X-ray Emission Lines in the Early Afterglows of Gamma-Ray Bursts

D. I. Kosenko[1*], S. I. Blinnikov[1,2], K. A. Postnov[1**], P. Lundqvist[3], and E. I. Sorokina[1]

[1]*Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119992 Russia*

[2]*Institute for Theoretical and Experimental Physics,*
ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia

[3]*Stockholm Observatory, Albanova, Stockholm, Sweden*

Received November 18, 2002

Abstract—We computed the thermal time-dependent X-ray spectrum of a hot plasma in the vicinity of a gamma-ray burst (GRB). An allowance for time-dependent processes in a hot rarefied plasma is shown to strongly affect the observed spectrum. These computations can give an alternative explanation for the observed X-ray emission lines in the early afterglows of GRBs (e.g., GRB 011211). Our technique allows the GRB collimation angle and the environment clumpiness parameters to be independently constrained.
© *2003 MAIK "Nauka/Interperiodica".*

Key words: *gamma-ray bursts, X-ray lines.*

INTRODUCTION

The discovery of X-ray afterglows from cosmic gamma-ray bursts (GRBs) aboard the Beppo-SAX satellite (Costa *et al.* 1997) has become a turning point in current studies of this interesting astrophysical phenomenon (see the reviews of Postnov (1998), van Paradijs *et al.* (2000), Mészáros (2002) and references therein). In the standard GRB model (Rees and Mészáros 1992, 1994; Piran 2000), the central source (whose nature has yet to be established) releases an energy of $\sim 10^{51}$ erg on a time scale of ~ 100 s. This energy can be thermal (a photon−lepton fireball; see Piran 2000) and electromagnetic (Usov 1994; Spruit 1999; Blandford 2002) and probably propagates in a narrow cone (jet). The observed gamma-ray features from GRBs (a nonthermal spectrum, rapid temporal variability) actually require an ultrarelativistic motion of the emitting plasma (see the reviews of Piran (2000) and Blinnikov (2000)) with characteristic Lorentz factors $\Gamma \sim 100-300$. The observed gamma-ray photons are generated by a nonthermal mechanism at the fronts of relativistic shock waves (although the apparent "nonthermal" spectrum can also be explained in terms of the model of optically thick shells moving at relativistic velocities; see Blinnikov *et al.* 1999). The X-ray and optical afterglows of GRBs in these models are also associated with the nonthermal (synchrotron) radiation of relativistic particles at the front of an external

shock wave that decelerates in the circumstellar or interstellar medium (Mészáros and Rees 1997). Apart from this standard model of GRB afterglows, we (Blinnikov and Postnov 1998) offered an alternative explanation for the first observed optical afterglow of GRB 970228, in which the GRB heated the envelope of a star (e.g., in a close binary) and the afterglow mechanism was purely thermal.

With the launch of the Beppo-SAX, Chandra, and XMM-Newton X-ray observatories in the late 1990s, a qualitatively new possibility of studying the soft X-ray spectra of early GRB afterglows with a high spectral resolution emerged. The first reliable data were obtained for GRB 991216. Emission lines were detected in the spectrum of the X-ray afterglow observed by Chandra ~ 1.5 days after the GRB. These lines were interpreted as the fluorescent $K\alpha$ iron line and the recombination jump of a hydrogenic iron ion in emission (Piro *et al.* 2000). To explain the observed line equivalent width required assuming a high metallicity of the emitting plasma (the iron ion abundance is more than a factor of 60 higher than its solar value, 1.8×10^{-3}, of the total mass), which was interpreted as evidence for the shell ejection of the supernova that preceded the GRB (the "supranova" model; Vietri and Stella 1998). Iron spectral lines were observed with lower confidence in the afterglows of several other GRBs (GRB 970508, GRB 970828, and GRB 000214). However, the results obtained for GRB 011211 by the XMM-Newton observatory (Reeves *et al.* 2002a, 2002b) proved to be of the greatest interest. The spectrum of the early X-ray

afterglow observed 11 h after the GRB exhibited a number of emission lines and evolved with time (line decay on a time scale of 10^4 s). The emission lines were interpreted as the $K\alpha$ lines of several ions with α-particle nuclei (Mg XI, Si XIV, S XVI, Ar XVIII, Ca XX) against the background of a thermal ff continuum with a temperature $T \sim 4.5$ keV and their time evolution was associated with the geometric factor in an expanding shell. This interpretation requires a line blueshift relative to the source's frame of reference. The authors believe that this blueshift could be caused by the motion of the emitting region at a velocity of ~ 0.1 s. Remarkably, no iron lines were found, and the authors attempt to explain this fact by peculiarities of the shell ejection of the supernova that may have preceded the GRB.

Further, as the XMM-Newton observers argue (Watson *et al.* 2002), in three of the four X-ray GRB afterglows observed with the EPIC (PN, MOS) CCD arrays of the XMM-Newton observatory, the model of thermal radiation from an optically thin plasma with collisional ionization (the coronal approximation) gives a better fit to the spectra than does a power-law continuum with absorption in the soft energy range, which is expected in models with the reflection of nonthermal X-ray radiation from a cloud of cold plasma and in photoionization models.

For the sake of justice, it should be noted that the procedure of fitting the XMM-Newton X-ray spectra was criticized because of the possible instrumental effects (Borozdin and Trudolyubov 2002) and the observed emission lines in GRB 011211 can be interpreted in a different way (Rutledge and Sako 2002). However, the very fact of the presence and time evolution of emission lines in the spectra of early X-ray GRB afterglows seems reliable enough. The physical explanation of the observed spectral features is clearly required, and this subject was discussed by several authors (Lazzati *et al.* 1999; Kumar and Narayan 2002). It was shown that none of the suggested models gave a completely satisfactory explanation of the observed data and a number of assumptions (physical or geometrical) are required. The general conclusion drawn from the analysis and interpretation of the observed X-ray spectral features is that the metallicity of the emitting plasma is much higher than its solar value.

In contrast to previous studies, here, we study the effects of *time-dependent* (collisional) ionization in a hot rarefied plasma heated by a GRB. This plasma can be associated with the preceding evolutionary phase of the GRB progenitor and can happen to be on the line of sight by chance. The time evolution in the continuum and lines observed in GRB 011211 can be direct evidence that the ionization in the GRB-heated plasma is time-dependent. We show that the lines of the observed species can be produced in astrophysically admissible conditions of the environment. However, their blueshift relative to the source is difficult to explain. In the thermal model, we should expect photorecombination jumps rather than individual lines from a supernova shell ejection, i.e., at an appreciable blueshift. This is because, at a normal composition of the hydrogen supernova shell, no lines are produced, a tenfold enhancement of metals is not natural, and the current views of the chemical composition of supernovae shells favor a purely metallic ejection.

THE PHOTOIONIZATION AND THERMAL MODELS

When modeling X-ray emission from a hot plasma, it is convenient to consider two extreme cases: the photoionization and thermal approximations (see, e.g., a recent review of Tatischeff (2002) and references therein). The photoionization (nebular) model describes the physical state of and emission from a plasma ionized by X-ray and gamma-ray photons (see Kallman and McCray 1982). In this case, matter reprocesses the incident hard photons into a softer line emission and/or into a continuum in neighboring spectral ranges. The basic tenet of the nebular model is that the temperature and ionization states of the medium are determined only by the plasma interaction with a radiation field; the other heating mechanisms are disregarded.

The thermal (coronal) model is used to describe an optically thin rarefied plasma whose temperature is high enough to ionize atoms by electron collisions. In this case, the radiation incident on the plasma plays no significant role, the ionization is produced by collisions, the ion and electron velocity distributions are Maxwellian, and the forming thermal (free−free, line, recombination, etc.) X-ray emission leaves the medium almost freely.

When considering time-variable processes (e.g., during impulsive irradiation/heating of gas) and when interpreting (fitting) the corresponding spectra, apart from the basic physical parameters of the medium (density, metallicity, optical depth, etc.), the relative heating and cooling rates and the relaxation time scales in the plasma should be taken into account to properly choose a radiation model.

TIME-DEPENDENT EMISSION LINES FROM A THERMAL PLASMA

The X-ray spectral diagnostics of an astrophysical plasma is widely used to interpret the observations. If the plasma is dense enough ($\equiv$ the thermal relaxation time is much shorter than the time

scale of the problem), then it is assumed to be time-independent. However, when interpreting actual astrophysical sources, allowing for time-dependent ionization processes can play a crucial role in determining the physical parameters of the medium.

Consider a thermal plasma with a particle density n and an effective ion charge Z in the coronal approximation. It is assumed to be impulsively heated to high temperatures, $T \sim 10^8$ K (we are not concerned with the details of the heating mechanism, but it is implied that the plasma was heated by hard photons directly from the GRB). The plasma emissivity due to free–free radiation is (Zeldovich and Raizer 1966; Rybicki and Lightman 1979)

$$\varepsilon_{\mathrm{ff}}(T) \approx 1.4 \times 10^{-27} [\mathrm{erg\ s^{-1}\ cm^{-3}}] T^{1/2} n_e n_i Z^2 \bar{g}_B, \tag{1}$$

where n_e and n_i are the electron and ion densities, respectively; Z is the ion charge; and $\bar{g}_B \sim 1.1–1.5$ is the Gaunt factor. For a plasma with a normal cosmic abundance, $\varepsilon_{\mathrm{ff}} \approx 1.7\varepsilon_{\mathrm{ff}}(\mathrm{H})$, where $\varepsilon_{\mathrm{ff}}(\mathrm{H})$ is the emissivity for a purely hydrogen, completely ionized plasma ($n_i = n_e$, $Z = 1$). The cooling time for a completely ionized plasma composed of electrons and ions with charge Z (due to free–free radiation alone; allowing for the recombination emission decreases this time severalfold, see Fig. 2) is

$$t_c \approx 2 \times 10^{15} [\mathrm{s}] \left(\frac{T}{10^8}\right)^{1/2} n_i^{-1} Z^{-2} \tag{2}$$

$$\approx 2 \times 10^{15} [\mathrm{s}] \left(\frac{T}{10^8}\right)^{1/2} n_e^{-1} Z^{-1}.$$

Thus, for a plasma with a given chemical composition, the use of formula (2) must be stipulated in each specific case and the numerical factor (correct to within $\sim 20\%$ for a purely hydrogen plasma) is only an upper limit for the cooling time.

