis of the system ($\tilde{\mathbf{r}}_2=0$)

$$\psi(\tilde{\mathbf{r}}_2=0,\tau_2) = \frac{iA}{(p_1/\tilde{\sigma}_0^2+i\sin^2\alpha_1(A\cot A(\alpha_2-\alpha_1)+\cot\alpha_1))}\frac{\sin\alpha_1\sin\alpha_2}{\sin A(\alpha_2-\alpha_1)}. \tag{20}$$

The maxima of $|\psi(\tilde{\mathbf{r}}_2=0,\tau_2)|^2$ determine the position of the focal points $z_c = z_R\tau_c = z_R\cot\alpha_c$, which satisfy the equation

$$D_1 = \sin^2(\alpha_1)[A\cot(A(\alpha_2-\alpha_1))+\cot\alpha_1] = 0. \tag{21}$$

The minimum width of the bean at the focal point $\tilde{\sigma}_c^2 = \tilde{\sigma}^2(\alpha_c)$ can be easily found from Eq. (19) as

$$\tilde{\sigma}_c^2 = \frac{1}{\tilde{\sigma}_0^2}\left(\frac{\lambda_{dB}}{\lambda}\right)^2\frac{1+\tau_c^2}{A^2\sin^2\alpha_1+\cos^2\alpha_1}. \tag{22}$$

We note that in the parabolic approximation (19) and (20) (in contrast to geometric optics), generally speaking, the wave function cannot have singularities outside sources.

For a source located at $z=-\infty(\alpha_1=\pi)$ the positions of the focal points and the beam widths there are determined by the expressions

$$\tau_c = -\cot(\pi n/A), \tag{23}$$

$$\tilde{\sigma}_c^2 = \frac{1}{\tilde{\sigma}_0^2}\left(\frac{\lambda_{dB}}{\lambda}\right)^2\frac{1}{\sin^2(\pi n/A)}, \tag{24}$$

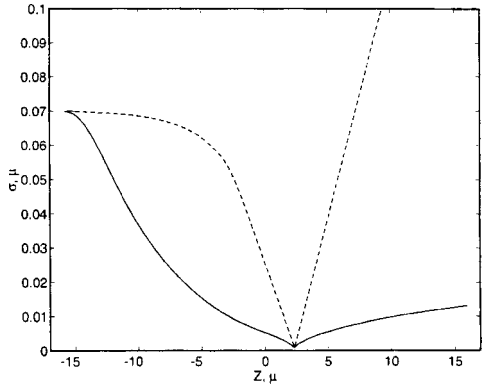

FIG. 3. Width σ of a beam of Na atoms versus the longitudinal distance ($\lambda = 0.59\ \mu$m, $\gamma/2\pi = 10$ MHz, $I_s = 10$ mW/cm^2, $I_0 = 10^8/2\pi$ W/cm^2, $V_z = 500$ m/s, and $\Omega/\gamma = \sqrt{I_0/I_s}$). The broken line shows the result of the numerical calculations in Ref. 7.

where $1 < A < 2$ for $n = 1$, $1 < A < 3$ for $n = 1,2$, and so on. Thus, depending on the magnitude of the focusing potential, several foci can be present in the system. Of course, the positions of the foci (21) and (23) are the same as the positions of the foci obtained in the geometric-optics approximation.[7,12,13]

Figure 3 shows the dependence of the beam width $\tilde{\sigma}$ on the longitudinal distance z for the parameters corresponding to Ref. 7. It is evident from this figure that the position of the focus and the minimum beam width, which were found in Ref. 7 using a numerical approach, agree quite well with the exact solution. However, on the whole, the behavior of the beam width in Ref. 7 along the axis deviates substantially from the exact solution (19) and (20).

The deviations will be even larger for strong focusing, i.e., low beam velocities. Figure 4 shows the three-dimensional probability density for the same parameters as in Fig. 3, with the exception of the initial beam velocity, which was chosen to be 100 m/s.

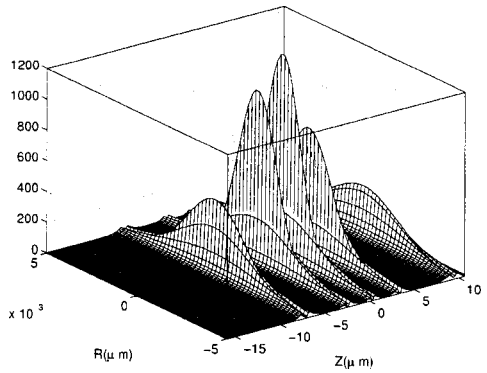

FIG. 4. Spatial distribution of the probability density $|\psi|^2$ for the same parameters as in Fig. 3, except $V_z = 100$ m/s.

In summary, in the present letter an analytic solution was found for the nonstationary Schrödinger equation with the potential

$$U \propto \mathbf{r}^2/(1+t^2)^2 \tag{25}$$

for any dimension of the space. The solution obtained could be helpful both for analyzing the propagation of atomic beams in the field of the TEM_{01}^{*} mode of laser radiation and for light propagation in a nonuniform medium with refractive index

$$n^2 \propto \mathbf{r}^2/(1+z^2)^2. \tag{26}$$

We thank the Russian Fund for Fundamental Research for financial support of this work.

*)e-mail: klimov@rim.phys.msu.su

[1] V. I. Balykin and V. S. Letokhov, *Atom Optics with Laser Light* (Harwood Academic Publishers, New York, 1995).
[2] C. S. Adams, M. Sigel, and J. Mlynek, Atom Optics – Phys. Rep. **240**, 143 (1994).
[3] J. P. Dowling and J. Gea-Banacloche, Adv. At., Mol., Opt. Phys. **37**, 1 (1996).
[4] V. I. Balykin and V. S. Letokhov, Zh. Éksp. Teor. Fiz. **94**, 140 (1988) [Sov. Phys. JETP **67**, 78 (1988)].
[5] V. I. Balykin and V. S. Letokhov, Usp. Fiz. Nauk **160**, 141 (1990) [Sov. Phys. Usp. **33**, 79 (1990)].
[6] G. M. Gallatin and P. L. Gould, J. Opt. Soc. Am. B **8**, 502 (1991).
[7] J. J. McClelland and M. R. Sheinfein, J. Opt. Soc. Am. B **8**, 1974 (1991).
[8] V. V. Klimov and V. S. Letokhov, J. Mod. Opt. **42**, 1485 (1995).
[9] W. W. Rigrod, Appl. Phys. Lett. **2**, 51 (1963).
[10] O. V. Rudenko and A. P. Sukhorukov, *Theory of Waves* (Nauka, Moscow, 1990).
[11] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965) [Russian translation, Mir, Moscow, 1968].
[12] M. Siladier, *Electron and Ion Optics* [Russian translation, Mir, Moscow, 1990].
[13] W. Glaser, *Grundlagen der Electronenoptik* (Springer, Vienna, 1952) [Russian translation, GITTL, Moscow, 1957].

Translated by M. E. Alferieff

Holographic-type parametric scattering: The universal character of the seed radiation

P. A. Prudkovskiĭ and A. N. Penin
*M. V. Lomonosov Moscow State University, 119899 Moscow, Russia**)

(Submitted 18 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 660–663 (25 November 1999)

It is concluded on the basis of experimentally obtained time dependences of the intensity of holographic-type parametric scattering induced by artificially produced seed radiation in a photorefractive lithium niobate crystal that the brightness of the natural seed radiation is independent of the pump intensity. This attests to a quantum origin of the seed radiation. © *1999 American Institute of Physics.* [S0021-3640(99)00522-8]

PACS numbers: 42.65.Hw, 42.70.Nq

Interest in quite complicated systems such as chemical autooscillatory reactions[1] or multicomponent polymers[2] has been increasing steadily in the last few years. This is due to progress in the fields of application of such systems and in the theory of self-organization and dynamic chaos. Photorefractive media, which are widely used today in dynamic holography, can serve as another example of such structures. One of the first self-organization phenomena in such media is holographic-type parametric scattering (HTPS), which was discovered in the mid-1980s.[3] It is characterized by the fact that narrow-angle frequency-degenerate light scattering is distinguished from wide-angle photoinduced light scattering (PILS) due to optical damage. However, even though photorefractive media have been studied for many years and have wide applications, only phenomenological models are still used in the theory of photorefraction. This is due to the difficulty of constructing mathematical models of the phenomena that take account of the entire spectrum of characteristic times from picoseconds to minutes or even hours for media such as lithium niobate and tantalate doped with copper or iron atoms.

As a result, a variety of phenomena occurring in photorefractive media cannot be explained on the basis of the existing theory of photorefraction. One unsolved problem is the nature of the seed radiation for HTPS and PILS. Despite the fact that at least three possible sources of the seed radiation were proposed more than 10 years ago — Rayleigh scattering of the pump by nonuniformities of the permittivity tensor or photogalvanic tensor,[4] flicker noise of the pumping intensity,[5] and vacuum fluctuations of the electromagnetic field[6] — the origin of the seed radiation is still unclear.

This work is devoted to an experimental investigation of the dynamical dependences of the HTPS intensity, which make it possible to draw definite conclusions about its seed radiation, attesting to its universal quantum origin. But, a final conclusion about the

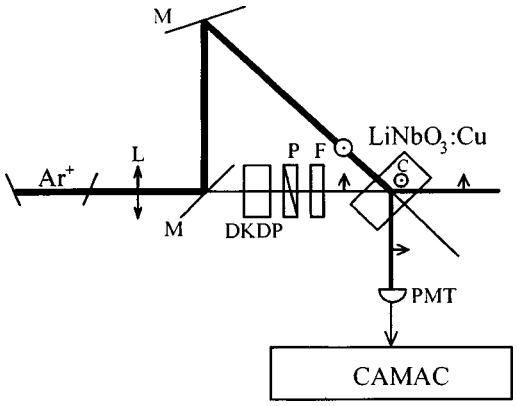

FIG. 1. Diagram of the experimental setup: Ar^+ — argon laser, L — long focal length lens, M — mirror, DKDP — birefringent dihydrophosphate crystal, P — polarizer, F — filter, $LiNbO_3$:Cu — photorefractive lithium niobate crystal, C — optic axis of the crystal, PMT — photomultiplier.

nature of the seed radiation must await a systematic theory explaining the mechanism of its appearance appears.

PILS and HTPS are explained by writing of many holographic gratings as a result of the interference of the pump with the so-called seed radiation, whose nature is still unclear, and the further self-consistent amplification of these gratings. A method for measuring the brightness of the seed radiation for HTPS, using the four-wave phase-matching condition, which is valid for this scattering and holds in lithium tantalate or niobate type ferroelectric crystals as a result of their anisotropy, is described in Ref. 6. For example, for *ee-oo*-type HTPS in $LiNbO_3$:Cu the following condition holds:

$$\mathbf{q}=\mathbf{k}_P^e-\mathbf{k}_{S_1}^o=\mathbf{k}_{S_2}^o-\mathbf{k}_P^e,$$

where $\mathbf{k}_P$ and $\mathbf{k}_S$ are, respectively, the wave vectors of the pump and scattered light; the superscripts determine the type of polarization; and, $\mathbf{q}$ is the wave vector of the holographic grating responsible for light scattering in two directions simultaneously. For this reason, if the seed radiation is produced artificially, by additional illumination with a light beam in one of the scattering directions (for example, in the direction $\mathbf{k}_S^1$), then the development of HTPS in the conjugate scattering direction ($\mathbf{k}_S^2$) will be determined by the writing of a holographic grating corresponding to a given brightness of the seed radiation. The difference between the time dependences of the development of HTPS with and without the additional light beam makes it possible to determine the excess brightness of this light beam above the brightness of the natural seed radiation. The brightness of the seed radiation was measured in Ref. 6 using this method, and it was concluded on the basis of an analysis of the brightness of various types of scattering that the seed radiation is of quantum nature. However, the method employed is extremely inaccurate, because it deals with absolute values of the brightness. A better method for investigating the seed radiation was used in the present work. This method uses only relative values and appeals to qualitative and not quantitative features of the experimental results obtained.

Figure 1 shows a diagram of the experimental setup. An argon laser beam, polarized

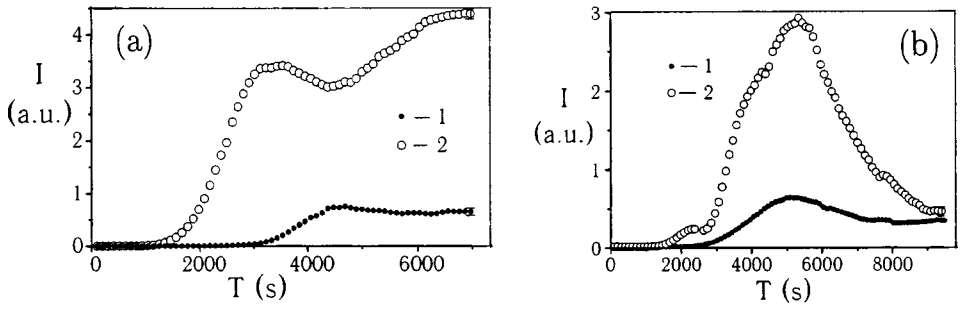

FIG. 2. Time dependences of the HTPS intensity with pump power $P_L \approx 2$ mW without (*1*) and in the presence of (*2*) an additional light beam attenuated by a filter by a factor of $\eta_2^{-1}=32$ (a) and $\eta_2^{-1}=250$ (b).

in an extraordinary direction, is incident along the normal direction into the $x-z$ surface of a lithium niobate crystal doped with copper atoms. A weak light beam is split off the pump on one of the mirrors, passes through a filter and a DKDT crystal — polarizer system, which rotates the polarization of the beam by 90°, and strikes the same point of the photorefractive crystal at an angle corresponding to HTPS $\theta=\cos^{-1}(\mathbf{k}^e/\mathbf{k}^o)$, determined from the condition of matching. The time dependence of the HTPS scattering intensity in the conjugate direction is recorded with an electronic system, which includes a photomultiplier, a CAMAC crate, and a IBM AT386 PC. The intensity of the additional light beam is related with the intensity of the pump beam by the ratio of the transmission and reflection coefficients of the mirror $\eta_1=T/R\approx 0.02$ and the transmission coefficient η_2 of the filter.