In the model of Reeves et al. (2002a), the electron density is $n_e \sim 10^{15}$ cm^{-3}; i.e., the plasma cools down in less than a second. If the plasma is not so dense, for example, $n_e \sim 10^{11}–10^{12}$ cm^{-3}, then its cooling time is much longer than its heating time, $t_{\mathrm{h}} \sim 90$ s (actually the GRB duration in a comoving frame), but is still comparable to or even shorter than the observing time of the X-ray radiation; i.e., the time dependency should be taken into account. There are two manifestations in the time dependency. First, over the integration time of X-ray photons in the detector (~ 5000 s for GRB 012111), the true spectrum can radically change because of a temperature drop even if the density of the medium is high and the parameter nt formally allows the model of time-independent ionization to be used. Second, the parameter nt is occasionally small and relaxation effects

for time-dependent plasma ionization can give rise to ions that were absent in the case of time-independent ionization. We emphasize that, since the cooling time strongly depends on the ion charge, $\propto Z^{-2}(!)$ [see formula (2)], the actual cooling time depends on the ionization stage and the ion density; for a low-density plasma, time-dependent ionization plays a significant role.

As an example, Fig. 1 plots the relative densities of some species against temperature for the time-independent case and against time for the model of time-dependent ionization. In the latter case, an initially cold plasma (10^5 K) with density $n = 10^{12}$ cm^{-3} instantly heats up to a temperature of 10^9 K. Thus, the second manifestation of time dependency is illustrated here.

We see from the figures that, during thermal relaxation, the ionization state of the medium at some time can be close to a steady ionization state at a different fixed temperature.

APPLICATION TO THE OBSERVATIONS OF THE X-RAY AFTERGLOW FROM GRB 011211

The GRB Afterglow

Below, we briefly repeat the main results of the observations of the X-ray afterglow from GRB 011211. The XMM-Newton X-ray observations were started 11 h (4×10^4 s) after the GRB. The redshift determined from the observations of the GRB host galaxy is $z = 2.14$. The emission lines in the X-ray spectrum were observed for 10^4 s and then became indistinguishable against the continuum background (actually, their equivalent widths became less than 100 eV). The spectrum observed in the first kiloseconds was fitted by a thermal model with a plasma temperature of 4.5 keV. Over the entire observing time (27 ks), the mean temperature was 2.1 keV. No iron lines were detected in the spectrum. We deduced the volume emission measure of the source $\mathrm{ME} = n_e^2 V = 10^{69}$ cm^{-3} from the X-ray luminosity averaged over the observing time (7×10^{45} erg s^{-1}); for a plasma with a normal cosmic abundance, it requires extreme parameters (a high density and/or a large volume) of the emitting medium. Actually, this is a model-dependent estimate in our case; we assumed the thermal continuum to be associated only with the free–free radiation of a purely hydrogen plasma. Clearly (see formulas (1) and (2)), for a plasma composed of heavy elements, the emission measure can be much smaller ($n_e^2 V \propto Z^{-1}$ for a completely ionized plasma) and the estimate of 10^{69} cm^{-3} is an upper limit.

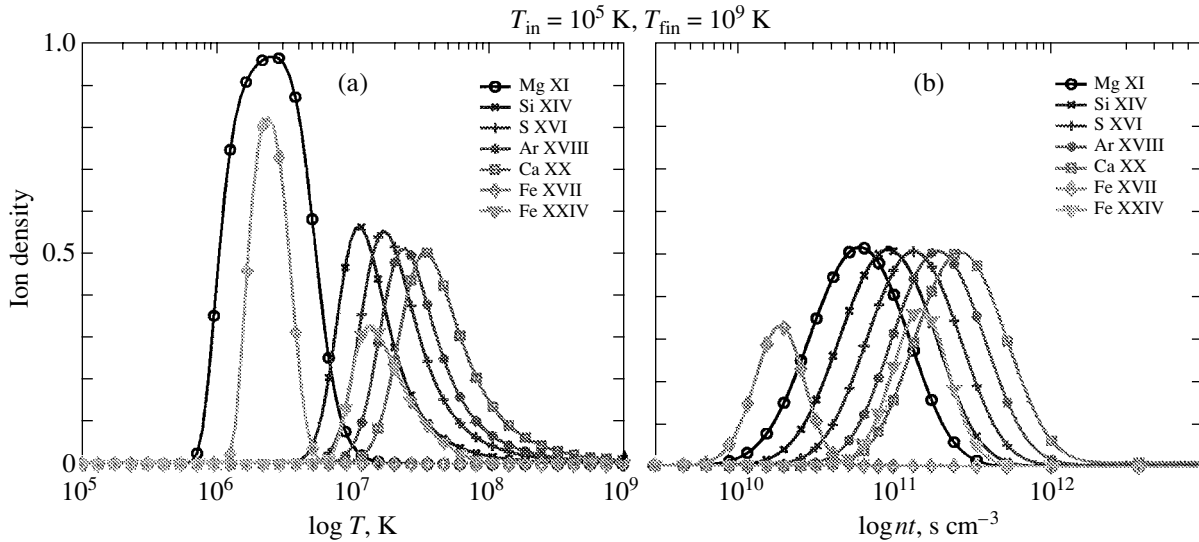$T_{in} = 10^5$ K, $T_{fin} = 10^9$ K

Fig. 1. Relative ion density versus temperature for the time-independent case (a) and versus time for the time-dependent case (b).

The Model of Reeves et al.

The observed spectrum was interpreted in terms of the model of an expanding dense shell (around the GRB) with radius $R = 10^{15}$ cm, thickness $h = 10^9$ cm, and density $n_e = 10^{15}$ cm^{-3}. The mean redshifts of the identified X-ray lines and the GRB are $z = 1.88$ and 2.14, respectively; i.e., the lines are blueshifted relative to the GRB. The shell velocity was found from these data to be $v \sim 0.1$ s. Since such a dense plasma cools down in 1 s (2), the observed line decay time was explained by a delay of the signal from different parts of the instantly illuminated shell (this geometric effect was invoked to account for the evolution time scales of X-ray afterglows, e.g., by Lazzati *et al.* (1999)).

Remarks on the Model of Reeves *et al.* (2002a). This model contains contradictions that call into question the possibility of its application.

A short plasma cooling time. If the gas cooling time is 1 s, then the bulk of the observed X-ray radiation must be produced when matter is irradiated by the GRB—the GRB duration was ~90 s in a comoving frame. The time difference between the gamma-ray and X-ray signals was then attributable only to the time delay of the radiation from the region being heated that lay slightly away from the line of sight to the GRB. However, this implies that we see the X-ray photons emitted by the plasma at the time when the plasma is being irradiated by an intense hard gamma-ray flux. Hence, for the corresponding spectrum to be explained, we must consider the photoionization approximation rather than the thermal one. However, since the photoionization model gives a spectrum that

does not reproduce the observations, it was rejected by observers (see Reeves *et al.* 2002a).

To remain within the framework of the thermal model, which, as argued by Reeves *et al.*, reproduces the observed X-ray spectrum, it should be assumed that the density of the emitting matter is much lower than the previously adopted value of 10^{15} cm^{-3}. If we assume that $n_e \sim 10^{11}-10^{12}$ cm^{-3}, then the cooling time $t_c \propto 1/n_e$ from expression (2) proves to be longer than the GRB duration, $t_c > t_h$. Only in this case can an attempt be made to describe the X-ray radiation in the thermal model.

A large emission measure. The emission measure in the model of Reeves *et al.* is determined from the mean X-ray continuum luminosity and, for the assumed density of $n_e = 10^{15}$ cm^{-3}, refers only to the layer volume $V = R^2 h = 10^{15} \times 10^{15} \times 10^9 = 10^{39}$ cm^3 that we see for 1 s. Since the X-ray radiation is recorded for 10^4 s, there must be 10 000 such volumes! This is in conflict with the assumption that one shell with radius $R = 10^{15}$ cm and thickness $h = 10^9$ cm emits over the entire observing time. For the model to be reconciled somehow, we must assume the presence of clumps with sizes $l = 10^{12} \gg h$ that are opaque at this density.

The apparent absence of iron lines. Reeves *et al.* interpreted this fact as follows: ^{56}Fe has not yet been formed out of the radioactive ^{56}Ni ejected by the supernova (SN). Recent calculations of nucleosynthesis during the explosion of a collapsing SN (see Rauscher *et al.* 2001) actually yield a deficit of all Fe isotopes compared to ^{56}Ni. However, these calculations were performed under special
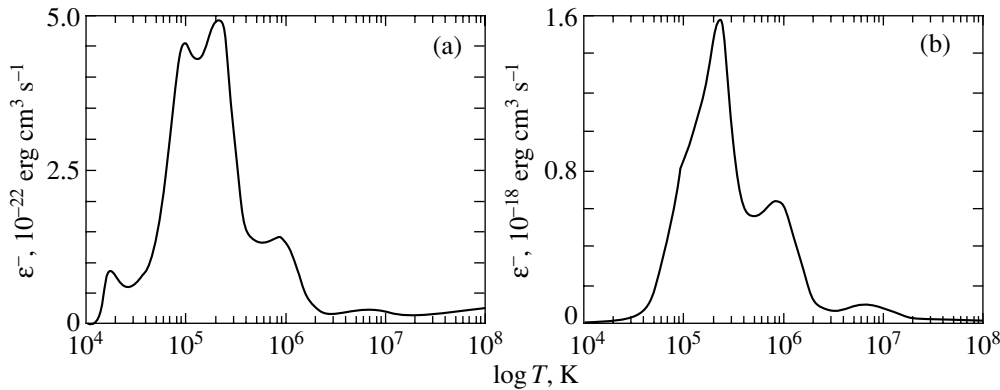
Fig. 2. The plasma cooling function ε^- in erg cm^3 s^{-1}: (a) solar chemical composition, then ME $= n^2 V = L_X/Q^- \simeq 3.7 \times 10^{68}$ cm^3; (b) a metal-enriched medium (H and He are absent) with ME $= n^2 V = L_X/Q^- \simeq 1.4 \times 10^{65}$ cm^3.

assumptions about the explosion mechanism for an ordinary type-II SN. Since this mechanism is not yet known, Rauscher *et al.* (2001) made special assumptions about the motion of an artificial piston that simulated a spherically symmetric explosion; the piston location was chosen from the mass cut so as to obtain the required amount of ^{56}Ni. However, if the shell in the model under discussion is actually the SN ejection, then this supernova is unlikely to have been ordinary. In particular, the explosion could have been highly asymmetric, when the initial iron, nickel, and cobalt abundances could be comparable. Asymmetric SN-explosion/GRB mechanisms can yield different results (Pruet *et al.* 2002): because of the large neutron excess, nucleosynthesis in the GRB accretion disk produces mainly iron, while ^{56}Ni is in deficit! Umeda *et al.* (2002) provided a table for the dependence of M(Fe) on the mass cut in hypernovae; some models give only ^{56}Ni, while other models give up to 40% iron. Thus, the amount of iron in SN ejecta is strongly model-dependent, and the fact that no iron lines are observed cannot serve as a solid argument for a SN associated with the GRB.

A high density and a small shell thickness. Difficulties also arise in the question of physical parameters for the SN ejection, which was collected in the model of Reeves *et al.* (2002a) into a thin dense layer that flies at a high velocity. Similar dense shells are occasionally reproduced in modeling the radiation hydrodynamics of supernovae (Falk and Arnett 1977; Chevalier and Klein 1979; Blinnikov and Bartunov 1993), but their properties are completely different. The total shell energy in the model of Reeves *et al.* (at a velocity reaching such high values as 2.6×10^4 km s^{-1}) approaches the characteristic SN explosion energy, while the hydrodynamically modeled dense shells are either of low mass or have much lower velocities—of the order of several thousand km s^{-1}

(Chugai *et al.* 2003)—and they carry away only a small fraction of the supernova energy.

In our opinion, the above contradictions make the model of thermal X-ray radiation from the early afterglow of GRB 011211 by Reeves *et al.* (2002a) completely inadequate.

A Model with Time-Dependent Ionization

Assume that decaying X-ray lines were actually observed in the afterglow spectrum of GRB 011211. We offer a different interpretation of the origin of the lines and their time evolution. Such an X-ray spectrum can be obtained in terms of a time-dependent thermal model under different assumptions about the density and geometry of the emitting matter. We identify the observed line decay time with the *thermal relaxation time* of the GRB-heated hot plasma.

Geometry of the emitting region. Assume that the GRB occurred inside (or near) a star-forming region with a complex clumpy distribution of matter. This assumption is quite natural, because the entire set of currently available observational data on GRBs points to their association with star-forming regions (van Paradijs *et al.* 2000). For the cooling time (2) to be about 10^4 s, the gas density must be about $n_e \sim 10^{11}$–10^{12} cm^{-3}. To obtain the observed emission measure ($n_e^2 V \lesssim 10^{69}$ cm^{-3}), the volume of the emitting region must be $V \lesssim 10^{47}$–10^{45} cm^3. At the same time, this volume must be optically thin for gamma-ray and X-ray radiation; i.e., $nl\sigma_T \sim 1$, where $\sigma_T \sim 10^{-24}$ cm^2 is the Thomson cross section for electrons. These two conditions are difficult to simultaneously satisfy in a homogeneous medium. However, for a more realistic, rarefied medium with dense condensations, these two conditions can be satisfied.

Indeed, consider $N_{cl} \sim 10^6$ gaseous clouds with a density of 10^{12} cm^{-3} and a scale size of $l_{cl} \sim$

10^{13} cm. The total volume of the dense component of the medium in such clouds corresponds to the required value of 10^{45} cm^3 and gives the required emission measure of 10^{69} cm^{-3} (recall that these are upper limits! Thus, a realistic cooling function yields ME $\sim 4 \times 10^{68}$ cm^{-3} for a plasma with a normal cosmic abundance and ME $\sim 10^{65}$ cm^{-3} for a plasma without H and He and with a solar ratio of the remaining elements; see Fig. 2). The total mass of the emitting clouds does not exceed $\sim 0.3 M_\odot$. Clouds of such sizes and densities are observed as maser condensations in the stellar winds of young and late-type stars (Bains *et al.* 2002, Lekht *et al.* 2001). The total volume occupied by such inhomogeneous matter is determined by the energy balance and the gamma-ray beaming cone angle.

The energy $\epsilon_0 = 10$ keV per nucleon, i.e., the the heating energy $E_{\mathrm{h}} = 10^{12}$ cm$^{-3} \times 10^{45}$ см$^3 \times$ 10 keV$= 10^{48}$ erg, is required to heat the matter to 10^8 K. The total isotropic GRB energy is $E_\gamma = 5 \times 10^{52}$ erg. The effective gamma-ray deposition cross section for heating the environment per nucleon can be assumed, with a high accuracy, to be $\sigma_\gamma \simeq 0.1\sigma_T$ (Ambwani and Sutherland 1988). Since the fluence at a distance d from the source is $F = E_\gamma/4\pi d^2$, the heating per nucleon is $\epsilon_0 = \sigma_\gamma F$, from which

$$d = \sqrt{\frac{\sigma_\gamma E_\gamma}{4\pi\epsilon_0}} \simeq 2 \times 10^{17} \quad \mathrm{cm};$$

i.e., the source energy is enough to heat up the surrounding matter up to distances $d \simeq 0.1$ pc. Note that F plays a more important role than E_γ does; F is measured by an instrument (at the distance from the source to the instrument rather than at d) and no assumptions should be made about radiation isotropy to estimate d.

The cone opening angle θ must be specified to calculate the total volume of the interstellar medium illuminated by gamma rays. Let one cloud illuminated by gamma rays be at a distance d from the central source exactly in the line of sight of the GRB and another cloud be at the same distance but be displaced from our line of sight by $R = d\sin\theta$; i.e., it lies at the edge of the gamma-ray cone. In this case,

$$d(1 - \cos\theta) = ct_l/(1+z) \simeq 10^{14} \quad \mathrm{cm},$$

where $t_l = 10^4$ s is the observed line decay time (if this decay is attributable to the delay effect rather than to the plasma cooling) and $z = 2.14$. If we take a large angle, $\theta = 20°$, as Reeves *et al.* did, then we find, as they did, that the cross section of the gamma-ray cone is $R \sim d = 10^{15}$ cm.

However, it would be more natural to assume that the collimation angle was smaller, because the isotropic energy of GRB 011211 was great. For small angles θ, we have $d(1 - \cos\theta) \simeq d\theta^2/2$; i.e., we obtain an estimate for the radius of the GRB cone cross section, $R \simeq \theta d = (2dct_l/(1+z))^{1/2} = (2 \times 10^{14}\mathrm{cm}\ d)^{1/2}$. In that case, for example, $R \simeq 10^{16}$ cm for $d = 2 \times 10^{17}$ cm and $\theta \simeq R/d = 0.05$, i.e., 3° (the total collimation angle is twice as large, $2\theta = 6°$).

It is important that the total volume of the GRB-illuminated medium is $V \sim R^2 d = 3 \times 10^{49}$ cm^{-3}. We note here that, although the gamma rays from the GRB traverse the distance d in several months, they reach the observer together with the X-rays produced (in the way!) by them. Therefore, an extended rarefied region with a size of the order of d filled with denser clouds rather than a narrow dense region 10^{13} cm in thickness can emit in the X-ray range. In order of magnitude, the filling factor can be $f \sim 10^{45}/10^{49} = 10^{-4}$. It can also be larger at smaller d, at lower densities in condensations or if condensation clumps begin to emerge only in a relatively thin layer rather than at all distances of the order of d. This simple estimate indicates that the physically acceptable structure of the clumpy interstellar medium can actually be an X-ray source. For example, if several hundred stars with intense stellar winds fall within a volume, then the number of such condensations can reach the required value of $10^5 - 10^6$ for the observed parameters of maser condensations (Banes *et al.* 2002). Such a high density of stars is not typical of our Galaxy, but the parameters of the interstellar medium in galaxies at $z \sim 2$ can be completely different.