Figure 2 shows typical time dependences of the PHST intensity with laser-pump power $P_L\approx 2$ mW with and without an additional light beam ($\eta_2^{-1}=32$ and $\eta_2^{-1}=250$). It is easy to see that in the presence of the additional light beam HTPS develops more rapidly and reaches high intensities. This means that the brightness of this beam is much higher than that of the natural seed radiation. As the transmission coefficient of the filter decreases, the difference between the two dependences decreases and essentially vanishes for $\eta_2^{-1}=2\times 10^3$ (Fig. 3). The intensity of the additional beam is approximately $I_0=\eta_1\eta_2 P_L/S\approx 2$ μW/cm^2, where $S\approx 10^{-2}$ cm^2 is the cross-sectional area of the ad-

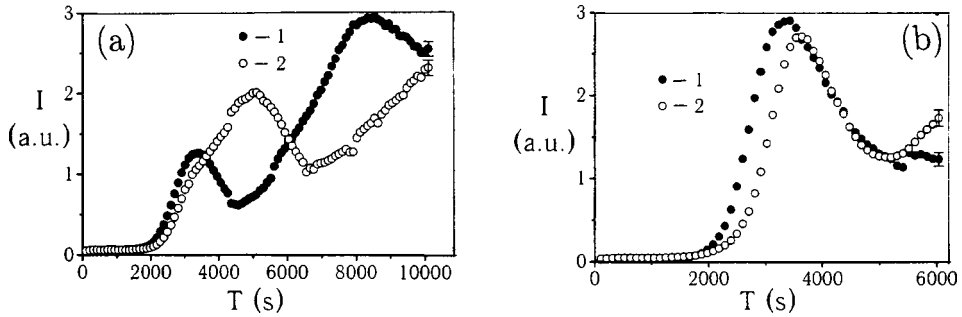

FIG. 3. Time dependences of the HTPS intensity with pump power $P_L\approx 2$ mW without (*1*) and in the presence of (*2*) an additional light beam attenuated by a filter by a factor of $\eta_2^{-1}=10^{-3}$ (a) and $\eta_2^{-1}=2\times 10^{-3}$ (b).

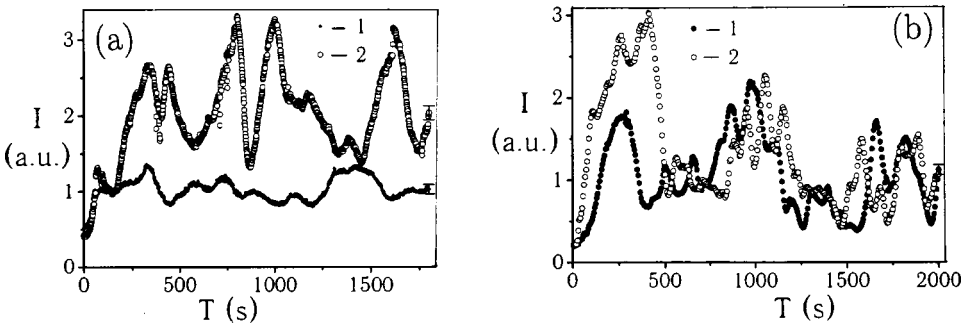

FIG. 4. Time dependences of the HTPS intensity with pump power $P_L \approx 20$ mW without (*1*) and in the presence of (*2*) an additional light beam attenuated by a filter by a factor of $\eta_2^{-1} = 2 \times 10^{-3}$ (a) and $\eta_2^{-1} = 2 \times 10^{-4}$ (b).

ditional beam. Using the value obtained for the intensity, it is easy to calculate the brightness of the seed radiation: $I_{\omega\Omega} = I_0 / \Delta\omega\Delta\Omega \approx 10^{-9}$ W/cm$^2 \cdot$s$^{-1} \cdot$sr, where $\Delta\omega = 5 \times 10^8$ s^{-1} and $\Delta\Omega = 4 \times 10^{-6}$ sr are the spectral and angular widths of the additional light beam. Similar dependences were obtained for the pump power $P_L \approx 20$ mW (Fig. 4). As one can see, in this case the dependence of the dynamics of development of HTPS on the presence of an additional beam becomes negligible only for $\eta_2^{-1} = 2 \times 10^4$, whence once again $I_0 \approx 2$ μW/cm^2 and $I_{\omega\Omega} \approx 10^{-9}$ W/cm$^2 \cdot$s$^{-1} \cdot$sr which is of the same order of magnitude as the value obtained in Ref. 6. However, more important here is the fact that the brightness of the seed radiation does not change as the pump intensity changes.

The absence of a dependence of the brightness of the seed radiation on the pump power contradicts most models of the appearance of seed radiation. Indeed, the model of Rayleigh scattering of the pump[4] and the flicker-noise model[5] assume, one way or another, that the seed radiation is generated by the pump and therefore its brightness depends on the pump intensity. The only model where the brightness of the seed radiation is completely unrelated with the pump intensity is the model in which holographic-type parametric scattering, just as spontaneous parametric scattering (down conversion), is induced by the zero-point fluctuations of the electromagnetic field,[6] whose brightness is one photon per mode.[7]

In summary, the results obtained in the present letter attest to a universal quantum origin of the seed radiation for HTPS, though, as already mentioned, a final conclusion in favor of this hypothesis must await the construction of a systematic quantum model of the HTPS writing process at the early stages of its development.

This work was supported by grant 99-02-16418 from the Russian Fund for Fundamental Research as part of the integration program ''Fundamental Optics and Spectroscopy.''

*)e-mail: postmast@qopt.ilc.msu.su

[1] A. Belmonte and J.-M. Flesselles, Phys. Rev. Lett. **77**, 1174 (1996).
[2] B. D. Ermi, G. Nisato, J. F. Douglas *et al.*, Phys. Rev. Lett. **81**, 3900 (1998).
[3] S. Odoulov, K. Belabaev, and I. Kiseleva, Opt. Lett. **10**, 31 (1985).

[4] V. V. Obukhovskiĭ, A. V. Stoyanov, and V. V. Lemeshko, Kvant. Élektron.**14**, 113 (1987) [Sov. J. Quantum Electron. **17**, 64 (1987)].
[5] B. I. Sturman, Zh. Éksp. Teor. Fiz. **100**, 1071 (1991) [Sov. Phys. JETP **73**, 593 (1991)].
[6] A. N. Penin and K. N. Zabrodin, Opt. Commun. **85**, 450 (1991).
[7] D. N. Klyshko and A. N. Penin, Usp. Fiz. Nauk **152**, 653 (1987) [Sov. Phys. Usp. **30**, 716 (1987)].

Translated by M. E. Alferieff

Subattosecond ($\leqslant 10^{-18}$ s) time measurements using coherent x-ray transition radiation

E. D. Gazazian, K. A. Ispirian,*) R. K. Ispirian, and M. I. Ivanian
Yerevan Physics Institute, 375036 Yerevan, Armenia

(Submitted 6 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 664–668 (25 November 1999)

The properties of the coherent x-ray transition radiation (CXTR) produced by intense, short electron beams microbunched in various devices, with periods corresponding to x-ray wavelengths $\sim 0.02-14$ nm and with modulation indices $\sim 0.001-1$, are investigated theoretically. The results of numerical calculations show that, just as in the case of incoherent and coherent transition radiation in the optical and microwave regions, CXTR can be used to study the microbunching processes and may thus find application for beam diagnostics in future x-ray FELs and advanced accelerator technology. © *1999 American Institute of Physics.* [S0021-3640(99)00622-2]

PACS numbers: 41.60.Cr, 41.50.+h, 06.30.Ft

Coherent transition radiation (CTR) is now used[1] for subpicosecond particle density distribution measurements. These measurements make use of the fact that the CTR spectrum is the Fourier transform of the longitudinal charge distribution and that, when the radiation wavelength λ is longer than the longitudinal σ_z and transverse σ_r sizes of the bunch, the radiation intensity is proportional to the square of the number N_b of electrons in the bunch. However, when x-ray FELs with self-amplified spontaneous emission (SASE) come online in the near future,[2,3] the femtosecond bunches will have subattosecond structures due to the particle grouping or microbunching (MB), with periods equal to the resonant radiation wavelength of the FEL.

The use of TR for the investigation of MB in the microwave region has been proposed[4] and now implemented.[5,6] Just as the first works[7–9] on the application of optical TR for high-energy particle beam diagnostics were, in the words of Ref. 9 ''initiated after the publication of Ref. 10,'' the present work follows Refs. 4–6, adopting the methods developed in them. Using a realistic model for MB, in this work we calculate the spectral and angular distributions of the CXTR and the dependence of the total CXTR photon number on the bunch and MB parameters. Numerical results given for the expected high-density, short bunches show the possibility of CXTR applications.

Suppose that after MB the particle distribution in a axially symmetric Gaussian bunch has the following form:

$$f(r,z)=\frac{N_b}{A}\left[\frac{\exp\left(-\frac{r^2}{2\sigma_r^2}\right)}{2\pi\sigma_r^2}\right]\left[\frac{\exp\left(-\frac{z^2}{2\sigma_z^2}\right)}{(2\pi)^{1/2}\sigma_z}[1+b_1\cos(k_r z)]\right], \tag{1}$$

where b_1 is the modulation index of the first harmonic of the MB (we neglect the much smaller higher harmonics), which takes place with a wavelength λ_r, $k_r=2\pi/\lambda_r$, and

$$A=1+2b_1\exp\left(-\frac{k_r^2\sigma_z^2}{2}\right) \tag{2}$$

is a factor determined by the normalization of the distribution function (1) and is practically equal to 1 in the case under discussion ($k_r\sigma_z\gg 1$) .

Using (1) and expressions for the spectral angular distributions of the incoherent XTR intensity and its relation with CXTR,[11] one can obtain the following expression for the number N_{CXTR} of CXTR photons produced by a bunch at an interface between vacuum and a substance with plasma frequency ω_p:

$$\frac{d^2N_{\mathrm{CXTR}}}{d\omega d\theta}=\frac{2\alpha N_b^2 b_1^2}{\pi A^2}\frac{\omega_p^4}{\omega^5}\frac{\theta^3}{(\gamma^{-2}+\theta^2)(\gamma^{-2}+\theta^2+\omega_p^2/\omega^2)}$$
$$\times\exp\left(-\left(\frac{\omega}{v}\sigma_r\theta\right)^2\right)\exp\left(-\left(\frac{\omega}{v}-k_r\right)^2\sigma_z^2\right), \tag{3}$$

where ω and θ are the CXTR photon emission frequency and polar angle, and v and $\gamma=E/mc^2$ are the particle velocity and Lorentz factor of the bunch, respectively.

As expected, the angular spectral distribution of the CXTR is mainly proportional to N_b^2 and b_1^2, decreases with increasing ω, and exceeds that of the incoherent XTR of N_b particles for $N_b b_1^2\exp(-(\omega_r\sigma_r\theta/v)^2)\gg 1$ when $\omega\simeq\omega_r$. It is also seen that CXTR is a diffraction-limited radiation, since it vanishes with increasing σ_r when $\omega\sigma_r\theta/v$ becomes much larger than 1.

The angular and spectral distributions of CXTR photons obtained by the integration of (6) can be written approximately in the forms

$$\frac{dN_{\mathrm{CXTR}}}{d\theta}\simeq\frac{\sqrt{\pi}v}{\sigma_z}\left.\frac{dN_{\mathrm{CXTR}}}{d\omega d\theta}\right|_{\omega=\omega_r}, \tag{4}$$

$$\frac{dN_{\mathrm{CXTR}}}{d\omega}=\frac{N_b^2 b_1^2\alpha}{\pi A^2\omega}\exp\left[-\left(\frac{\omega}{v}-k_r\right)^2\sigma_z^2\right]\begin{cases}\tau^4/(1+\omega_\gamma^2)^2, & \text{if} \quad \tau\ll 1,\\ (1+2\omega_\gamma^2)\ln\gamma, & \text{if} \quad \tau\gg 1,\xi\gg 1.\end{cases} \tag{5}$$

The total number of CXTR photons per bunch is obtained by integrating expression (4) over θ or expression (5) over ω and is equal to

$$N_{\mathrm{CXTR}}=\frac{\sqrt{\pi}v}{\sigma_z}\left.\frac{dN_{\mathrm{CXTR}}}{d\omega}\right|_{\omega=\omega_r}. \tag{6}$$

In Eqs.(4)–(6) $E_1(x)$ is the exponential integral function, and the following notation is used: $\omega_\gamma=\omega/(\omega_p\gamma)$ and $\xi=\omega_p\sigma_r/v$ are dimensionless frequencies, and $\tau=\gamma/q_{\mathrm{dif}}$,

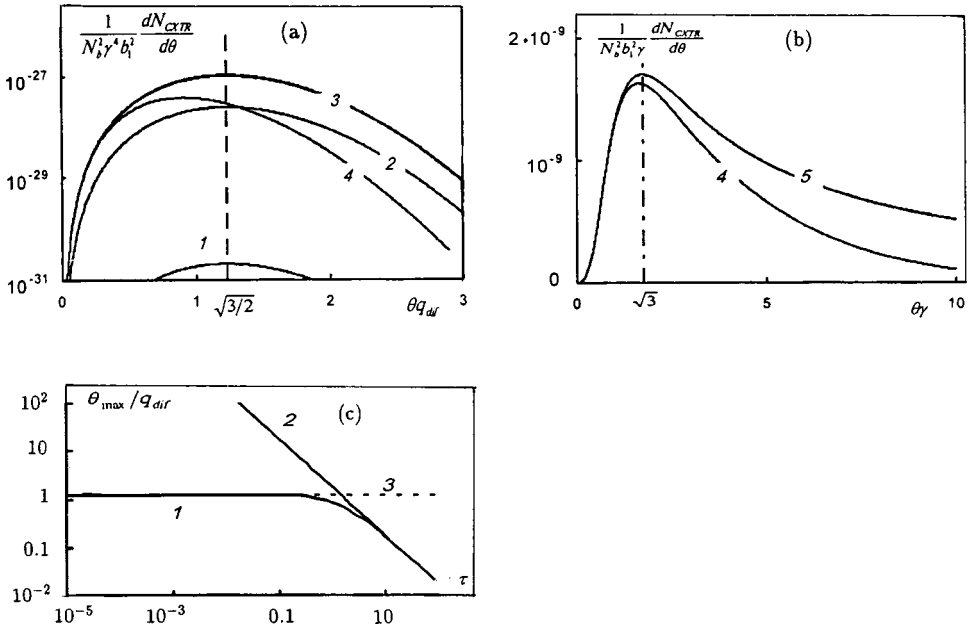

FIG. 1. a,b) Angular distributions of CXTR photons per particle in a bunch at a copper–vacuum interface for $q_{\rm dif}=1.26\times10^5$, normalized to γ^4 (a) and to γ (b) for γ = 25, 200, 10^3–10^5, 10^6, and 10^7–10^8 or $\tau=2.0\times10^{-5}$, 1.6×10^{-4}, 8.0×10^{-4}–8.0×10^{-2}, 0.8 and 8–80 (curves *1–5*, respectively). c) Dependence on τ of the angle $\theta_{\rm max}$ corresponding to the maximal CXTR and XTR intensity (curves *1* and *2*, respectively). Curve *3* is the direction $\theta=\sqrt{3}/\gamma/q_{\rm dif}$.

where $q_{\rm dif}=\omega_r\sigma_r/v$ is the dimensionless transverse size of the bunch, which determines the diffraction limit. Particular cases of expressions (4) and (5) for various relationships among the values of τ, ξ, and γ will be analyzed in Ref. 12.