Lazzati (2002) independently reached a similar conclusion that the medium surrounding a GRB must be clumpy when analyzing the conditions necessary for the realization of (time-independent) thermal models for early X-ray GRB afterglows. We emphasize once again that, in contrast to Lazzati, we do not require the presence of a shell around the GRBs; it will suffice that a region with the required plasma parameters be on the line of sight. This region may not be genetically associated with the GRB at all.

Modeling the spectrum. To model the X-ray spectra of an optically thin thermal plasma, including those for the afterglow of GRB 011211, we used a numerical algorithm developed by P. Lundqvist to compute the state of ionization and radiation of a time-independent plasma that takes into account the basic elementary processes of ionization by electron collisions, autoionization, photorecombination, dielectronic recombination, and ion charge exchange, and that computes the spectra of free−free, free−bound, line, and two-photon emission. This algorithm was adapted by Blinnikov and Sorokina to the time-dependent case ($\dot{n}_i = f(T(t), n(t), Z, ...)$ for

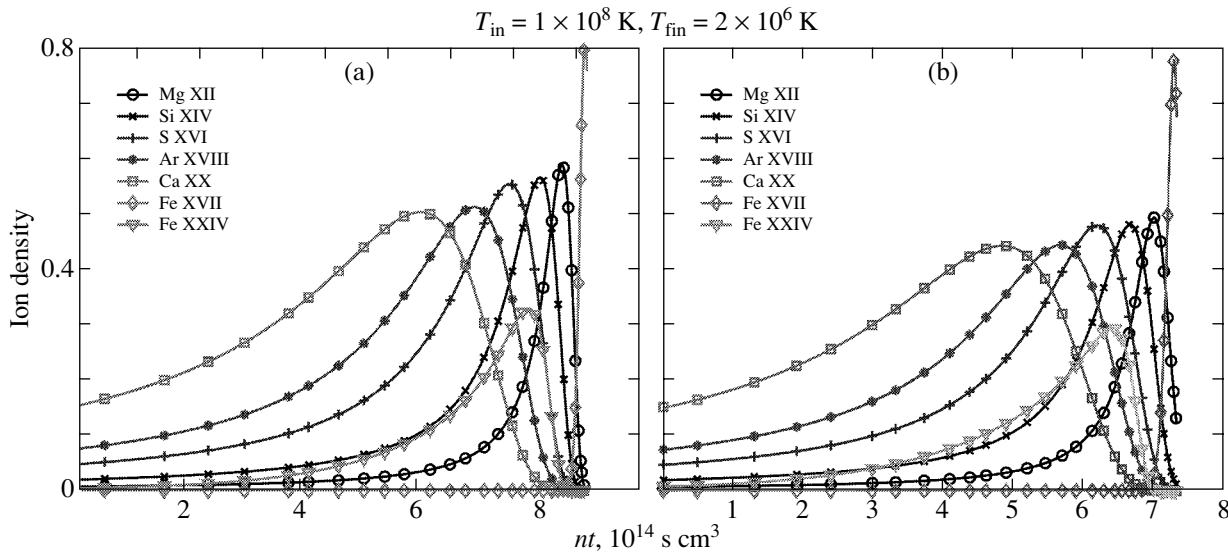$T_{in} = 1 \times 10^8$ K, $T_{fin} = 2 \times 10^6$ K

Fig. 3. Relative ion density versus nt for a linear cooling law $T(t)$ (3): (a) solar chemical composition, (b) a metal-enriched medium (without H and He).

the densities of ions of all types). The computational parameters are the current temperature, the particle density, and the chemical composition.

To calculate the time dependences of the spectra, we must find the law of temperature variations with time $T(t)$. The plasma cooling function ε^- (erg cm^3 s^{-1}) in the temperature range $2 \times 10^6-10^8$ K is almost independent of the temperature (see Fig. 2). In this case, the temperature behavior with time can be determined from the differential equation

$$\frac{dT}{dt} = -\frac{\varepsilon^-(T_0)n_e^2}{C_V \rho},$$

where $C_V = (3/2)\Re(1/A + X_e)$ is the specific heat at constant volume, $\Re$ is the universal gas constant, A is the mean ion mass in atomic units, X_e is the number of free electrons per nucleon, and $T_0 = 10^8$ K is the initial temperature.

In the above range, the temperature then linearly decreases with time:

$$T(t) = T_0 - \varepsilon^-(T_0)n_e^2 t/(C_V \rho). \qquad (3)$$

The time evolution of the ion densities (see Fig. 3) can be computed numerically and determines the behavior of the corresponding spectral lines (see Fig. 4).

The figures show a strong time evolution of the thermal spectra for the cooling optically thin plasma heated by the GRB to temperatures $\sim 10^8$ K. The emission line and continuum intensities are seen to evolve. For a normal heavy-element abundance, we immediately see iron ion lines of a moderate degree of ionization; hydrogenic and helium-like iron lines

appear as the plasma cools down (see Fig. 4). If we assume that a SN ejection with a heavy-metal over-abundance (and, accordingly, with an underabundance of hydrogen and helium) is observed, then the metal emission lines become more intense against the continuum background. At late cooling stages, photorecombination jumps of metals become prominent in the continuum (see Fig. 4).

DISCUSSION

We computed the X-ray spectra for an optically thin thermal plasma with time-dependent collisional ionization. Our analysis indicates that a moderate density of the emitting plasma ($n_e \sim 10^{11}-10^{12}$ cm^{-3}) and a clumpy circumstellar (interstellar) medium are required to interpret the line spectra of early X-ray afterglows (as, e.g., from GRB 011211) in terms of thermal models.

In our model of an early afterglow, the X-ray line emission does not necessarily originate in the matter genetically associated with the GRB source; it will suffice that a region with such dense condensations at distances from the source less than ~ 0.1 pc be found within the GRB radiation cone. For example, about $0.3 M_\odot$ of matter, with a volume of $\sim 10^{49}$ cm^3 contained in individual clouds with linear sizes of ~ 1 AU illuminated by gamma rays in a cone with an opening angle of $2\theta \sim 6°$, is required for the parameters of the observed X-ray radiation from GRB 011211, $L_X(0.6-30)$ keV $\approx 7 \times 10^{45}$ erg s^{-1} (in the source's frame). Such angles are obtained in the standard interpretation of the decay of GRB afterglows (Frail *et al.* 2001). The parameters of individual clouds
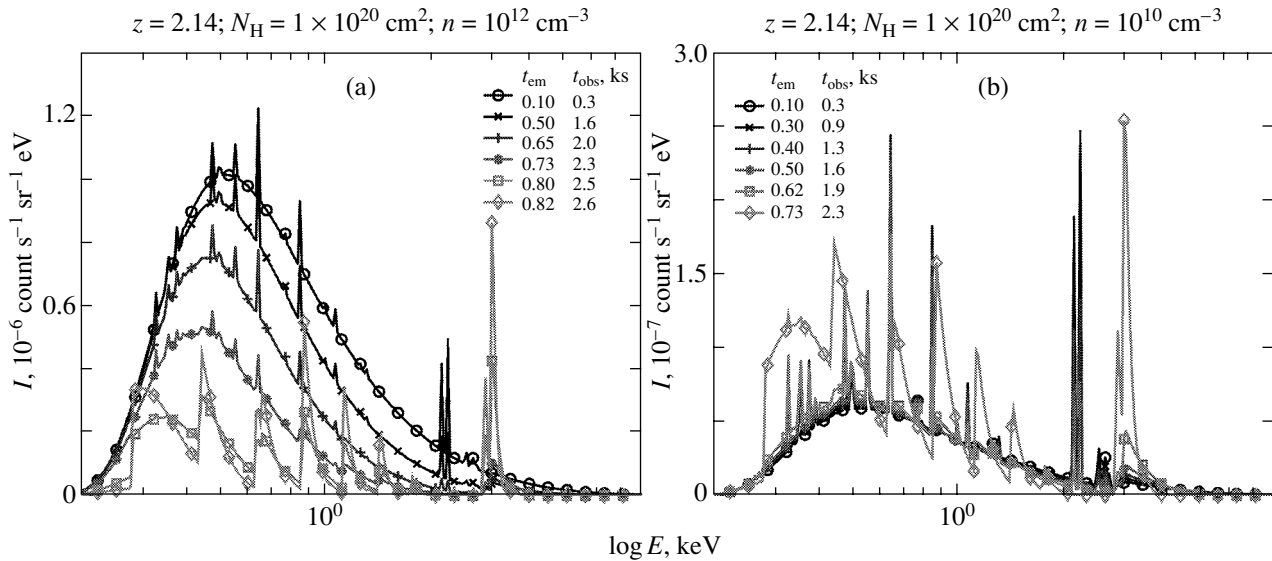
Fig. 4. Time evolution of the spectrum (with an allowance made for the redshift $z = 2.14$) at various times: (a) solar chemical composition and (b) a metal-enriched medium (without H and He).

are also astrophysically acceptable and correspond to dense maser condensations in star-forming regions (Bains *et al.* 2002; Lekht *et al.* 2001).