In the discussion above it was assumed that the particles in the bunches are monochromatic. The expected high-energy bunches for SASE FELs will have uncorrelated and correlated energy spreads equal to $\Delta E/E=\Delta\gamma/\gamma=0.0002$ and 0.001, respectively.[3] It is clear that such energy spreads will have negligible influence on the XTR and CXTR angular distributions. However, they can result in substantial broadening of the narrow CXTR spectral distributions. Indeed, since the resonance wavelength in FELs is equal to $\lambda_r=l(1+K^2)/2\gamma^2$, where l and K are the undulator period and parameter, particles with different energies will have different resonance frequencies, and the CXTR photons they produce will have wider spectral distributions. Assuming the particles have a Gaussian energy distribution, after convolution with the above expressions for the CXTR spectral distribution, one can obtain lengthy expressions (not given in this work) for the real CXTR spectral distribution.

Figure 1a and 1b shows the angular distributions of CXTR photons produced at a vacuum–copper interface for $\lambda_r=1$ Å at different values of γ in the case $\sigma_z=2.5\times10^{-3}$ cm, $\sigma_r=2\times10^{-3}$ cm. These bunch sizes correspond to the beam parameters for the TESLA x-ray FELs and the SLAC LCLS. As to the CXTR angular distributions, we can say that they have different behaviors, depending on the ratio $\tau=\gamma/q_{\rm dif}$. If $\tau\ll1$

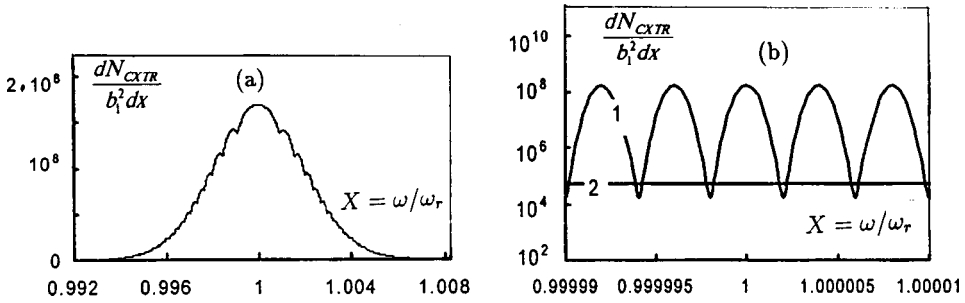

FIG. 2. a) The envelope of the spectral distribution of the number of CXTR photons for a bunch at a copper–vacuum interface in a relatively wide frequency region around $\omega=\omega_r$. b) The spectral distribution in a narrow, central frequency region, showing the CXTR oscillations (curve *2*) and smooth behavior of the XTR spectrum for $\gamma=5\times10^4$, $N_b=6\times10^9$, and $\lambda_r=1$ Å and with an energy spread $\Delta E/E=10^{-3}$.

(curves *1*, *2*, *3* in Fig. 1) the maximum of the curves, which is determined by the transverse size of the bunch, occurs at approximately $\theta_{\rm CXTR}^{\rm max}=\theta_{\rm dif}\simeq\sqrt{3/2}/q_{\rm dif}$. This value is much smaller than the angle of maximal incoherent XTR, $\theta_{\rm XTR}^{\rm max}\simeq\sqrt{3}/\gamma$. In the case $\tau\gg1$ the maximal radiation angle equals the maximal radiation angle of incoherent XTR (curve *5* in Fig. 1b). If $\tau\sim1$ (see curves *4* in Fig. 1a and 1b), the maximal radiation angle decreases with increasing τ in the range from $\tau\simeq0.32$ to $\tau\simeq5$, and it remains smaller than both $\theta_{\rm dif}$ and $\theta_{\rm XTR}$; this can be seen in Fig. 1c, which shows the dependence on τ of the maximal angle. At $\tau\geqslant5$ the maximal radiation angle approaches $\theta_{\rm XTR}$. The results of these numerical calculations show that by using angular discrimination one can separate the CXTR from the incoherent XTR, just as in the case of x-ray FELs,[2,3] where the stimulated and spontaneous radiations can be separated by angular discrimination.

As to the CXTR spectral distributions shown in Fig. 2a and 2b, their width is of the order of expected SASE x-ray width, but they exhibit oscillations with periods equal to the resonant frequency (Fig. 2b). By detecting these oscillations one can study the MB, at least at high values of N_b and b_1.

Figure 3 shows how the the normalized CXTR photon number depends on τ and γ (lower axis) for the DESY FEL bunch sizes. One can see that the behavior is different for

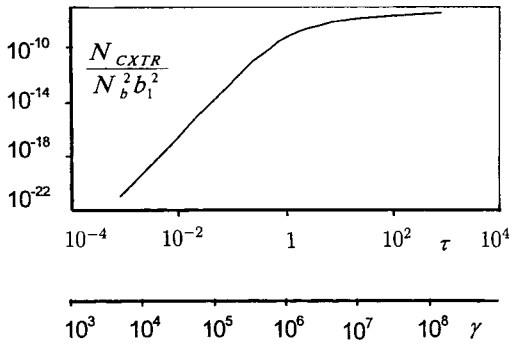

FIG. 3. Normalized total number of CXTR photons versus τ or γ (the lower x axis) for the TESLA x-ray FEL beam parameters, $E=25$ GeV, $\sigma_r=20$ μm, $\sigma_z=25$ μm, and $N_b=6\times10^9$.

$\tau \leqslant 1$ and $\tau \geqslant 1$ or for $\gamma \leqslant 10^6$ and $\gamma \geqslant 10^6$. In the latter case the photon number increases by a logarithmic law. For the parameters[2,3] of the TESLA x-ray FEL and SLAC LCLS the values of N_{CXTR}/b_1^2 are equal to 2.5×10^5 and 3.1×10^4, respectively.

As we have said, measurement of the MB parameters is essential for SASE FELs, IFELs, and advanced methods of particle acceleration. In all these cases it is necessary to measure the particle grouping period and the modulation index at various distances from the undulator entrance. Measurement of the CXTR spectrum is of special interest and presents some specific difficulties, similar to those which will be encountered in SASE FELs.[2,3] The microbunched beam will be accompanied by undulator or SASE x-ray photons having much higher intensities and almost the same spectral distribution as the CXTR, but they can be separated by absorption in the relatively thick CXTR radiator foil at low intensities or by magnetic deflection of the beam. As is planned for the SASE FELs,[2,3] the CXTR spectral measurements can be carried out after a large attenuation to avoid pile-up processes. From the experimentally measured CXTR spectrum and its maxima one can determine the grouping period. Measurement of the modulation index b_1 can be carried out by measuring the dependence of the CXTR yield on the number of particles in the bunches N_b or on the beam current. By measuring this dependence with the help of a simple ionization chamber and performing some calculations and fine spectral measurements, one can determine b_1 above certain values (depending on γ, N_b, etc).

Many simplifying assumptions have been made in the above calculations. For instance, we adopted a model of the particle distribution which took into account only the first harmonic of the MB. This assumption is justified by the expectation that the modulation indices b_2, b_3, etc. of the second and higher harmonics are much smaller, if no special measures are taken to enhance the higher harmonics. The correctness of this assumption is also supported by spectral measurements.[5,6] Another assumption, that MB does not affect the transverse particle distribution, is again justified by the experimental data.[5] The factor[13] taking the angular divergence of the beam into account may be neglected, because the primary beams of SASE FELs have very small emittances.

In conclusion, we note that although at present there are no very short bunches with bunch lengths $\sim 10^{-17}$–10^{-18} s, the fact that short laser pulses with lengths less than or of the order of the wavelength have been obtained recently leads one to hope that soon, when the difficulties connected with the Coulomb repulsion have been overcome, it will be possible to obtain charged particle beams with lengths less than 1 nm. Then the proposed method could be used for measuring the lengths of those beams as well as for observing and monitoring microbunching.

*)e-mail: ispirian@vx1.yerphi.am

[1] *Proceedings of the Third European Workshop on Beam Diagnostics and Instrumentation for Particle Accelerators*, DIPAC 97, Frascati (1997), LNF-97/048(IR), pp. 195–251.
[2] *Conceptual Design of a 500 GeV Linear Collider with Integrated X-Ray Laser Facility*, edited by R. Brinkman *et al.*, DESY 1997-048/ECFA (1997), p. 182.
[3] J. Arthur *et al.*, LCLS, Design Study Report, SLAC-R-521, UC-414 (1998).
[4] J. Rosenzweig, G. Travish, and A. Tremaine, Nucl. Instrum. Methods Phys. Res. A **365**, 255 (1995).
[5] Y. Liu *et al.*, Phys. Rev. Lett. **80**, 4418 (1998).
[6] A. Tremaine *et al.*, Phys. Rev. Lett. **81**, 5816 (1998).

[7] R. Wartski, S. Roland, J. Lasulle *et al.*, J. Appl. Phys. **46**, 3644 (1975).
[8] R. Wartski, T. Marcon, and S. Roland, IEEE Trans. Nucl. Sci. **NS-20**, 544 (1973).
[9] R. Wartski, S. Roland, and P. Brunet, *Proceedings of the International Symposium on Transition Radiation of High Energy Particles*, Yerevan (1977), p. 561.
[10] A. I. Alikhanian, K. A. Ispirian, and A. G. Oganesian, Zh. Éksp. Teor. Fiz. **56**, 1796 (1969) [Sov. Phys. JETP **29**, 964 (1969)].
[11] G. M. Garibian and Yan Shi, *X-Ray Transition Radiation* [in Russian], Publishing House of the Academy of Science of Armenian Republik, Yerevan (1983).
[12] E. D. Gazazian, K. A. Ispirian, R. K. Ispirian, and M. I. Ivanian, *Proceedings of the International Symposium RREPS99*, Baikal (1999), to be published in Nucl. Instrum. Methods B in 2000.
[13] Y. Shibata *et al.*, Phys. Rev. E **50**, 1479 (1994).

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

Decay of an electronic vortex of a collisionless shock wave as a possible mechanism of a ''Coulomb explosion''

A. B. Gordeev
Russian Science Center ''Kurchatov Institute'' 123182 Moscow, Russia

T. V. Loseva
Institute of the Dynamics of Geospheres, Russian Academy of Sciences, 117979 Moscow, Russia

(Submitted 23 September 1999; resubmitted 18 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 669–674 (25 November 1999)

A new mechanism of a ''Coulomb explosion,'' where ions are accelerated by the electric field separating charges at the magnetic Debye radius $r_B \sim B/4\pi e n_e$, is proposed on the basis of a nonquasineutral model of electronic vortices in a magnetic field. It is shown by means of numerical calculations that in the process of acceleration of the ions a collisionless shock wave, whose front has an effective width of the order of $\delta \sim r_B$, determined by the breakdown of quasineutrality, is formed in a time of the order of ω_{pi}^{-1}, where ω_{pi} is the ion plasma frequency. The origin of such explosive dynamics is the formation of ''holes'' in the electron density at characteristic times of the order of ω_{pe}^{-1} (ω_{pe} is the electronic plasma frequency) as a result of the generation of electronic vorticity by the Weibel instability of an electromagnetic wave. Calculations for a laser pulse with intensity $J \sim 6 \times 10^{18}$ W/cm^2 show that the ions expand in the radial direction with velocities up to 3.5×10^8 cm/s. © *1999 American Institute of Physics.* [S0021-3640(99)00722-7]

PACS numbers: 52.35.Tc, 52.40.Nk

1. The development of super-high-intensity lasers over the last several years has led to many possible applications.[1,2] The laser intensities here are $J > 10^{18}$ W/cm^2 with laser wavelengths $\lambda \leqslant 1$ μm and pulse durations $< 10^{-12}$ s.[3] The standard object discussed in the literature is a filament with initial radius $r_0 \sim c/\omega_{pe}$, out of which electrons are forced during the passage of a laser pulse. The discussion of the effects arising during the evolution of such a filament led to the concept of a ''Coulomb explosion,'' where ions fly apart with very substantial velocities. Various methods for the appearance of such filaments are discussed in the laser-plasma literature.[4] Exact measurements of the rapid laser-induced plasma expansion have shown experimentally the presence of velocities $\sim (2-4) \times 10^8$ cm/s for a laser pulse with intensity $J \sim 6 \times 10^{18}$ W/cm^2.[5] The acoustic mechanism of expansion proposed in Ref. 3 is questioned in Ref. 5. We share this doubt and propose a mechanism of plasma expansion based on a new model of electronic

vortices.[6] Here it is assumed that an electronic vortex arises with the passage of a laser pulse and results in the ion acceleration studied in what follows, while initially a ponderomotive acceleration by the electromagnetic field of the laser pulse operates.[7]

The main conclusion of the model of Ref. 6 is that breakdown of quasineutrality near the vortex axis can arise as a result of the generation of vorticity in a laser plasma. This happens if the size r_0 of the region of vorticity generation is comparable to the effective Debye radius $r_B \sim B/4\pi e n_e$ (B is the magnetic field) and the electron density n_e drops sharply in this region. Such ''holes'' in the electron density can be created in a laser plasma as a result of the generation of a magnetic field. In this case electrons are forced out of the region near the axis of an electronic vortex of size $r_0 \leqslant r_B \sim c/\omega_{pe}$ for quasistationary magnetic fields $eB/m_e w_{pe} c \geqslant 1$, which can be produced as a result of the Weibel instability for relativistic laser intensities $a \equiv eE_{\text{laser}}/m_e \omega_0 c \geqslant 1$, where ω_0 is the frequency of the laser radiation.[8] The characteristic magnitude of such a quasistationary magnetic field can be estimated on the basis of the results of Refs. 9 and 10. Estimates in accordance with Refs. 9 and 10 show that, in order of magnitude, such a magnetic field $B \simeq 4\pi e n_e \lambda a$, where λ is the wavelength of the laser radiation. Substituting here the parameters from Ref. 5 $a \simeq 2$, $n_e \simeq 10^{19}$ cm^{-3}, and $\lambda \simeq 1$ μm, we obtain the characteristic magnetic field $B \simeq 10^3$ T.