In our model, the time evolution of the spectra is associated with thermal ion relaxation and plasma cooling rather than with a geometric factor, as in the model of Reeves *et al.* (2002a). However, comparison with the observed spectrum of GRB 011211 in our model also requires an enhanced abundance of elements with α-particle nuclei. The observed absence of X-ray iron lines either requires an iron underabundance, which, according to current views of nucleosynthesis in supernovae, is unlikely, or can be explained by rapid line variability or line blurring against the background of a high nonthermal (power-law) continuum level, which cannot be eliminated by fitting the spectra.

If, following the model of Reeves *et al.* (2002a), we assume the presence of a metal-enriched ejection from the supernova that preceded the GRB, then, in the thermal model, it would be more natural to expect the appearance of photorecombination jumps rather than time-independent emission lines from a dense shell whose time evolution is associated with a purely geometric factor.

In our model, it would be most natural to assume that there is no blueshift of the X-ray lines relative to the gamma-ray source, if the clumpy medium is the circumstellar shell produced by a slow superwind from the GRB progenitor, especially if this is a neighboring dense cloud in a region of violent star formation. If, alternatively, the significant line blueshift will be proved, then it will actually serve as an argument for the supernova that preceded the GRB. However,

if these rapidly flying clouds are rich in hydrogen, then they cannot be placed so close to the gamma-ray source as in the model of Reeves *et al.* (2002a). Hence, such a supernova explosion must occur several months or even years before the GRB (the supranova model).

A proper explanation of the observed spectrum for the X-ray afterglow of GRB 011211 should probably be sought in terms of the combined model of the photoionization heating of the inhomogeneous medium around the GRB progenitor and the time-dependent collisional ionization in dense condensations with an optical depth to scattering of the order of unity. Of course, radiative transfer effects should be taken into account to compute the actual X-ray spectrum emergent from such media.

CONCLUSIONS

We computed the thermal X-ray radiation from an optically thin thermal plasma with time-dependent ionization in the coronal approximation. We showed that, when such a medium was instantly heated to high temperatures ($\sim$10 keV), a variable emission spectrum with lines of ions with α-particle nuclei and iron at different ionization stages appeared. We proposed a model for the thermal X-ray radiation from the early afterglow of GRB 011211, in which a narrow cone of gamma-ray radiation passes through the clumpy interstellar medium that falls within this cone and is not necessarily genetically associated with the GRB source. In this case, the time dependency of the X-ray spectrum (line and continuum decay) is attributed not only to a geometric factor (a delay from

different parts of the shell) but also to the physical time dependency of the collisional ionization and cooling of a plasma with a density $10^{11}-10^{12}$ cm^{-3}. The parameters of such a medium are similar to those of maser condensations in star-forming regions. We conclude that an allowance for time-dependent effects in the plasma is inevitable in the spectral X-ray diagnostics of the regions around GRBs in terms of thermal models.

ACKNOWLEDGMENTS

REFERENCES

1. K. Ambwani and P. G. Sutherland, Astrophys. J. **325**, 820 (1988).
2. I. Bains, R. J. Cohen, A. Louridas, *et al.*, Mon. Not. R. Astron. Soc. (2002) (in press); astro-ph/0211473.
3. R. Blandford, *Lighthouses of the Universe*, Ed. by M. Gilfanov, R. Sunyaev, and E. Churasov (Springer-Verlag, Berlin, 2002), p. 381.
4. S. I. Blinnikov, Surv. High Energy Phys. **15**, 37 (2000); astro-ph/9911138.
5. S. I. Blinnikov and O. S. Bartunov, Astron. Astrophys. **273**, 106 (1993).
6. S. I. Blinnikov, A. V. Kozyreva, and I. E. Panchenko, Astron. Rep. **43**, 739 (1999).
7. S. I. Blinnikov and K. A. Postnov, Mon. Not. R. Astron. Soc. **293**, L29 (1998).
8. K. N. Borozdin and S. P. Trudolyubov, astro-ph/0205208 (2002).
9. R. A. Chevalier and R. I. Klein, Astrophys. J. **234**, 597 (1979).
10. N. N. Chugai, S. I. Blinnikov, P. Lundqvist, *et al.*, Astrophys. J. (2003) (in press).
11. E. Costa, F. Frontera, J. Heise, *et al.*, Nature **387**, 783 (1997).
12. E. Falk and D. Arnett, Astrophys. J., Suppl. Ser. **33**, 515 (1977).
13. D. A. Frail, S. R. Kulkarni, R. Sari, *et al.*, Astrophys. J. Lett. **562**, L55 (2001).
14. T. R. Kallman and R. McCray, Astrophys. J., Suppl. Ser. **50**, 263 (1982).
15. P. Kumar and R. Narayan, astro-ph/0205488 (2002).
16. D. Lazzati, astro-ph/0210301 (2002).
17. D. Lazzati, S. Campana, and G. Ghisellini, Mon. Not. R. Astron. Soc. **304**, L31 (1999).
18. E. E. Lekht, N. A. Silant'ev, J. E. Mendosa-Torres, M. I. Pashchenko, and V. V. Krasnov, Astron. Astrophys. **377**, 999 (2001).
19. P. Mészáros, Ann. Rev. Astron. Astrophys. **40**, 137 (2002).
20. P. Mészáros and M. J. Rees, Astrophys. J. **476**, 232 (1997).
21. T. Piran, Phys. Rep. **333**, 529 (2000).
22. L. Piro, G. Garmire, M. Garcia, *et al.*, Science **290**, 955 (2000).
23. K. A. Postnov, Usp. Fiz. Nauk **169**, 545 (1998).
24. J. Pruet, S. E. Woosley, and R. D. Hoffman, astro-ph/0209412 (2002).
25. T. Rausher, A. Heger, R. D. Hoffman, and S. E. Woosley, Astrophys. J. **576**, 323 (2002); astro-ph/0112478 (2001).
26. M. J. Rees and P. Mészáros, Mon. Not. R. Astron. Soc. **258**, 41 (1992).
27. M. J. Rees and P. Mészáros, Astrophys. J. Lett. **430**, L39 (1994).
28. J. N. Reeves, D. Watson, J. P. Osborne, *et al.*, Nature **416**, 512 (2002a).
29. J. N. Reeves, D. Watson, J. P. Osborne, *et al.*, Astron. Astrophys. (2002b) (in press); astro-ph/0206480 (2002).
30. R. E. Rutledge and M. Sako, astro-ph/0206073 (2002).
31. G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley, Ney York, 1979).
32. H. Spruit, Astron. Astrophys. **341**, L1 (1998).
33. V. Tatischeff, astro-ph/0208397 (2002).
34. H. Umeda, K. Nomoto, T. G. Tsuru, and H. Matsumoto, Astrophys. J. **578**, 855 (2002).
35. V. V. Usov, Mon. Not. R. Astron. Soc. **267**, 1035 (1994).
36. J. van Paradijs, C. Kouveliotou, and R. A. M. J. Wijers, Ann. Rev. Astron. Astrophys. **38**, 379 (2000).
37. M. Vietri and L. Stella, Astrophys. J. Lett. **507**, L45 (1998).
38. D. Watson, J. N. Reeves, J. Osborne, *et al.*, Astron. Astrophys. **393**, L1 (2002).
39. Ya. B. Zeldovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1966).

Translated by V. Astakhov

Globular Clusters as Candidates for Gravitational Lenses to Explain Quasar–Galaxy Associations

Yu. L. Bukhmastova[*]

Astronomical Institute, St. Petersburg State University, Universitetskii pr. 28, Petrodvorets, 198504 Russia

Received September 25, 2002

Abstract—We argue that globular clusters (GCs) are good candidates for gravitational lenses in explaining quasar–galaxy associations. The catalog of associations (Bukhmastova 2001) compiled from the LEDA catalog of galaxies (Paturel 1997) and from the catalog of quasars (Veron-Cetty and Veron 1998) is used. Based on the new catalog, we show that one might expect an increased number of GCs around irregular galaxies of types 9 and 10 from the hypothesis that distant compact sources are gravitationally lensed by GCs in the halos of foreground galaxies. The King model is used to determine the central surface densities of 135 GCs in the Milky Way. The distribution of GCs in central surface density was found to be lognormal.
© *2003 MAIK "Nauka/Interperiodica".*

Key words: *quasar–galaxy associations, gravitational lensing, globular clusters, hidden halo mass.*

INTRODUCTION

What the nature of the association between some quasars and nearby galaxies is has remained an open question for more than three decades. The essence of this problem is that some quasars are observed in the immediate vicinity of galaxies, with the redshifts of the galaxies being much lower than those of the quasars close to them. A number of publications are devoted to this problem. Among these, we note the monographs on gravitational lensing by Bliokh and Minakov (1989) and Schneider *et al.* (1992) and the review article by Zakharov (1997).

Arp was among the first to gave rapt attention to close quasar–galaxy pairs. In his paper (Arp 1987), he provided examples of four galaxies at $z \sim 0.01$ that have two or three quasars at a distance of $\sim 2'$ from the galactic center. These quasars have $z \sim 1$ and are brighter than $m_v \sim 20^m$. Since the density of quasars with m_v brighter than 20^m is 20 objects per square degree, it seems unlikely to detect two or three quasars around any point. According to the estimates by Schneider *et al.* (1992), the probabilities of detecting two and three such quasars at an angular distance of $2'$ are 0.07 and 0.005, respectively.

Two major strategies of searching for quasar–galaxy pairs are being discussed in the literature. The first strategy of searching for such pairs is to consider quasars around galaxies and investigate the possible excess of the density of quasars around galaxies above the mean quasar density.