2. We shall show that an ideal object for realizing a ''Coulomb explosion''[4] could be an electronic vortex in a magnetic field. Such a vortex arises in a plasma produced as a result of the passage of a powerful laser wave through a gas. The main difference between the new model of electronic vortices and the Abrikosov model is that, together with the basic equation for the electronic vorticity[6]

$$\mathbf{\Omega}_e = \mathbf{B} - \frac{c}{e} \operatorname{curl} \mathbf{p}, \tag{1}$$

the Poisson equation

$$\operatorname{div} \mathbf{E} = 4\pi e (Z n_i - n_e) \tag{2}$$

must be used because of the breakdown of quasineutrality on a scale $\sim r_B$.

Since quasineutrality breaks down on the axis of an electronic vortex, the electron density decreases and the ions start to accelerate under the action of the electric field, so that a collisionless shock wave forms at times $\sim \omega_{pi}^{-1}$. In the process of such acceleration, the structure of the electronic vortex now depends on the nonuniform ion density, as a result of which the magnetic field readjusts and the structure of the electronic vortex breaks down. The key object in this process is the ''hole'' in the electron density with nonzero vorticity. The mechanism leading to the appearance of such a ''hole'' is immaterial, so that the nonquasineutral vortex model[6] makes it possible to describe the electronic ''hole'' arising for any reason. Here the difference in the geometry between the conventional problem of the propagation of a laser pulse and the formulation of the problem considered below, the quasistatic magnetic field being oriented along the z axis in both cases, should be underscored. Then, the azimuthal direction along φ in the vortex corresponds to laser pulse propagation along the y axis and ion motion in a direction along the radius r corresponds to transverse expansion of the plasma along the x axis.

In the proposed approach the Coulomb explosion process is divided into two independent stages. At the first stage the force structure readjusts at times $\sim \omega_{pe}^{-1}$ with sta-

tionary ions and generation of electronic vorticity and breakdown of quasineutrality. At the second stage the electric field arising because of nonquasineutrality leads to breakdown of the vortex as a result of ion motion at times of the order of ω_{pi}^{-1}.

It is not difficult to show that in the axisymmetric case, when the ions expand under the action of the radial electric field E_r, together with ion motion in the radial direction there also arises ion motion in the angular direction φ. Although this velocity is low compared with the radial velocity $v_{i\varphi} \sim (r_B/ct_0)v_{ir}$, where t_0 is the characteristic ionic time, the electric field E_φ is of fundamental importance, since only it is capable of bringing about a change in the magnetic field B_z with $\partial/\partial\varphi \equiv 0$.

In the equation for the Lagrangian invariant $I \equiv \Omega_{ez}/n_e$ (Ref. 6)

$$\frac{\partial I}{\partial t} + \mathbf{v}_e \cdot \nabla I = 0, \tag{3}$$

because of the axial symmetry of the problem the radial component of the electronic velocity

$$v_{er} = \frac{Zn_i}{n_e} v_{ir} + \frac{1}{4\pi e n_e} \frac{\partial E_r}{\partial t} \tag{4}$$

does not containg a term with the magnetic field and the radial dynamics I occurs with a characteristic ionic velocity.

Since the electric field E_φ is weak, a predominantly radial motion of the ions arises in the electronic vortex, as a result of which the magnetic field and all other quantitites in the vortex readjust. We underscore that although in the nonstationary case, for the symmetry employed, there arises an induction electric field E_φ, which brings about a change in the magnetic field B_z, the field E_φ does not appear explicitly in the final equations, as will be shown below, and it can be shown that the induction equation is equivalent to Eq. (3) for the electronic Lagrangian invariant I.

The ordinary hydrodynamic equations with zero pressure are used for the ionic component:

$$m_i \frac{d\mathbf{v}_i}{dt} = Ze\mathbf{E} + \frac{Ze}{c}[\mathbf{v}_i \times \mathbf{B}], \quad \frac{\partial n_i}{\partial t} + div(n_i \mathbf{v}_i) = 0. \tag{5}$$

The magnetic field is determined from

$$\text{curl}\, \mathbf{B} = \frac{4\pi e}{c}(Zn_i \mathbf{v}_i - n_e \mathbf{v}_e) + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \tag{6}$$

and the displacement current in the r direction is of the same order of magnitude as the ionic current.

3. Therefore, to investigate a Coulomb explosion, ion motion under the action of the electric field of a vortex will be calculated on the basis of the hydrodynamic equations in the absence of kinetic pressure. We note that now, in contrast to Ref. 6, the structure of the electronic vortex is determined substantially by the density profile of the ionic component but it does not depend on the ionic velocity.[11] Introducing the characteristic time t_0 and the characteristic ionic velocity v_0 as

$$t_0=\sqrt{\frac{m_i}{4\pi e^2 Z n_{e\infty}}} \quad \text{and} \quad v_0=c\sqrt{\frac{Zm_e}{m_i}}, \tag{7}$$

Eqs. (3)–(5) presented above can be put into the dimensionless form

$$\frac{\partial u}{\partial \tau}+u\frac{\partial u}{\partial \rho}=-\frac{\partial \gamma}{\partial \rho}-i\frac{\partial b}{\partial \rho}, \quad \frac{\partial u}{\partial \tau}+\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho n u)=0, \tag{8}$$

$$\frac{\partial i}{\partial \tau}+\left\{\frac{n}{\nu}u+\frac{1}{\nu}\frac{\partial}{\partial \tau}\left(\frac{\partial u}{\partial \tau}+u\frac{\partial u}{\partial \rho}\right)\right\}\frac{\partial i}{\partial \rho}=0. \tag{9}$$

Here and below $\tau=t/t_0$, $u=v_{ir}/v_0$, $i=I\sqrt{n_{e\infty}/4\pi m e_e c^2}$, $b=B/\sqrt{4\pi n_{e\infty} m_e c^2}$, $\rho=r\sqrt{4\pi e^2 n_{e\infty}/m_e c^2}$, and $n_{e\infty}=Zn_{i\infty}$ is the electron density at infinity. To determine the dimensionless quantities v, $\gamma=1/\sqrt{1-v^2}$, b, and ν appearing here, equations must be added to determine the structure of the vortex:[6]

$$\frac{\partial g}{\partial \rho}=\frac{\nu i-g}{\rho}-\frac{\gamma f}{\rho}-\nu\rho f, \tag{10}$$

$$\frac{\partial f}{\partial \rho}=\frac{\nu i-g}{\rho\gamma^3}-\frac{f}{\rho}, \tag{11}$$

$$\nu=\frac{\gamma^3 n+g^2+\gamma^4 f^2}{\gamma+i(g+\gamma^3 f)}, \tag{12}$$

where $g=b+\gamma f$ and $v=-\rho f$. Here, in contrast to Ref. 6, the expression for ν (12) takes account of the change in the ion density $n=n_i/n_{i\infty}$, which results in readjustment of the structure of the electronic vortex. It can be shown that the system of equations (10)–(12) determines an effective adiabat for ions — the functional $b\equiv b[n(\rho)]$, which makes it possible to close the ionic equations. The description using Eqs. (10)–(12) is applicable in the quasistatic limit $r_B/ct_0<1$ for electrons, which corresponds to $B^2<4\pi n_i m_i c^2$.

Numerical integration of the system of equations (8)–(12) revealed the formation of a collisionless shock wave with characteristic scale $\delta\sim r_B$, examined by Sagdeev.[12–14] Figure 1 displays profiles of the dimensionless Lagrangian invariant i and magnetic field b in an electronic vortex, which arise in a shock wave. At large distances the velocity of the shock wave decreases and inertial expansion of the ions remains.

The calculations show that the initial strong nonquasineutrality near the vortex axis vanishes after the passage of a nonlinear wave. It is evident from Fig. 2 that the ion and electron densities behind the wave front are identical and a quasineutral plasma is formed, and the characteristic dimension of the front is $\delta\sim r_B$ with respect to the electron density at the maximum. It is also evident that when the ions accelerate in the electric field, the wave gradually steepens. Figure 3 shows the dimensionless velocity w of the nonlinear wave and the maximum dimensionless velocity $u_{\max}$ of the ions in the wave as a function of the initial dimensionless magnetic field b_0 on the vortex axis. It follows from the figure that for the characteristic values $b_0\sim 1$ corresponding to the measurements in Ref. 5, $B\simeq 10^3$ T and $n_e\simeq 10^{19}$ cm^{-3}, the maximum dimensionless radial velocity is $u_{\max}\sim 0.7$. This gives a maximum radial velocity of the ions $V_r\sim 3.5\times 10^8$ cm/s, which agrees well with Ref. 5.

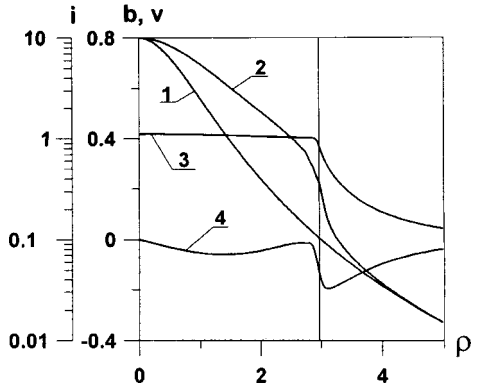

FIG. 1. Profiles of the Lagrangian invariant i (*2*), magnetic field b (*3*), and electron velocity v (*4*) in a collisionless shock wave for time $t=3.5\omega_{pi}^{-1}$. The initial profile of the Lagrangian invariant (*1*) is also displayed in the figure.

The structure of the electronic vortex is determined by the Lagrangian invariant I and the characteristic size r_0 of the vortex. The calculations showed a weak dependence of the ion energy on the Lagrangian invariant: As i varies from 5 to 130 with $r_0 = c/\omega_{pe}$, the initial magnetic field on the axis varies by less than a factor of 3, and the dimensionless velocity w of the wave varies from 0.41 to 0.95. This gives an estimate for the ion energy $\epsilon_i \simeq Zm_e c^2 u_{\max}^2/2$, which in contrast to Ref. 4 takes account of the intensity J of the laser pulse. In Ref. 15, for the above-mentioned propagation of a laser pulse along the y axis, the measurements showed that the maximum expansion velocity of the plasma occurs at an angle of 90° with respect to the direction of the laser pulse — along the x axis. This agrees well with the radial acceleration of ions in an electronic vortex. The ''tail'' there due to ions accelerated in a direction opposite to the propagation direction of the laser pulse — in the direction along the negative y axis — corresponds exactly to induction acceleration in the vortex along φ for $v_{i\varphi}<0$.

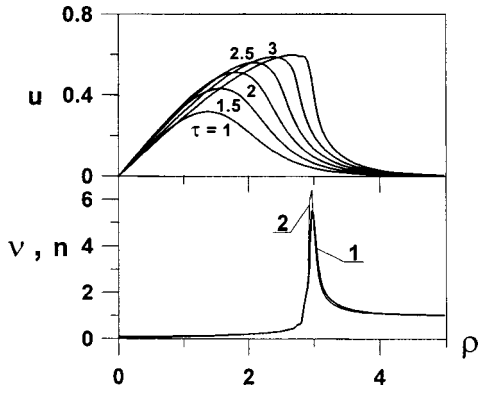

FIG. 2. Profiles of the electron ν (*1*) and ion n (*2*) densities and ion velocity (top) in a collisionless shock wave for time $t=3.5\omega_{pi}^{-1}$. The velocity profiles for earlier times are also displayed in the top panel.

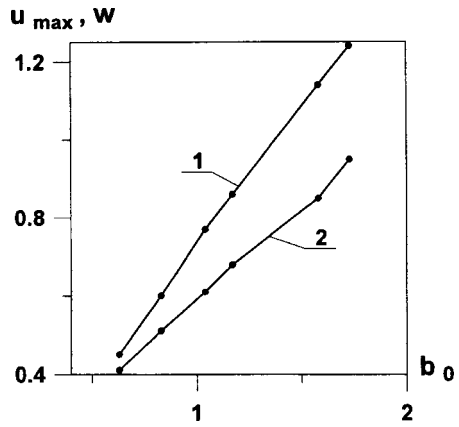

FIG. 3. Maximum ion velocity $u_{\max}$ (*1*) and velocity w (*2*) of a collisionless shock wave versus the initial magnetic field b_0 on the axis of the vortex.

4. To describe the reasons for the formation of a collisionless shock wave, we shall employ the well-known theoretical concepts of Refs. 12–14. Since I steepens, there appears on the wave front a characteristic value I_0 on a small scale near maximum density. In this case, for a planar geometry, an analytic approach is possible for large values of the magnetic field $B^2>4\pi n_e m_e c^2$, which for the parameters considered is equivalent to the condition $a>1$. Specifically, the following equation can be derived for the magnetic-field profile $B(x)$:

$$\frac{\partial}{\partial x}\left(\frac{1}{B}\frac{\partial B}{\partial x}\right)=\left(\frac{4\pi e}{I_0}\right)^2\frac{B(W-v_{ix})-n_{e\infty}I_0W}{B(W-v_{ix})}, \tag{13}$$

where W is the velocity of the nonlinear wave propagating along the x axis, and the ion velocity v_{ix} can be expressed in terms of B. It is evident that the characteristic scale of the wave front is $\delta\sim r_B$.

In the approach considered, the transition to a shock wave can be initiated by the appearance of singularities on the front of the nonlinear wave on account of Eq. (3) (see also Ref. 14).

5. A universal model has been proposed for a Coulomb explosion. A nonquasineutral electronic vortex in a magnetic field — a hole arising in the electron density in a laser plasma as a result of vorticity generation by the Weibel instability — was chosen as the initial state. Such a hole is a quasistationary intermediate state, which inevitably forms over very short periods of time of the order of ω_{pe}^{-1} in the presence of electronic vorticity and is characterized by a Lagrangian invariant I and size r_0. As the size r_0 increases, the energy of the vortex increases, and for $r_0\sim c/\omega_{pi}$ the ion velocity is of the order of the velocity of light, so that the motion of ions in the vortex must be taken into account and the quasistatic approximation breaks down.[16] The quasiequilibrium state obtained is fundamentally nonstationary at times of the order of ω_{pi}^{-1}, when the ions are accelerated in an electric field and the electronic rotation is transformed into ion motion. The calculations presented show that the velocity and angular distribution of the ions agree well with the measurements performed in Refs. 5 and 15.

This work was supported by a grant from the Russian Fund for Fundamental Research (No. 97-02-16980). We thank V. T. Tikhonchuk for bringing Ref. 15 to our attention.

[1] E. Esarey *et al.*, IEEE Trans. Plasma Sci. **24**, 252 (1996).
[2] M. Tabak, J. Hammer, M. E. Glinsky *et al.*, Phys. Plasmas **1**, 1626 (1994).
[3] K. Krushelnick, A. Ting, C. I. Moore *et al.*, Phys. Rev. Lett. **78**, 4047 (1997).
[4] N. H. Burnett and G. D. Enright, IEEE J. Quantum Electron. **26**, 1797 (1990).
[5] G. S. Sarkisov, V. Yu. Bychenkov, V. T. Tikhonchuk *et al.*, JETP Lett. **66**, 828 (1997); G. S. Sarkisov, V. Yu. Bychenkov, V. N. Novikov, and V. T. Tikhonchuk, Phys. Rev. E **59**, 7042 (1999).
[6] A. V. Gordeev and S. V. Kevchenko, JETP Lett. **67**, 482 (1998).
[7] G. S. Sarkisov, V. Yu. Bychenkov, and V. Yu. Tikhonchuk, JETP Lett. **69**, 20 (1999).
[8] E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
[9] S. V. Bulanov, M. Lontano, T. Zh. Esirkepov *et al.*, Phys. Rev. Lett. **76**, 3562 (1996).
[10] G. A. Askar'yan, S. V. Bulanov, F. Pegoraro, and A. M. Pukhov, JETP Lett. **60**, 251 (1994).
[11] A. V. Gordeev, A. S. Kingsep, and L. I. Rudakov, Phys. Rep. **243**, 215 (1994).
[12] R. Z. Sagdeev, in *Plasma Physics and the Problem of Controlled Thermonuclear Reactions*, Vol. 4 (Pergamon Press, New York, 1959) [Russian original, Soviet Academy of Sciences Press, Moscow, 1958, p. 384).
[13] R. Z. Sagdeev, in *Reviews of Plasma Physics*, pp. 23–91, edited by M. A. Leontovich, (Consultants Bureau, New York, 1966) [Russian original, Atomizdat, Moscow, 1964].
[14] V. V. Krasnosel'skikh, Zh. Éksp. Teor. Fiz. **89**, 498 (1985) [Sov. Phys. JETP **62**, 282 (1985)].
[15] K. Krushelnick, E. L. Clark, Z. Najmudin *et al.*, Phys. Rev. Lett. **83**, 737 (1999).
[16] A. V. Gordeev, S. V. Levchenko, Electromagnetic Waves and Electronic System, **3**, 25 (1998).

Translated by M. E. Alferieff

Dynamical correlations of two-dimensional vortex-like defects

V. V. Lebedev*)

Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117940 Moscow, Russia; Department of Physics of Complex Systems, Weizmann Institute of Science, 76100 Rehovot, Israel

(Submitted 27 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 675–679 (25 November 1999)

High-order dynamical correlations of defects (quantum vortices, disclinations, etc.) in thin films are examined by starting from the Langevin equation for the defect motion. It is demonstrated that the dynamical correlation functions F_{2n} of the vorticity or disclinicity behave as $F_{2n} \sim y^2/r^{4n}$, where r is the characteristic scale and y is the renormalized fugacity. Therefore below the Berezinskii–Kosterlitz–Thouless transition temperature the F_{2n} are characterized by anomalous scaling exponents. The behavior differs strongly from the normal law $F_{2n} \sim F_2^n$ obeyed by equal-time correlation functions; the unequal-time correlation functions appear to be much larger. The phenomenon resembles intermittency in turbulence. © *1999 American Institute of Physics.* [S0021-3640(99)00822-1]

PACS numbers: 68.55.Ln

It is well known that defects such as quantum vortices, spin vortices, dislocations and disclinations play an essential role in physics of low-temperature phases of thin films. Berezinskii[1] and then Kosterlitz and Thouless[2] recognized that two-dimensional (2D) systems have a class of phase transitions related to defects. The main idea of their approach is that in 2D the defects can be treated as point objects interacting like charged particles. The low-temperature phase corresponds to a fluid consisting of bound uncharged defect–antidefect pairs, which is an insulator, whereas the high-temperature phase contains free charged particles and can be treated as a plasma. Correspondingly, in the low-temperature phase the correlation length is infinite, whereas in the high-temperature phase it is finite. An enormous number of papers have been devoted to different aspects of the problem (see, e.g., the surveys in Refs. 3–7). The scheme proposed by Kosterlitz and Thouless can be applied to superfluid and hexatic films and planar 2D magnets. It admits a generalization for crystalline films (see Refs. 8 and 9). There are also applications to superconductive materials, especially to high-T_c superconductors (see, e.g., Ref. 10).

The dynamics of films in the presence of such defects was considered in Refs. 11 and 12. In those papers a complete set of equations was formulated describing both the motion of the defects and the hydrodynamic degrees of freedom. Then, to obtain macroscopic dynamical equations, an averaging over an intermediate scale was performed. In

that procedure the ''current density'' related to the defects was replaced by an expression proportional to the average ''electric field'' and to gradients of the temperature and of the chemical potential. The resulting equations perfectly correspond to the problems solved in those papers.[11,12] Unfortunately, in that procedure information concerning high-order correlations of the defect motion is lost. That is the motivation for the present study, in which we shall examine these high-order correlations.

The static properties of the system of vortex-like defects in thin films can be described quite universally. The starting point of the description is the free energy $\mathscr{F}$ associated with the defects:

$$\mathscr{F}/T=-\sum_{i\neq j}\beta n_i n_j \ln\left(\frac{|x_i-x_j|}{a}\right)-\sum_j \ln y, \tag{1}$$

where the subscripts i,j label defects, x_i are the positions of the defects, $n_j=\pm 1$ are the ''charges'' of the defects, a is a cutoff parameter of the order of the size of the defect core, and the coupling constant β and the fugacity y are dimensionless T-dependent factors. Expression (1) is correct for quantum vortices in superfluid films, for disclinations in hexatic films, and for spin vortices in 2D planar magnets. For dislocations in crystalline films, expression (1) has to be modified slightly.[8]

The presence of the pairs in the system leads to nontrivial ''dielectric'' properties of the medium. As a result, the interaction between the charges is modified; the effect can be described in terms of a scale-dependent coupling constant β (Ref. 2). The dependence can be found in the framework of the scheme proposed by Kosterlitz.[13] Excluding pairs with separations from a to r, we arrive at renormalized values of the parameters β and y which obey the following renormalization-group equations:

$$\frac{d\beta}{d\ln(r/a)}=-cy^2, \quad \frac{dy}{d\ln(r/a)}=(2-\beta)y, \tag{2}$$

where c is a numerical factor of order unity. In the low-temperature phase, the coupling constant β tends to a constant on large scales. The asymptotic value of β is larger than 2; the critical value $\beta_c=2$ corresponds to the transition temperature. In the asymptotic region, where β can be treated as r-independent, the renormalized fugacity y remains r-dependent. Its asymptotic behavior can easily be extracted from Eq. (2): $y\propto r^{2-\beta}$. We see that in the low-temperature phase y tends to zero as the scale increases. Thus the inequality $y\ll 1$ is satisfied for large scales in the low-temperature phase and probably in some region of scales above T_c.

We consider the correlation functions

$$F_{2n}(t_1,\ldots,t_{2n};x_1,\ldots,x_{2n})=\langle\rho(t_1,x_1)\ldots\rho(t_{2n},x_{2n})\rangle \tag{3}$$

of the ''charge density''

$$\rho(x)=\sum_j n_j\delta(x-x_j). \tag{4}$$

For superfluid films the ''charge density'' (4) is proportional to the vorticity $\nabla\times\mathbf{v}_s$. In statics, one gets the estimate[14]

$$F_2(r)\sim y^2(r)/r^4, \tag{5}$$

where $r=|x_1-x_2|$. For high-order equal-time correlation functions the normal estimate $F_{2n}\sim F_2^n$ is valid under the condition $y\ll 1$.[15]

Following Ref. 11, we adopt the stochastic equation

$$\frac{dx_{j,\alpha}}{dt}=-\frac{D}{T}\left[\frac{\partial \mathscr{F}}{\partial x_{\alpha j}}+n_j\gamma\epsilon_{\alpha\beta}\frac{\partial \mathscr{F}}{\partial x_{\beta j}}\right]+\xi_{j,\alpha}, \tag{6}$$

which determines the trajectory of the jth vortex in a superfluid film. Here $\mathscr{F}$ is the free energy (1), D is a diffusion coefficient, γ is a dimensionless parameter, and ξ_j are Langevin forces with the correlation function

$$\langle \xi_{i,\alpha}(t_1)\xi_{j,\beta}(t_2)\rangle = 2D\delta_{ij}\delta_{\alpha\beta}\delta(t_1-t_2). \tag{7}$$

Equation (6) can be derived in the spirit of the procedure proposed by Hall and Vinen for a 3D superfluid (see Ref. 16). Huber[17] argued that the same equation is correct for spin vortices in planar 2D magnets. We believe that for $\gamma=0$, Eq. (6) is applicable to the dynamics of disclinations in hexatic films such as membranes, freely suspended films, and Langmuir films. We should also take into account annihilation and creation processes. They are characterized by a creation rate $\overline{R}(r)$, which is the probability density for a defect–antidefect pair with separation r to be created per unit time per unit area, and by an annihilation rate $R(r)$, which is the probability for a defect–antidefect pair to annihilate per unit time if the pair is separated by a distance r. Actually $\overline{R}(r)$ and $R(r)$ are nonzero only if r is of the order of the core size a.

To examine the correlation functions (3) we use the Doi technique,[18] which enables one to treat systems of classical particles in which creation and annihilation processes occur. The Doi technique is formulated in terms of creation $\hat{\psi}$ and annihilation ψ operators which satisfy the same commutation rules as do those for Bose particles; we must introduce the fields $\psi_{\pm}$ and $\hat{\psi}_{\pm}$ corresponding to defects and antidefects, respectively. An effective Hamiltonian can be written in terms of the fields. Then one can formulate a conventional diagrammatic expansion in terms of R, $\overline{R}$, β, and the propagators of defects. Extracting blocks with small separations between the defects, one can pass to renormalized quantities. Details of the procedure will be published elsewhere.[19] Here we only present the results. The fugacity is expressed in terms of the renormalized values of the creation and annihilation rates as

$$\overline{R}(x)=\frac{y^2}{r^4}R(x). \tag{8}$$

The renormalized value of the annihilation rate is

$$R(x)=8\pi\beta\delta(x). \tag{9}$$

The renormalized values of β and y obey the same renormalization-group equations (2) as in statics. For D and γ we get the following renormalization-group equations:

$$\frac{dD}{d\ln(r/a)}\sim -y^2, \quad \frac{d\gamma}{d\ln(r/a)}\sim y^2, \tag{10}$$

which are analogous to the equation (2) for β. We conclude that the corrections to D and γ are small (and irrelevant) due to the small value of the fugacity.

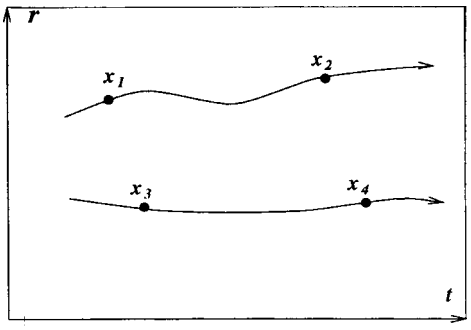

FIG. 1. Two trajectories passing through given points.

Next, we can examine the correlation functions (3), rewriting the charge density (4) as

$$\rho = \hat{\psi}_+ \psi_+ - \hat{\psi}_- \psi_- . \tag{11}$$

One should distinguish between contributions related to different numbers k of defect–antidefect pairs passing through the points x_m at given times t_m. They can be estimated as

$$F_{2n} \sim y^{2k}(r_*) r_*^{-4n}, \tag{12}$$

where we assume that all spatial separations are of the order of r_* and all time intervals are of the order of r_*^2/D. We see that the ratio in (12) contains the $2k$th power of a dimensionless small parameter y. Thus we conclude that the leading contribution to F_{2n} is due to a single defect–antidefect pair, which corresponds to $k=1$. The situation is illustrated in Fig. 1. Though the contribution associated with a number of defect–antidefect pairs contains an additional huge entropy factor, it has also an additional small factor associated with the small probability of observing defect–antidefect pairs with separations larger than the core radius; the smallness is due to the strong Coulomb attraction. In the large-scale limit when β is saturated we have $F_{2n} \propto r_*^{-4(n-1)-2\beta}$. If some spatial separations among $|r_i - r_j|$ and/or some time intervals differ strongly, then one can formulate some simple rules. For example, if one of the time intervals τ is much larger than all values of $|r_i - r_j|^2/D$, then the correlation function behaves as $F_{2n} \propto \tau^{-\beta}$.