The overlap between the Second Reference Catalogue of Bright Galaxies containing 4364 objects and the catalog of quasars (Hewitt and Burbidge 1980) containing 1356 objects showed a statistical significance of the correlation between the sky projections of quasars at different redshifts and bright galaxies at $z < 0.05$ (Chu *et al.* 1984). However, the result was questioned, because the catalog included mostly radio-active quasars. Moreover, analysis of the catalogs of quasars and galaxies (Nieto and Seldner 1982) revealed no statistical significance for the associations. Nieto (1977, 1978, 1979) and Nieto and Seldner (1982) found no significant excess of the density of any type of quasars around nearby galaxies but revealed an excess of the density among optically variable quasars.

The second strategy of searching for pairs is to consider galaxies around quasars to determine whether there is an excess of the density of galaxies around quasars above the mean galaxy density. Rózyczka (1972) pointed out that there is such an excess of galaxies around nearby quasars and that low-redshift quasars are located in groups of galaxies at the same redshift. Tyson (1986) found an excess of the density of galaxies around low- and high-redshift quasars.

Burbidge *et al.* (1990) conducted a computer-aided search for close quasar–galaxy pairs with an angular separation of less than $10'$. As a result of this search, a catalog of 577 quasars and 500 galaxies was published. The observed number of associations is several times larger than the expected number of chance coincidences. For example, the probability of

[*]E-mail: `bukh@astro.spbu.ru`

detecting a randomly thrown object at an angular distance of $2'$ from a galaxy brighter than 15^m is estimated to be $\sim 10^{-3}$. Thus, the expected number of chance pairs for 5000 observed quasars is 5, while the catalog contains 38 such pairs. Such a simple estimate indicates that the physical association between quasars and galaxies in pairs is real. This association is also supported by the angular separation–galaxy redshift diagram constructed for 392 pairs. According to this observational relation, the mean linear galaxy–quasar separation in projection onto the sky, $\sim 100\,\mathrm{kpc}$, remains constant. These data led Burbidge *et al.* (1990) to conclude that quasars tend to be located in the halos of normal galaxies much more often than is expected for chance projections and that this physical association requires an explanation.

Based on the LEDA catalog of galaxies (Paturel 1997) containing some $100\,000$ objects and the catalog of quasars and active galactic nuclei (AGNs) (Veron-Cetty and Veron 1998), we previously (Bukhmastova 2001) searched for close quasar–galaxy pairs. The search for pairs was conducted in such a way that the linear separation between the galaxy and the quasar projection onto the galactic plane did not exceed 150 kpc. The following criteria were used:

(1) The spatial coordinates α, δ, and z are available for quasars and galaxies and the magnitudes m_Q are known for quasars;

(2) The quasar must be farther than the galaxy (i.e., $z_Q > z_G$);

(3) The quasars must be projected onto the halos of galaxies up to 150 kpc in size;

(4) The galaxy must have $z_G > 4 \times 10^{-4}$.

$77\,483$ galaxies and $11\,358$ quasars satisfied criterion (1). 1054 galaxies and 3164 quasars were in pairs, which amounted to 8382 quasar–galaxy pairs. Below, such pairs selected by using criteria (2)–(4) are called quasar–galaxy associations.

A statistical analysis of such associations based on computer simulations and theoretical calculations indicates that their observed number is much larger than the number of chance projections (1200 pairs), suggesting a physical association between the galaxy and quasar images.

When analyzing the catalog of close quasar–galaxy pairs (Burbidge *et al.* 1990), Baryshev and Ezova (1997) found the number of quasar–galaxy associations to depend on $a = z_G/z_Q$, where z_G and z_Q are the redshifts of the galaxy and the quasar from each pair. According to this dependence, galaxies in pairs tend to be located either close to the observer (so that $a < 0.1$) or close to the quasar ($a > 0.9$) and avoid the central position in the observer–quasar segment.

Based on new data for 8382 pairs, we previously (Bukhmastova 2001) analyzed the dependence of the number of pairs on $a = z_G/z_Q$. This dependence confirmed the previous result that there is an excess of pairs with $a < 0.1$ and $a > 0.9$.

There are at least two hypotheses that can explain the appearance of close quasar–galaxy pairs. One of these is that the quasars in associations are in the halos of nearby galaxies; thus, their redshifts are not distance indicators for these objects. Arp (1966, 1987, 1990) adhered to this hypothesis. The second hypothesis is that quasars in associations are only projected onto the halos of nearby galaxies; their redshifts remain cosmological in origin and the enhanced visible luminosity of such quasars is the result of gravitational lensing by objects in the galactic halos.

Here, we hypothesize that at least some of such pairs result from gravitational lensing. If this is the case, then the quasars in associations are the images of distant sources. The following question naturally arises: What astrophysical object for each pair acts as the source and what object is the gravitational lens?

Canizares (1981) was the first to attempt to explain the appearance of quasar–galaxy associations as due to gravitational lensing by galactic halo stars (microlensing). Thus, the individual halo stars falling on the line of sight acted as the lenses. Faint quasars were chosen as the sources; because of the low surface density of faint quasars, it was subsequently shown that the hypothesis of gravitational lensing could not explain the observed effect. It is important to note that microlensing makes it possible to weakly amplify the source images.

If the gravitational lenses are not stars but globular clusters (GCs), dwarf galaxies, or clustered hidden mass objects with masses $10^5-10^9 M_\odot$ ("mesolenses"), then it becomes possible to strongly amplify the sources (Baryshev and Ezova 1997). The King model is applicable to such objects. This model well fits the density profile in GCs, dwarf galaxies, and, possibly, other self-gravitating hidden mass objects. These objects are assumed to be located in the halos of galaxies in associations.

The existence and abundance of such hidden mass objects in galactic halos were discussed by Paolis *et al.* (1998). Dalal and Kochanek (2001) noted that the hidden mass in galactic halos could be clustered. The authors investigated seven radio sources with four-component images, in which they detected gravitational lensing. Six lensed sources exhibited an anomalous flux that could be explained by the presence of a substructure. Analysis showed that the masses of the objects of the gravitational lens substructure lay within the range characteristic of GCs, $10^6-10^9 M_\odot$.

The nature of the gravitational lenses responsible for the appearance of quasar−galaxy associations and for the assumed enhancement of the quasar flux can differ widely. Gravitational lenses can be made up of dark baryonic matter. In his review article, Carr (1994) discussed in detail the various components of dark baryonic matter. Snowballs produced through the condensation of cold hydrogen, brown dwarfs, Jupiter-like objects composed of hydrogen, dim red stars (M dwarfs), clouds of molecular hydrogen, neutron stars, black holes produced by the collapse of baryonic matter, and supermassive black holes can be these components.

The possibility that the gravitational lenses could be formed out of nonbaryonic cold dark matter (CDM) is not ruled out. The need for this mass component in the Universe follows from the prediction of critical total matter density $\Omega_{tot} = 1$ in the inflationary model and from the prediction of baryonic matter density $\Omega_{bar} = 0.05 h_{65}^{-2}$ in the standard model of primordial nucleosynthesis. Thus, 95% of the mass in the Universe can be in the mysterious nonbaryonic form. The Einstein cosmological constant or dark energy is currently believed to account for 70% of the dark matter (Peebles and Ratra 2002). The remaining 25% is accounted for by nonbaryonic CDM, for which the thermal energy and pressure can be disregarded compared to the energy of the rest mass. Dark energy represents a uniform background, while CDM is clustered to form structures of all scales (for a detailed review of the present views of dark energy, see Sahni and Starobinsky (2000)).

In the standard cosmological model, CDM is a dominant dynamical factor that governs the formation of galactic halos and the large-scale structure. High-resolution numerical calculations show that the hierarchy of CDM structures extends from the scales of galaxy superclusters down to the halos of individual galaxies composed of fundamental CDM clusters with masses in the range $10^6 - 10^8 M_\odot$ (Klypin et al. 1999; Morr et al. 1999, 2001).

Thus, if the clusters with such masses are assumed to account for 0.2%, then, as follows from an estimate of the lensing probability, on the order of several tens of thousands of quasars could be amplified by gravitational lensing (Baryshev and Ezova 1997).

To explain the large number of quasar−galaxy associations, we accept the Barnothy−Tyson hypothesis (Barnothy 1965) that quasars, at least partially, are the gravitationally amplified images of AGNs. Thus, in our model, compact massive objects located in the central regions of active galaxies, such as Seyfert galaxies and radio galaxies of all types, rather than faint quasars act as the sources. The number of these objects is large enough to provide the observed properties of the associations for amplification factors of 3−5 magnitudes. Kormendy (2000a, 2000b), Kormendy and Ho (2000), Ho and Kormendy (2000), and Bartusiak (1998) gave evidence that one might expect the presence of compact massive objects in the central regions of almost all galaxies. This large number of potential sources, naturally, exceeds the probability of occurrence of the lensing effect.

Stuart et al. (2002) assumed that black holes with masses $> 3 \times 10^9 M_\odot$ that are located in the central parts of quasars. The authors showed that a third of the distant quasars at $z \sim 6$ could be amplified by gravitational lenses on the line of sight, with the observed flux being enhanced by several tens of times or more.

This study is a logical continuation of the studies of the properties for quasar−galaxy associations initiated by Baryshev and Bukhmastova (1997, 2001). We argue that GCs are suitable candidates for gravitational lenses when we deal with quasar−galaxy associations. We discuss GC parameters, in particular, the distribution of GCs in central surface density. We give the possible locations of GC lenses relative to the centers of the galaxies in associations and analyze the types of galaxies in pairs. We summarize our main conclusions and discuss possible observational tests.