For the pair correlation function we have the same estimate (5) as in statics. The high-order correlation functions are much larger than their normal estimates in terms of the pair correlation function. Namely, in accordance with Eqs. (5) and (12) we have

$$F_{2n}/F_2^n \sim y^{-2n+2} \gg 1. \tag{13}$$

Let us explain in terms of our scheme the origin of the estimate $F_{2n} \sim F_2^n$ for the equal-time correlation functions. This estimate corresponds to $k=n$ in Eq. (12). The reason is that two defects cannot pass simultaneously through $2n$ points, and at least $k=n$ defect–antidefect pairs must be taken in order to obtain a nonzero contribution to the equal-time correlation function F_{2n}. The situation is illustrated in Fig. 2. Thus we have two different regimes: for equal-time and for unequal-time correlation functions. To establish the boundary between the regimes we should consider small time intervals in which the

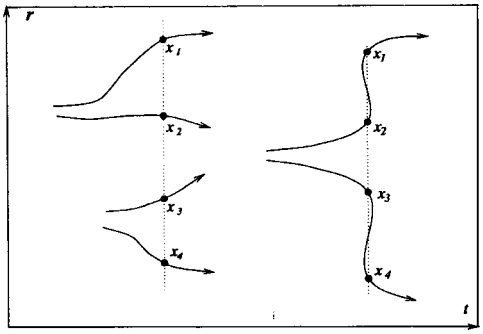

FIG. 2. Possible and impossible trajectories passing through four points at a given moment in time.

single-pair contribution is finite but small. The smallness is associated with diffusive ''smearing.'' The characteristic time over which the simultaneous (equal-time) regime passes into the nonsimultaneous (unequal-time) regime can be estimated from $Dt \sim r^2/|\ln[y(r)]|$, where r is the characteristic spatial separation.

The effect considered resembles intermittency in turbulence (see, e.g., Ref. 20), which leads to large r-dependent factors in the ratios like (13) in the velocity correlation functions of a turbulent flow. However, in the case of defects the large r-dependent factors are related to the ultraviolet cutoff parameter, whereas for intermittency in turbulence the large r-dependent factors are related to the infrared (pumping) scale. Our situation is thus closer to the inverse cascade (see Ref. 21) realized on scales much larger than the pumping length. Our results can also be compared with nontrivial tails of probability distribution functions in the physics of disorder materials (see, e.g., Refs. 22 and 23).

I am grateful to G. Falkovich, K. Gawedzki, I. Kolokolov, and M. Vergassola for helpful discussions and to E. Balkovsky, A. Kashuba, and S. Korshunov for valuable remarks. This research was supported in part by a grant from the Israel Science Foundation, by a grant from the Minerva Foundation, and by the Landau–Weizmann Prize program.

*)e-mail: lebede@landau.ac.ru and lwlebede@wicc.weizmann.ac.il

[1] V. L. Berezinskii, Zh. Éksp. Teor. Fiz. **59**, 907 (1970) [Sov. Phys. JETP **32**, 493 (1971)]; Zh. Éksp. Teor. Fiz. **61**, 1144 (1971) [Sov. Phys. JETP **34**, 610 (1972)].
[2] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **5**, L124 (1972); **6**, 1181 (1973).
[3] J. M. Kosterlitz and D. J. Thouless, in *Progress in Low Temperature Physics*, Vol. VII B, edited by D. F. Brewer, North-Holland, Amsterdam (1978), p. 373.
[4] D. R. Nelson, in *Fundamental Problems in Statistical Mechanics*, Vol. V, edited by E. G. D. Cohen, North-Holland, New York (1980), p. 53.
[5] D. R. Nelson, in *Phase Transitions and Critical Phenomena*, Vol. 7, edited by C. Domb and J. L. Lebowitz, Academic Press, London (1983), p. 1.
[6] P. Minnhagen, Rev. Mod. Phys. **59**, 1001 (1987).
[7] Z. Gulacsi and M. Gulacsi, Adv. Phys. **47**, 1 (1998).
[8] D. R. Nelson and J. M. Kosterlitz, Phys. Rev. Lett. **39**, 1201 (1977); B. I. Halperin and D. R. Nelson, Phys. Rev. Lett. **41**, 121 (1978); B. I. Halperin and D. R. Nelson, Phys. Rev. Lett. **41**, 519 (1978); A. P. Young, Phys. Rev. B **19**, 1855 (1979); D. R. Nelson and B. I. Halperin, Phys. Rev. B **19**, 2457 (1979).

[9] K. J. Strandburg, Rev. Mod. Phys. **60**, 161 (1988).
[10] G. Blatter, M. V. Feigel'man, V. B. Geshkenbein *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
[11] A. Ambegaokar, B. I. Halperin, D. R. Nelson, and E. D. Siggia, Phys. Rev. Lett. **40**, 783 (1978); Phys. Rev. B **21**, 1806 (1980).
[12] A. Zippelius, B. I. Halperin, and D. R. Nelson, Phys. Rev. B **22**, 2514 (1980); R. Bruinsma, B. I. Halperin, and A. Zippelius, Phys. Rev. B **25**, 579 (1982).
[13] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[14] J. V. Jose, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, Phys. Rev. B **16**, 1217 (1977).
[15] L. P. Kadanoff, Ann. Phys. **120**, 39 (1979); L. P. Kadanoff and A. B. Zisook, Nucl. Phys. B **180**, 61 (1981).
[16] H. E. Hall and W. F. Vinen, Proc. R. Soc. London, Ser. A **238**, 215 (1956); W. F. Vinen, in Progress in *Low Temperature Physics*, Vol. III, North-Holland, Amsterdam (1961), p. 1; P. Nozieres and W. F. Vinen, Philos. Mag. **14**, 667 (1966).
[17] D. L. Huber, Phys. Rev. B **26**, 3758 (1982).
[18] M. Doi, J. Phys. A: Math. Gen. **9**, 1465, 1479 (1976).
[19] V. V. Lebedev, Phys. Rev. E, submitted.
[20] U. Frisch, *Turbulence: the Legacy of A. N. Kolmogorov*, Cambridge University Press, New York (1995).
[21] R. Kraichnan, Phys. Fluids **10**, 1417 (1967); R. Kraichnan, J. Fluid Mech. **47**, 525 (1971).
[22] I. Lifshits, S. Gredeskul, and A. Pastur, *Introduction to the theory of disordered systems*, Wiley Interscience, New York (1988).
[23] K. B. Efetov, *Supersymmetry in Disorder and Chaos*, Cambridge University Press (1997).

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

Low-temperature magnetic properties of the 2D molecular ferrimagnet $N(n\text{-}C_5H_{11})_4\,[Fe^{II}Fe^{III}(C_2O_4)_3]$

L. Bottyán
KFKI, Research Institute for Particle and Nuclear Physics, 1525 Budapest, P. O. B. 49, Hungary

L. Kiss
Institute for Solid State Physics and Optics, 1525 Budapest, P. O. B. 49, Hungary

N. S. Ovanesyan*), A. A. Pyalling, and N. A. Sanina
Institute of Problems of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

A. B. Kashuba
L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 14 October 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 680–683 (25 November 1999)

^{57}Fe Mössbauer and magnetometric studies of the molecular ferrimagnet $N(n\text{-}C_5H_{11})_4\,[Fe^{II}Fe^{III}(C_2O_4)_3]$ are indicative of a 2D magnetic character with strong uniaxial anisotropy in the basal plane of the crystal. It is established that the change in the sign of the net magnetization of this compound is related to a compensation between Fe^{III} and Fe^{II} sublattice magnetizations at $T_{comp}=31.2$ K. The form and parameters of the magnetic Hamiltonian describing the temperature dependence of the Fe^{III} sublattice and the net magnetizations are determined. © *1999 American Institute of Physics.* [S0021-3640(99)00922-6]

PACS numbers: 76.80.+y, 75.50.Gg, 75.60.Ej

In the search for new molecular magnets, considerable attention in the last decade has been paid to the synthesis of bimetallic transition-metal complexes with spontaneous magnetization. Okawa *et al.* suggested the use of the *tris*-oxalates of chromium $[Cr^{III}(ox)_3]^{3-}$ and iron $[Fe^{III}(ox)_3]^{3-}$, with D_3 symmetry, as building blocks (ox $=(C_2O_4)^{2-}$). This kind of synthesis led to the discovery of a whole family of ferromagnetic and ferrimagnetic materials of the form $NBu_4\,[M^{II}M^{III}(ox)_3]$, where M^{II}=Mn, Fe, Co, Ni, Cu; Bu=n-C_4H_9 (Refs. 1 and 2). These and similar compounds with different organic cations crystallize in a trigonal structure with space group $R3c$ (Refs. 3 and 4). The honeycomb-like 2D lattice, with the two different magnetic ions in the corners of the hexagon, form alternating layers $[M^{II}M^{III}(ox)_3]_n^{n-}$ separated by the organic cations. In such a structure, ferro- and ferrimagnetism can both be brought about by the intralayer superexchange interaction along the path $M^{II}-(ox)^{2-}-M^{III}$. The interlayer interaction is weak on account of the considerable distance between layers, which is determined by the organic cation ($d=8$ to 15 Å; Refs. 5 and 6). Consequently, these compounds are well represented by a stack of two-dimensional magnets.

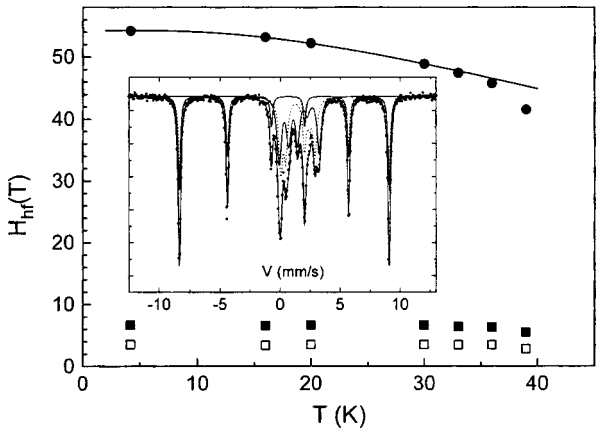

FIG. 1. Temperature dependences of Fe^{III} (●) and two Fe^{II} (■ and □) hyperfine fields in NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$. The solid line represents the Fe^{III} magnetization derived from the model Hamiltonian (*1*). The inset shows an example Mössbauer spectrum at 4.2 K fitted with the combined magnetic-quadrupole hyperfine total Hamiltonian.

Here we report on Mössbauer and magnetometric investigations of ferrimagnetic NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$, where $Pn=n\text{-}C_5H_{11}$, in the temperature range 4.2–100 K (the Néel temperature $T_N=48$ K). This compound is of particular interest in respect to the reported negative sign of the magnetization.[7] We have fitted the temperature dependences of the Fe^{III} and Fe^{II} sublattice magnetizations in the framework of a relevant microscopic model and have determined an exchange constant, magnetic anisotropy constants, and the *g* factor of Fe^{II}. It is shown that this compound is an *N*-type ferrimagnet according to the Néel classification.[8]

The polycrystalline compound was synthesized by the standard method.[6] Magnetization measurements were performed by a Quantum Design MPMS-5S SQUID magnetometer in the reciprocating-sample (RSO) mode at external magnetic fields and temperatures in the ranges 1 mT to 5 T and 5 to 100 K. A powder specimen of 58 mg was compacted into the sample holder. ^{57}Fe Mössbauer spectra were recorded at various temperatures using a He gas-flow cryostat (Oxford Instruments CF506).

The basic distortion of the coordination octahedron in the metal-oxalate layers of NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$ is compression along the trigonal axis, parallel to the crystallographic *c* axis. In addition, a small rhombic distortion is present, reducing the space group from $R3c$ to $C222_1$ and the site symmetry of both Fe^{II} and Fe^{III} from D_3 to C_2 (Ref. 9). The distance between nearest (Fe^{II} and Fe^{III}) ions in neighboring layers is $d=10.23$ Å.[5]

The ^{57}Fe Mössbauer spectrum of the compound at $T=4.2$ K is displayed in the inset of Fig. 1. It shows the magnetic hyperfine structure of Fe^{III} ($S=5/2$) and Fe^{II} ($S=2$) ions, with an intensity ratio of 1:1. The spectra are analyzed in terms of the electron–nuclear Hamiltonian including the magnetic hyperfine and electric quadrupole interactions. We find that the *z* component of the electric field gradient (EFG) V_{zz} is *negative* at Fe^{II}, which corresponds to the trigonal *compression* of the oxygen octahedron along the *c* axis and a nonzero asymmetry parameter $(V_{xx}-V_{yy})/V_{zz}\leqslant 0$, which manifests the rhombic distor-

tion. The angles θ^{II} and θ^{III} between the principal z direction of the EFG and the hyperfine fields at the nuclei of Fe^{II} and Fe^{III} are both 90°, i.e., the magnetization is confined to the basal plane of the crystal. The hyperfine field H_{hf} at the nucleus of Fe^{II} is lowered considerably due to a significant unquenched orbital moment contribution.[10] Moreover, there exist two nonequivalent sites for the Fe^{II} ion, as is indicated by the two slightly different values of H_{hf} in Fig. 1. The intensity ratio of the Mössbauer lines corresponds to a random orientation of the easy axes of magnetization of the crystallites in the polycrystalline sample.

The temperature dependence of the hyperfine fields $H_{hf}(T)$ at Fe^{II} and Fe^{III} is shown in Fig. 1. At the Fe^{II} site H_{hf} is practically independent of temperature, indicating an identical temperature dependence of the spin and orbital components of the magnetic moment. In order to describe the temperature dependence $H_{hf}(T)$ at the Fe^{III} site (proportional to the Fe^{III} sublattice magnetization), the following form of the 2D microscopic Hamiltonian of a single layer of NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$ is assumed:

$$\hat{H}=J\sum_{ab}\mathbf{S}_a\cdot\mathbf{S}_b+\frac{\lambda_z}{2}\sum_b(S_b^z)^2+\frac{\lambda_y}{2}\sum_b(S_b^y)^2-2\mu_B\sum_a\mathbf{S}_a\cdot\mathbf{H}-g_b\mu_B\sum_b\mathbf{S}_b\cdot\mathbf{H}, \tag{1}$$

where the indices a and b run over the Fe^{III} and Fe^{II} sublattices, respectively. We have estimated the variation of the parameters in the Hamiltonian (1) for the two nonequivalent Fe^{II} sites and found it to be negligible.