DATA ON GLOBULAR CLUSTERS

The idea to consider GCs as possible candidates for gravitational lenses was first put forward by Barnothy (1974) using the associations of the quasar 3C 455 with the galaxy NGC 7413 and an M15-type GC as an example. GCs as mesolens candidates were also considered by Yushchenko et al. (1998), Ugolnikov (2001), and Kurt and Ugolnikov (2000). Here, to obtain specific quantitative results, we use the parameters of 135 GCs in the Milky Way; thus, the previously used sample of 56 GCs analyzed by Baryshev and Ezova (1997) was expanded. For each of the GCs, we determined the core radius in pc and the central surface density by using the King model. These parameters are used for rough estimates.

The list of Milky Way GCs in Table 1[1] contains the following parameters: Column 1 gives the cluster number; column 2, the concentration parameter $c = \log(r_t/r_c)$, where r_t is the outer cluster radius and r_c is the core radius; column 3, the core radius r_c in arcminutes; column 4, the distance to the Sun R in kpc; column 5, the cluster mass in solar masses $m =$

---

[1] Table 1 is published in electronic form only and is accessible via ftp cdsarc.u-strasbg.fr/pub/cats/J (130.79.128.5) or http://cdsweb.u-strasbg.fr/pub/cats/J.
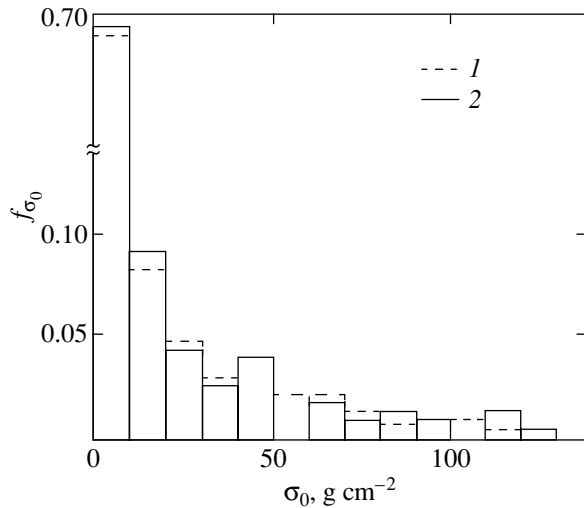
Fig. 1. The differential distribution function f_{σ_0} of the GC central surface density: (*1*) the lognormal distribution with $\sigma = 2.68$ and $\mu = 0.58$; (*2*) the observational data for 135 objects.
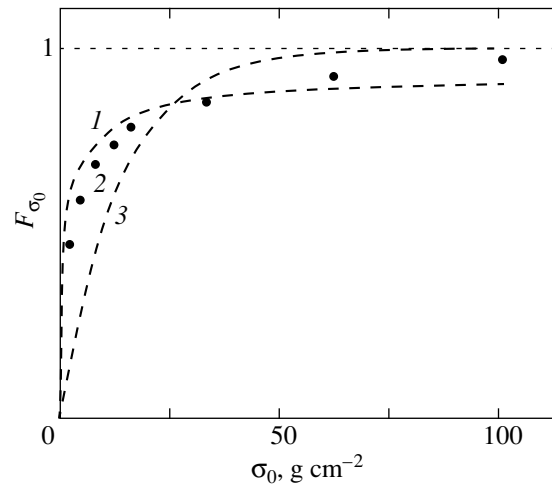

Fig. 2. The cumulative distribution function F_{σ_0} of the GC central surface density: (*1*) the lognormal distribution with $\sigma = 2.68$ and $\mu = 0.58$, (*2*) the observational data, (*3*) the exponential distribution $\beta \exp(-\beta x)$ with $\beta = 0.056$ and $x = \sigma_0$.

$\log(M/M_\odot)$; column 6, the core radius R_c in pc; and column 7, the central surface density σ_0 in g cm^{-2}.

We took the data on (ID), (c), (r'_c), and (R) from Harris (1996) and on ($\log(M/M_\odot)$) from Mandushev *et al.* (1991). We calculated (R_c) from (r_c) and (R) and (σ_0) from formulas of the King model. The mass of an object in the King model is defined as

$$M = \pi r_c^2 \sigma_0 \Lambda_0, \qquad (1)$$

where

$$\Lambda_0 = 2\ln(r_t/r_c) - 3. \qquad (2)$$

For $r_t/r_c \sim 100$, the surface density can be calculated with a high accuracy from the formula

$$\sigma_0 = \frac{10^{-5}M/M_\odot}{r_c^2\,[\text{pc}]}. \qquad (3)$$

The Surface-Density Function

The results of Table 1 reveal that $\sim 72\%$ of the Milky Way GCs have central surface densities up to 10 g cm^{-2}. Figure 1 shows an empirical GC surface-density distribution (histogram *1*). The sharp decrease at low σ_0 and the long tail at high σ_0 suggests that the distribution of this quantity may be lognormal (histogram *2* in Fig. 1). The distribution of σ_0 is then

$$f_{\sigma_0}(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \qquad (4)$$

where μ and σ are the distribution parameters and $x = \sigma_0$. The values $\mu = 0.58$ and $\sigma = 2.68$ were determined by the maximum-likelihood method.

To test the hypothesis of a lognormal distribution of the central surface density, we used the χ^2 test. According to this test, the GC data are consistent with the hypothesis of a lognormal σ_0 distribution with 95% confidence (at an $\alpha = 5\%$ significance level).

Figure 2 shows the cumulative distribution function of the GC surface density. Curve *1* represents the lognormal distribution with the above parameters and curve *2* represents the experimental dependence for 135 objects. For comparison, this figure
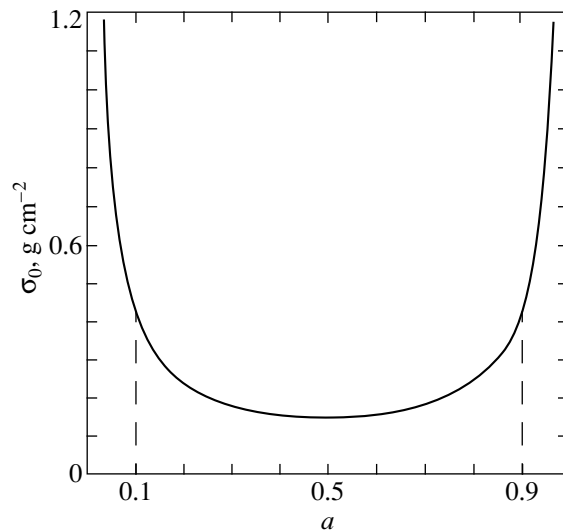

Fig. 3. GC central surface density versus GC position on the observer−distant source axis: $\sigma_{0cr}(a, z_S) = \dfrac{0.077}{a(1-a)z_S}$. The sources are assumed to be at $z_S = 2$.

also shows curve *3* for the exponential distribution function $\beta e^{-\beta x}$, where $x = \sigma_0$, with $\beta = 0.056$.

Our analysis leads us to conclude that the lognormal distribution well describes the experimental data on the central surface density of Milky Way GCs. Our significantly expanded (compared to the sample of Baryshev and Ezova (1997)) sample of GCs confirmed that the distribution function of the GC central surface density is lognormal and allowed the parameters of this distribution to be refined.

How to physically interpret this law remains an open question; i.e., whether it is the result of selection effects or is inherent in the physical formation and evolution of GCs.

The Critical Central Surface Density of a Lens

The central surface density of the putative lenses considered in the previous section is involved in many gravitational lensing formulas.

The main manifestations of gravitational focusing include the possible amplification of the image of a distance source. A significant amplification factor is possible only if the observer is not between the lens focus and the lens. Since the quasar brightness is assumed to be significantly amplified, by 3−5 magnitudes, when interpreting quasar−galaxy associations, the position of the lens focus relative to the observer becomes important. In turn, the focus position is determined by the central surface density of the gravitational lens. Thus, we can determine the minimum lens density that corresponds to the case where the observer is at the lens focus (and, hence, a large image amplification is achieved):

$$\sigma_{0cr}(a, z_S) = \frac{0.077}{a(1-a)z_S} \text{ g cm}^{-2}. \quad (5)$$

Assume that the quasars are located at the distance corresponding to the redshift $z_S = 2$. A plot of this dependence for $z_S = 2$ is shown in Fig. 3. We see from the figure that, if the lenses in quasar−galaxy pairs tend to be located at distances $a < 0.1$ and $a > 0.9$ (as follows from Bukhmastova (2001) and Baryshev and Ezova (1997)), then this corresponds to a lens surface density $\sigma_0 > 0.42 \text{ g cm}^{-2}$. An upper limit can be determined if we set $a = 10^{-4}$ and $z_S = 2$. In this case, $\sigma_0 < 385 \text{ g cm}^{-2}$, which is in close agreement with the central surface density for GCs (Table 1).

Thus, the fact that the required central surface densities of putative gravitational lenses lie within the range of their values for GCs is one of the arguments that GCs are good candidates for gravitational lenses.

The number of quasar−galaxy pairs found previously (Bukhmastova 2001) is 8382. These pairs include 3164 quasars. Assume that each quasar in an association is lensed and that GCs in the halos of galaxies in associations act as the lenses. Given the distances to the galaxy and the quasar, we can then calculate the critical central densities of the putative GC lenses for each pair that correspond to the condition that the observer be near the focus (and, hence, a large amplification factor is achieved). If our assumption is valid, then at least 3164 calculated critical densities must be characteristic of the GC central surface densities shown in Fig. 1.