The first term of Hamiltonian (1) is the antiferromagnetic exchange interaction of strength J. The second and third terms of the Hamiltonian describe the magnetic anisotropy on the Fe^{II} sublattice, due to the orbital moment of the Fe^{II} ion.[7] The magnetic anisotropy at the Fe^{III} site is neglected, since its orbital moment $L=0$. The last two terms are the Zeeman interactions with the external field. For the Fe^{III} ion, the g factor has the value 2. An effective spin-wave Hamiltonian is obtained from (1) by a standard Holstein–Primakoff transformation. The temperature dependence of the magnetization of the two sublattices was calculated without taking into account the interaction between spin waves. The best agreement with the experimental data for $H_{hf}(T)$ at Fe^{III} is achieved with the following values of the parameters in the Hamiltonian (1):

$$J=12.4(2)\ \mathrm{K},\quad \lambda_y=9.6(8)\ \mathrm{K},\quad \lambda_z=20-50\ \mathrm{K}. \tag{2}$$

The value of the λ_z parameter has little effect on the Fe^{III} magnetization.

SQUID magnetometry of the sample was performed in the following way. First, the sample was cooled from $T=100\,\mathrm{K}$ to $T=5\,\mathrm{K}$ in external magnetic fields ranging from $H=1$ mT to $H=5$ T, and then the magnetization, $M(T,H)$ (where H is the cooling field) was recorded in a measuring field $H_{meas}=10\,\mathrm{mT}$ as a function of increasing temperature. The experimental results, displayed in Fig. 2, have two prominent features. First, the magnetization of the sample changes sign at around $T_{comp}=31.2\,\mathrm{K}$. Second, the polycrystalline specimen did not reach magnetic saturation at 5 K and hence it is frozen in a magnetically metastable state with a certain domain structure. Those domains that coincide with crystallites of the specimen hold their magnetic orientation upon heating even if the direction of their magnetization is opposite to H_{meas}. The net magnetization of such a domain follows the *microscopic* magnetization $M(T)$. In the case of a specimen having single domain structure within each crystallite, one can extract this microscopic magne-

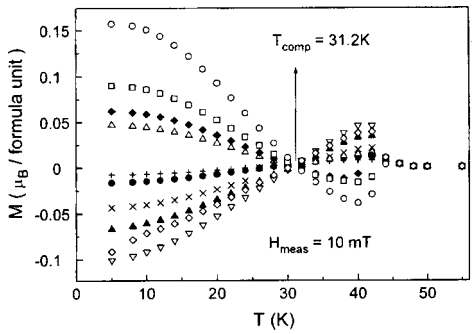

FIG. 2. Temperature dependence of the net magnetization as measured by a SQUID in NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$ cooled in various external magnetic fields: 15 mT(+), 20 mT(●), 50 mT(◇), 0.1 T(∇), 0.2 T(▲), 0.5 T(×), 0.8 T(Δ), 1 T(◆), 2 T(□) and 5 T(○). The measuring field was 10 mT, throughout.

tization $M(T)$ by normalizing the magnetization curve to its value at $T=0$. In Fig. 3 we have plotted seven normalized curves $M(T,H)/M(5\,\mathrm{K},H)$ for seven cooling magnetic fields: 5 T, 2 T, 0.8 T, 0.5 T, 0.2 T, 0.1 T and 50 mT. The magnetization curve $M(T)$ for a magnet described by Hamiltonian (1) was calculated with the parameters (2), leaving the Fe^{II} g factor as a variable parameter for fitting. The solid line shown in Fig. 3 was obtained with $g_b=2.34(2)$. Note that the $M(T,H)/M(5\,\mathrm{K},H)$ curves for strong cooling fields approach a limiting curve which we identify with the microscopic $M(T)$. To cross-check this result we calculated the saturation magnetization for our specimen, assuming random directions of the c axis in the crystallites, and obtained a value of $M(0)=0.16(2)\mu_B$ per Fe^{III}–Fe^{II} pair.

On the other hand, the considerable deviation of the $M(T,H)/M(5\,\mathrm{K},H)$ curves from $M(T)$ for weaker cooling fields can be interpreted as a ''freezing-out'' of the domain wall motion within the crystallites. We mention also that in two cooling fields, 0.5 T and 0.2 T, the dc and RSO signals for the experimental curves $M(T,H)$ are not identical below T_{comp}. The apparent convergence of all $M(T,H)/M(5\,\mathrm{K},H)$ curves in

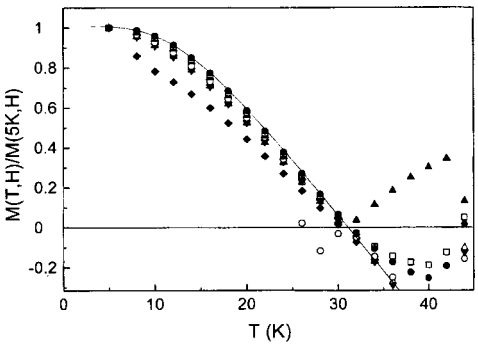

FIG. 3. Normalized magnetization curves of NPn_4 $[Fe^{II}Fe^{III}(ox)_3]$ cooled to 5 K in external magnetic fields ranging from 50 mT(+) to 5 T(○). The solid line represents a fitting curve derived from Hamiltonian (*1*) with parameters (*2*) and $g_b=2.34(2)$.

the vicinity of T_{comp} illustrate the fact that two domains of opposite order parameter have identical magnetization at T_{comp}.

We surmise that in cooling fields $H>0.1$ T and for $T<T_{comp}$ the sample is magnetized up to its equilibrium magnetization. In strong cooling fields $H>0.8$ T most single-domain crystallites reverse their magnetization during the cooling below T_{comp}, whereas in weaker cooling fields $H<0.8$ T only a few crystallites reverse their magnetization completely, so that many domain walls inside the crystallites remain to the end of the cooling at 5 K. In very weak cooling fields $H<0.1$ T the crystallites do not reach magnetic saturation even at $T>T_{comp}$, and with decreasing cooling fields the magnetization of the sample decreases. Recent external magnetic field Mössbauer experiments have shown that even $H_{ext}=7$ T is not sufficient to polarize the powder sample.[11] This result is again in agreement with the large value of the fitted anisotropy constants, λ_y and λ_z, and the large orbital contribution to H_{hf} at Fe^{II}.

This work was supported in part by the Hungarian Scientific Research Fund (OTKA) under Contract T022559 and by the Russian Fund for Fundamental Research, Project 99-03-32581.

*)e-mail: ovanesyan@icp.ac.ru

[1] H. Tamaki, Z. J. Zhong, N. Matsumoto *et al.*, J. Am. Chem. Soc. **114**, 6974 (1992).
[2] H. Okawa, N. Matsumoto, H. Tamaki, and M. Ohba, Mol. Cryst. Liq. Cryst. **233**, 257 (1993).
[3] L. O. Atovmyan, G. V. Shilov, R. N. Lyubovskaya *et al.*, JETP Lett. **58**, 766 (1993).
[4] S. Decurtins, H. W. Schmale, H. R. Oswald *et al.*, Inorg. Chim. Acta **216**, 65 (1994).
[5] C. Mathoniere, C. J. Nutall, S. G. Carling, and P. Day, Inorg. Chem. **35**, 1201 (1996).
[6] G. V. Shilov, L. O. Atovmyan, N. S. Ovanesyan *et al.*, Russian J. Coord. Chem. **24**, 288 (1998).
[7] C. J. Nutall and P. Day, Chem. Mater. **10**, 3050 (1998).
[8] J. L. Néel, Ann. Phys. (Paris) **3**, 137 (1948).
[9] S. G. Carling, C. Mathoniere, P. Day *et al.*, J. Chem. Soc. Dalton Trans. 1839–1843 (1996).
[10] N. S. Ovanesyan, G. V. Shilov, L. O. Atovmyan *et al.*, Mol. Cryst. Liq. Cryst. **273**, 175 (1995).
[11] N. S. Ovanesyan, G. V. Shilov, N. A. Sanina *et al.*, Mol. Cryst. Liq. Cryst. (1999), in press.

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

Relativistic quantum coin tossing

S. N. Molotkov and S. S. Nazin
Institute of Solid-State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 28 September 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 10, 684–689 (25 November 1999)

A relativistic quantum exchange protocol enabling ''coin tossing at a distance'' between two participants is proposed. The exchange protocol is based on the fact that to distinguish a pair of orthogonal states with certainty in relativistic quantum mechanics requires a finite time that depends on the structure of the states themselves. © *1999 American Institute of Physics.* [S0021-3640(99)01022-1]

PACS numbers: 03.67.Dd, 03.67.Hk

The coin tossing protocol is one of the simplest cryptographic protocols and reduces to the following. Let us assume that two participants A and B, who mistrust one another, wish to cast lots, for example, using a coin or a random-number generator that generates 0 or 1 with equal probability. If a 0 occurs, then A wins; if a 1 occurs, then B wins. If A and B are at the same location, the problem is trivial. However, if A and B are separated in space and can communicate only along a communication channel, then the problem can even appear to be unsolvable, since A and B can always cheat for their own benefit. The case where A and B can exchange information only along a classical communication channel was solved by Blum.[1] In the strict sense the protocol of Ref. 1 is not secret with respect to cheating by one of the participants, since it is based on the unproved computational complexity of calculating a discrete logarithm.[1] For example, if one participant has a quantum computer (which, it is true, has not yet been implemented experimentally) at his disposal, then he would always be able to win because of the fast calculation of a discrete logarithm.[2,3]

If besides a classical communication channel there is also a quantum communication channel between A and B, then various quantum exchange protocols whose secrecy is based not on computational complexity but rather on the fundamental laws of nature (quantum mechanics) are possible. Protocols for quantum key distribution,[4–6] a quantum protocol for quantum bit commitment,[7–9] and protocols for quantum coin tossing,[10] quantum gambling,[11] and quantum secret sharing[12] have been proposed and investigated. It has been shown that ideal quantum coin tossing on the basis of nonrelativistic quantum mechanics is impossible.[13,14] An ideal protocol is one where the probability of both participants acknowledging that no cheating has occurred is strictly 1 and the probability of 0 or 1 occurring is 1/2. However, a protocol where the probability of both participants acknowledging the lots to be honest is as close to 1 as desired is possible.[10]

A bit commitment protocol and a coin tossing protocol that take into account the

existence of a finite signal propagation velocity has been proposed recently.[15] The information carriers in this scheme are classical states. The protocols are based on the fact that the participants A and B each possess two spatially separated sites which they control (see details in Ref. 15).

An example of a real-time relativistic quantum coin tossing protocol will be presented below.

All quantum cryptographic protocols in one way or another employ two facts. One is the no-cloning theorem[16] — the impossibility of copying an unknown quantum state, i.e., the impossibility of the process

$$|A\rangle|\psi\rangle \to U(|A\rangle|\psi\rangle) = |B_\psi\rangle|\psi\rangle|\psi\rangle,$$

where $|A\rangle$ and $|B_\psi\rangle$ are states of the apparatus before and after copying and U is a unitary operator. This process is forbidden because of linearity and unitary evolution in quantum mechanics. An even weaker process is impossible — it is impossible to obtain information about one state in a pair of nonorthogonal states without perturbing the states:[17]

$$|A\rangle|\psi_1\rangle \to U(|A\rangle|\psi_1\rangle) = |A_{\psi_1}\rangle|\psi_1\rangle, \quad |A\rangle|\psi_2\rangle \to U(|A\rangle|\psi_2\rangle) = |A_{\psi_2}\rangle|\psi_2\rangle.$$

It is impossible for $|A_{\psi_1}\rangle \neq |A_{\psi_2}\rangle$ to occur if $\langle\psi_1|\psi_2\rangle \neq 0$, i.e., it is impossible to distinguish nonorthogonal states with certainty. No such prohibition exists for orthogonal states. For this reason, almost all cryptographic protocols employ nonorthogonal states as information carriers. An exception is the protocol proposed in Ref. 18.

Orthogonal states are distinguishable with certainty, and in the nonrelativistic quantum-mechanical theory of measurements two orthogonal states can be distinguished instantaneously. Basically for these reasons it is impossible to construct cryptographic protocols based on orthogonal states in nonrelativistic quantum mechanics.

The situation is different in relativistic quantum field theory. The physical observables of a field referring to points separated by a space-like interval cannot be linked by a cause–effect coupling, and the commutator of the field operators vanishes outside the light cone[19]

$$[u^-(\hat{x}_1), u^+(\hat{x}_2)]_\pm = -iD^-(\hat{x}_1 - \hat{x}_2), \tag{1}$$

where $u^\pm(\hat{x})$ are field operators, $\hat{x}_{1,2}$ are points of a four-dimensional space–time, and $D^-(\hat{x}_1 - \hat{x}_2)$ is the negative-frequency commutator function.[19] The latter circumstance imposes a restriction on the time required to distinguish (in one measurement) a pair of orthogonal states with certainty.

Before presenting the protocol, we shall discuss the states and measurements employed in the protocol. Any single-particle state of a field can be represented in the form

$$|\psi_{1,2}\rangle = \int \psi_{1,2}(\hat{p})\,\delta(\hat{p}^2 - m^2)u^+(\hat{p})\,d\hat{p}|0\rangle, \tag{2}$$

where the integration extends over the mass shell, $\psi_{1,2}(\hat{p})$ is the amplitude of the field, and $|0\rangle$ is the vacuum state. In what follows we shall deal with massless particles (for

example, photons). In this case, we shall interpret the field operator $u^+(\hat{p})$ as the operator creating a photon in the Coulomb gauge. We shall also assume that the amplitudes $\psi_{1,2}(\hat{p})$ are such that the states $|\psi_{1,2}\rangle$ are orthogonal:

$$\begin{aligned}\langle\psi_1|\psi_2\rangle &= \int\int \psi_1^*(\mathbf{p}')\psi_2(\mathbf{p})\langle 0|u^-(\mathbf{p}')u^+(\mathbf{p})|0\rangle \frac{d\mathbf{p}'d\mathbf{p}}{\sqrt{2p_0'}\sqrt{2p_0}} \\ &= \int \psi_1^*(\mathbf{p})\psi_2(\mathbf{p})\frac{d\mathbf{p}}{2p_0} = 0, \\ [u^-(\mathbf{p}'),u^+(\mathbf{p})]_- &= \delta(\mathbf{p}'-\mathbf{p}).\end{aligned} \tag{3}$$

In contrast to nonrelativistic quantum mechanics, a final theory of measurements is absent in quantum field theory. However, for single-particle states of the field we shall proceed by analogy to the nonrelativistic case. A measurement making it possible to distinguish two orthogonal states with certainty is given by a decomposition of unity in the subspace of single-particle states

$$\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_\perp=\mathcal{I}, \quad \mathcal{P}_\perp=\mathcal{I}-\mathcal{P}_1-\mathcal{P}_2, \quad \mathcal{P}_i\mathcal{P}_j=\delta_{ij}\mathcal{P}_i, \tag{4}$$

$$\mathcal{I}=\int u^+(\mathbf{p})|0\rangle\langle 0|u^-(\mathbf{p})\frac{d\mathbf{p}}{2p_0},$$

$$\mathcal{P}_{1,2}=\left(\int \psi_{1,2}(\mathbf{p}')u^+(\mathbf{p}')|0\rangle\frac{d\mathbf{p}'}{2p_0'}\right)\left(\int \langle 0|u^-(\mathbf{p})\psi_{1,2}^*(\mathbf{p})\frac{d\mathbf{p}}{2p_0}\right). \tag{5}$$

The probability of obtaining a result with the input state $|\psi_1\rangle$ is

$$\mathrm{Pr}_1(\psi_1)=\langle\psi_1|\mathcal{P}_1|\psi_1\rangle\equiv 1, \quad \mathrm{Pr}_{2,\perp}(\psi_1)=\langle\psi_1|\mathcal{P}_{2,\perp}|\psi_1\rangle\equiv 0, \tag{6}$$

and similarly for the state $|\psi_2\rangle$. The measurements (4) and (5) are nonlocal in the sense that access to the entire region of space where a field is present is required. It is intuitively obvious that if we are dealing with the electromagnetic field in a certain region of space (this means that in the state presented to us at a certain moment in time the quantum-mechanical averages of the observables of the field are different from zero in this region), to determine the state of the field the measuring apparatus must have access to this entire region. Even if at each point information obtained about the state of the field as a result of a local interaction between the field and the measuring apparatus arises instantaneously, a certain time is still required to transmit this information to an observer located at some fixed point in space. It is obvious that irrespective of the observer's location this time cannot be less than $L/2c$ in order of magnitude, where c is the speed of light and L is the diameter of the region where the field is nonzero.

For the protocol it is only important that the orthogonal states in the relativistic case can be distinguished with certainty only in a finite time that depends on the structure of the states themselves. In other words, orthogonal states are effectively indistinguishable (indistinguishable with certainty) during a certain finite time and become distinguishable with certainty after this time has elapsed.

To obtain information about the state of a field a measurement must encompass the region of space where this field is concentrated. For this reason, if initially the field is

prepared in a certain controllable (inaccessible for one of the exchange participants) region of space–time and then propagates into a region accessible for measurement, then the state becomes completely accessible in a finite time. The propagation amplitude of the field satisfies the causality principle

$$\langle \psi_{1,2}(\hat{x}_1)|\psi_{1,2}(\hat{x}_2)\rangle = -i\psi^*_{1,2}\left(-i\frac{\partial}{\partial \mathbf{x}_1}\right)\psi_{1,2}\left(i\frac{\partial}{\partial \mathbf{x}_2}\right)D_0^-(\hat{x}_1-\hat{x}_2), \tag{7}$$

where $D_0^-(\hat{x})$ is the negative-frequency function

$$D_0^-(\hat{x}) = \frac{i}{(2\pi)^{3/2}}\int d\hat{k}\,\delta(\hat{k}^2)\,\theta(-k^0)\exp(i\hat{k}\hat{x}) = \frac{1}{4\pi}\varepsilon(x^0)\,\delta(\lambda), \tag{8}$$

$$\varepsilon(x^0) = \theta(x^0)-\theta(-x^0), \quad \lambda^2 = (x^0)^2-\mathbf{x}^2; \tag{9}$$

$|\psi_{1,2}(\hat{x})\rangle$ is a state in the $\hat{x}$ representation:

$$|\psi_{1,2}(\hat{x})\rangle = \int \psi_{1,2}(\mathbf{p})e^{i\hat{p}\hat{x}}u^+(\mathbf{p})\frac{d\mathbf{p}}{\sqrt{2p_0}}|0\rangle. \tag{10}$$

We note that for $\hat{x}_1=\hat{x}_2$

$$\begin{aligned}\mathrm{Pr}_{1,2}(\psi_{1,2}) &= \langle \psi_{1,2}|\mathcal{P}_{1,2}|\psi_{1,2}\rangle \\ &= |\langle \psi_{1,2}(\hat{x}_1)|\psi_{1,2}(\hat{x}_2)\rangle|_{\hat{x}_1=\hat{x}_2}|^2 = -|\psi^*_{1,2}\left(-i\frac{\partial}{\partial \mathbf{x}_1}\right)\psi_{1,2}\left(i\frac{\partial}{\partial \mathbf{x}_2}\right)|^2 D_0^- \\ &\quad \times(\hat{x}_1-\hat{x}_2)D_0^+(\hat{x}_2-\hat{x}_1)|_{\hat{x}_1=\hat{x}_2}.\end{aligned} \tag{11}$$

Despite the product of two singular generalized functions ($D_0^-(\hat{x}) = -D_0^+(-\hat{x})$) at $\hat{x}^2=0$ in Eq. (11), such a product determines a generalized function, since there always exists a convolution of two generalized functions with a carrier in the forward part of the light cone.[19]

We now proceed to the description of the protocol. The form of the states $|\psi_1\rangle$ (corresponds to 0) and $|\psi_2\rangle$ (corresponds to 1) is stipulated in advance. The protocol starts at $t=0$. At $t=0$ each participant starts to prepare N states. The number N is also stipulated in advance. The worse case is where on the one hand each participant completely controls only the immediate neighborhood of his own laboratory while on the other he can arrange his apparatus as close as desired to the laboratory of the other participant, if he wishes to deceive the latter. This means that the protocol must be stable in a situation where one user can instantaneously transmit information to the other, i.e., when the length of the communication line between them is effectively zero. Actually, it is sufficient to require that the effective size of the region of localization of a state is much greater than the length of the communication line. Formally, the situation where the length of the communication line is zero means that the users cannot control the space outside their laboratories near $x_{A,B}$. At $t=0$ each participant switches on the source of the states, which immediately begin to propagate in the communication channel and become accessible for measurement. A finite time T is required to distinguish the states

with certainty (or with probability as close as desired to 1). A measurement to distinguish states with certainty is described by a decomposition of unity (4) and (5), which is written out unambiguously.

After a time $T/2$ has elapsed, first A reports to B which states he has transmitted, and A reports this only for $N/2$ of his own states. Then, after obtaining the information from A, B reports for the other half $N/2$ of the states what he actually transmitted. Only after receiving the classical information from B does A communicate to B the remaining second half $N/2$ of the states. After receiving from A, B reveals his own remaining half of the states.

Ultimately, the participants can check the correspondence between the classical information which they receive from one another and their quantum-mechanical measurements. If there is any irregularity, for example, A reports that an i state was present, for example, $|\psi_1\rangle$ (0), while B detects $|\psi_2\rangle$ (1), the protocol is aborted. The presence of classical information from one another and the ability to distinguish orthogonal quantum states with certainty do not permit switching even one bit, which will be important for calculating the resulting parity bit. Further, if as a result of exchange it is accepted that no cheating has taken place (the quantum measurements correspond to the classical information), a parity bit is calculated $c=c_1\oplus c_2\oplus \ldots c_N$ ($c_i=a_i\oplus b_i$, where a_i and b_i are bits from A and B, respectively), which is the lot.

We shall now discuss possible cheating strategies, for example, for participant A. To begin with, the participants must exchange classical information so as to rule out the possibility of the states which A receives being resent immediately back to B without first being identified. For photons, it is sufficient for A to place a mirror immediately at the point where B states enter the communication channel. Further, let it be stipulated in advance, for example, that A wins if 0 occurs. Then, if there was no requirement to report classical information, A receives states from B and, without attempting to identify them, immedidately sends them back to B. It is obvious that the parity bit in this case will always be zero (A wins), since $a_i\equiv b_i$, $c_i=a_i\oplus b_i=b_i\oplus b_i\equiv 0$, $c=c_1\oplus c_2\oplus \ldots c_N\equiv 0$. However, if A must report along a classical channel what he has actually transmitted to B, then such cheating is ruled out.

Alternate revelation of half of the states through a classical channel is necessary in order to rule out the following cheating strategy (which, incidentally, is easy to overlook). Since B controls only the space of his own laboratory near the point x_B, A can position himself near x_B, and after resending the quantum states back to B, at the stage of exchanging classical information A can transmit back to B virtually instantaneously the information which he has received from B along the classical channel. If B reports that he has transmitted immediately all N states (which A has transmitted back), then A can instantaneously return the classical information to B. Such a strategy does not work with alternate revelation of half the states.

For the protocol it is important that the states are quantum states. If the states were classical states, then, for example, A could always ''select,'' successively at each moment in time during the protocol, from each state obtained from B as small as desired a part for measurements (this is not forbidden in classical physics) and simultaneously resend the states back to B without distorting them. The arbitrarily small ''selected'' part of the states could be used to identify the B states during the protocol and then to report

them along a classical channel. In this case A always wins. For quantum states, such a strategy is impossible, since it is impossible to perform measurements at successive moments in time without introducing distortions.

Since a finite time is required to distinguish orthogonal states with certainty, during the time $0 \leqslant t \leqslant T$ the states are indistinguishable with certainty and are effectively nonorthogonal. The probability of correct detection of a state is an increasing function of time $P(t)$: ($P(0)=1/2$ (simple guess) and $P(T)=1$). The specific form $P(t)$ depends on the concrete states and is immaterial for us. The participant A cannot delay the transmission of his own states. Otherwise they will not be detected in the time T (for large N there will be events outside $0 \leqslant t \leqslant T$). One possible cheating strategy could be for participant A to adjust the states that have already been transmitted depending on the results of measurements performed on a state received from B. In this case A must already obtain the result of measurements on a state from B by a certain time t [the probability of this $p(t)$], and B must not have had enough time to detect a state from A (the probability of this event is $1-p(t)$). The probability of successful cheating is $P_{ch}=p(t)[1-p(t)]$. The maximum of P_{ch} occurs for $p(t_c)=1/2$ and is 1/4, which is less than the probability of a simple guess, which is 1/2. In principle, A could perform a collective measurement immediately on N states transmitted from B,[20] but because the states are effectively nonorthogonal for time T the probability of identifying the parity bit correctly in this case once again reaches 1 only for $t=T$.

In closing, I shall make one remark. The possibility of identifying a state with probability 1 during a finite time interval T depends on whether or not there exist states, for the type of particles employed, with a finite carrier in space. For photons, today only states with exponential localization of energy are known.[21] For a protocol this is not too significant, since the time interval can be chosen to be long enough to guarantee that the probability of distinguishing two orthogonal states is exponentially close to 1.

This work was supported by the Russian Fund for Fundamental Research (Project 99–02–18127) and Project 02.04.5.2.40.T.50 of the program ''Promising Technologies and Devices for Micro- and Nanoelectronics.''

[1] M. Blum, in *Proceedings 24th IEEE Comp. Conf.*, 1982, p. 133-137; also in SIGACT News **15**, 23 (1983).
[2] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, Santa Fe, NM, USA, edited by S. Goldwasser (IEEE Comput. Soc. Press, Los Alamitos) 1994, p. 124.
[3] A. Yu. Kitaev, Usp. Mat. Nauk **52**, 54 (1997).
[4] S. Wiesner, SIGACT News **15**, 78 (1983).
[5] C. H. Bennett and G. Brassard, in *Proceedings IEEE Int. Conf. on Computers, Sytems, and Signal Processing* (IEEE, New York, 1984) p. 175.
[6] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[7] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois, in *Proceedings of the 34th Annual IEEE Symposium on the Foundation of Computer Science* (IEEE Cmp. Soc., Los Alamitos, California, 1993) p. 362.
[8] G. Brassard and C. Crépeau, *Advances in Cryptology: Proceedings of Crypto'90, Lecture Notes in Computer Science,* Vol. 537, p. 49 (Springer-Verlag, Berlin, 1991).
[9] M. Ardehali, quant-ph/9603015.
[10] D. Mayers, L. Salvail, and Y. Chiba-Kohno, quant-ph/9904078.
[11] L. Goldenberg, L. Vaidman, and S. Wiesner, quant-ph/9808001.
[12] H. F. Chau, quant-ph/9901024.
[13] H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
[14] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).
[15] A. Kent, quant-ph/9810067; quant-ph/9810068; Phys. Rev. Lett. **83**, 1447 (1999); quant-ph/9906103.
[16] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).

[17] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[18] L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239 (1995).
[19] N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, 3rd edition (Wiley, New York, 1980) [Russian original, Nauka, Moscow, 1973].
[20] Tal Mor, quant-ph/9906073.
[21] I. Bialynicki-Birula, Phys. Rev. Lett. , **80**, 5247 (1998).

Translated by M. E. Alferieff

ERRATA

Erratum: Texture and magnetic anisotropy of carbon nanotubes in cathode deposits obtained by the electric-arc method [JETP Lett. 70, No. 7, 476–480 (10 October 1999)]

A. S. Kotosonov
Scientific-Research Institute of Graphite Materials, 111524 Moscow, Russia
Pis'ma Zh. Eksp. Teor. Fiz. **70**, No. 10, 690 (25 November 1999)

[S0021-3640(99)01122-6]

PACS numbers: 81.05.Ys, 61.46.+w

In the paper by A. S. Kotosonov, published in **70**, No. 7, in the second sentence on page 480 the citations Refs. 9 and 10 should be Refs. 11 and 12.

Translated by M. E. Alferieff