Below, we present the calculated critical central surface densities of the putative lenses for 8382 quasar−galaxy pairs. The pairs can be found at http://www.astro.spbu.ru/staff/Baryshev/gl_dm/htm. The left column gives the critical density in g cm^{-2} calculated from formula (5), and the right column gives the number of lenses with this density:

0−10	708
10−20	688
20−30	556
30−40	375
40−50	272
50−60	226
60−70	295
70−80	212
80−90	470
90−100	396
>100	4184

It turns out that 3332 pairs give central surface densities up to 80 g cm^{-2}, which closely match the central surface densities for GCs. This is yet another argument for the theory of the gravitational lensing of AGNs by GCs in galactic halos.

Dwarf galaxies, like GCs, also belong to the objects with a King mass distribution. They are difficult to detect because of their low luminosities. The data on dwarf galaxies (Ferguson and Binggeli 1994; Gallagher and Wyse 1994) are too scarce to reach a final conclusion about their central surface density. Dwarf galaxies contain a large amount of hidden mass, which leads to disagreements regarding the determination of the core radius. The parameters of dwarf galaxies from Yakovlev *et al.* (1983) give an idea of their basic parameters:

$\log M/M_{\odot}$	r_c, kpc	$\log r_t/r_c$	σ_0, g cm^{-2}
5.0−9.5	0.2−1	0−0.7	10^{-4}−0.2

If dwarf galaxies with characteristic surface densities of 10^{-4}−0.2 g cm^{-2} acted as the lenses, then an excess of pairs with $a \sim 0.5$ would be observed in

the best case; i.e., galaxies would prefer the central position on the observer–quasar segment, which is in conflict with the observational data from the catalog of quasar–galaxy associations. According to these data, $a < 0.1$ or $a > 0.9$. Thus, for dwarf galaxies to be considered as good candidates for gravitational lenses, they must contain a large amount of hidden mass. The latter must be strongly concentrated toward the center in order to increase the central surface density to $1-80\,\mathrm{g\,cm^{-2}}$.

If clustered transparent hidden mass objects act as the lenses, then we can specify the boundaries of their central surface densities: $\sigma_0 > 0.28$ ($z_Q \sim 3$), $\sigma_0 > 1.68$ ($z_Q \sim 0.5$). The higher is the central surface density of a lens, the closer it is to the observer or the quasar. If the halo objects of our Galaxy (or the halo objects of the galaxy in which a quasar may be located) are assumed to be the lenses, then the central surface densities of these lenses must be $\sim 10^5-10^8\,\mathrm{g\,cm^{-2}}$, which correspond to the parameters of actually existing objects—asteroids and minor planets. At present, such objects are difficult to invoke as the gravitational lenses to explain quasar–galaxy associations, because the model of a pointlike lens should most likely be used for them. For the lensing probability in this model to be high enough, the halo mass must be too large.

The Caustics and Cross Sections of Gravitational Lenses

If we use the models of a pointlike gravitational lens applicable to stars or those of an isothermal sphere applicable to galaxies to explain the enhanced brightness of the quasars from associations, then an accurate alignment of the observer, the lens, and the source on one straight line (on the axial caustic) is required to achieve large amplification factors.

An important feature of the gravitational lenses with a King mass distribution is that, in this case, no accurate alignment of the observer, the lens, and the source on one straight line is required to achieve large amplification factors. This is because large amplification factors are achieved near the caustics. In addition to the axial caustic, GC lenses have the conical caustic. Its presence and the wide range of possible GC central surface densities ensure great freedom of the observer's position relative to the lens and the source and, hence, increase the probability of strong lensing (see Baryshev and Ezova (1997) for more details).

The probability that a distant object will be amplified by a foreground lens depends on the lens cross section. The larger this cross section, the higher the probability. For lenses with a King mass distribution, there is a total cross section that consists of the
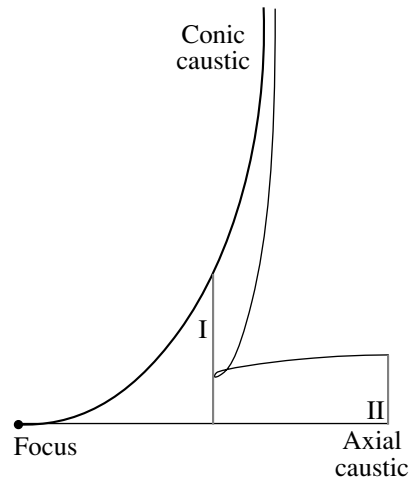

Fig. 4. A scheme of the caustic regions for a lens with a King mass distribution: I, the near-focus region where the axial and conical caustics meet; a large cross section can be reached, because these regions can merge together; and II, the region of maximum cross section of the axial caustic ($a = 0.5$).

cross sections of the two (axial and conical) caustic regions. Therefore, the following two regions of maximum cross section exist for lenses with a King mass distribution: region I where all caustics meet (near-focus region) and region II on the line of sight halfway between the observer and the lens. This is schematically shown in Fig. 4. To take region II into account requires an accurate alignment of the observer, the lens, and the source on one straight line. In addition, as was mentioned above, a much lower central surface density than the GC surface density is required for the lenses located halfway between the observer and the source. To take region I into account, we need not impose stringent requirements on the observer's position relative to the source–lens straight line. Besides, the GC central surface densities exactly correspond to the case where the observer falls within region I, which explains the preferential positions of the galaxies in associations close to the observer and the quasar.

Thus, the presence of a conical caustic in GC lenses gives them yet another advantage over other possible objects.

AROUND WHAT TYPES OF GALAXIES SHOULD GC LENSES BE SEARCHED FOR?

The new catalog of quasar–galaxy associations contains 8382 pairs, which comprise 3164 quasars and 1054 galaxies. If galactic halo GCs are considered as the lenses located in the halos of these galaxies, then the following question naturally arises: What types of galaxies are encountered more often in

Table 2. The distribution of LEDA galaxies in type

1	2	3	4	5	6
−5	2500	0.061	37	0.048	0.79
−4	1474	0.036	15	0.019	0.54
−3	2872	0.070	34	0.044	0.63
−2	3295	0.081	48	0.063	0.77
−1	1554	0.038	24	0.031	0.82
0	2053	0.050	30	0.039	0.78
1	1968	0.048	25	0.033	0.67
2	2327	0.057	32	0.042	0.73
3	3967	0.098	58	0.076	0.78
4	5232	0.129	57	0.075	0.58
5	5995	0.148	51	0.067	0.45
6	2730	0.067	55	0.072	1.07
7	1053	0.026	38	0.050	1.92
8	1171	0.028	50	0.066	2.88
9	918	0.022	67	0.088	3.90
10	1366	0.033	136	0.179	5.32

pairs? Or, what galaxies contain GCs around them and at what distances from the galactic center should GCs be searched for?

The original LEDA catalog of galaxies (Paturel 1997), which was used to compose pairs, contains data on the types of 40 475 galaxies. Types −5, 5, and 10 correspond to elliptical, spiral, and irregular galaxies, respectively. Table 2 contains data on the number of galaxies of each type: column 1 gives the type of galaxy; column 2, the number of galaxies of a given type in the original LEDA catalog of galaxies; column 3, the fraction of the galaxies of a given type of the total number of galaxies; column 4, the number of galaxies of a given type in quasar–galaxy pairs; column 5, the fraction of the galaxies in pairs of the total number of galaxies in pairs for which data on their types are available (there are 757 such galaxies); column 6 reflects the relative contribution of the galaxies of each type in pairs. We see from the sixth column of this table that the relative number of GC lenses in the types galaxies from −5 to 5 is approximately the same. Starting from type 6, the relative number of putative lenses increases; in the irregular galaxies of type 10, the number of GCs is expected to be approximately a factor of 7 larger than that in the spiral and elliptical galaxies of each type.

Table 3[2] presents irregular galaxies of types 9 and 10 in quasar–galaxy associations: column 1 gives the galaxy name; column 2, the galaxy type; column 3, the galaxy redshift; column 4, the assumed distances (in kpc) from the galactic center at which the clustered object (CO) lenses are located; column 5, the number of quasars projected onto the galaxy; and column 6, the number of compact star clusters detected in these galaxies to date. The data of this table make it possible to check specific galaxies for the presence of clustered objects in their halos and, in particular, in the immediate vicinity of the quasars.

The absolute magnitudes of the galaxies of types 9 and 10 in pairs lie within the range from −11 to −21, with ∼80% of them having absolute magnitudes from −15 to −19. The distances from the centers of the irregular galaxies at which one might expect GCs are ∼50−100 kpc. To all appearances, these are young globular clusters.

The currently available observational data confirm our conclusion that the halos of irregular galaxies contain clustered objects with masses typical of the GCs. Based on the Hubble Space Telescope (HST) data, Hunter *et al.* (2001) investigated the irregular galaxy NGC 1569. In this galaxy, they found 47 clusters whose core radii lie within the range typical of the GCs. The absolute magnitude of the galaxy is $M_V = -18$. The parameters of two clusters in this galaxy are known: the core radii are 2.3 and 3.1 pc and the masses are $\sim 3 \times 10^5 M_\odot$. These authors provided data on seven more irregular galaxies in which compact star clusters were detected. Sixty compact star clusters were identified in the irregular galaxy NGC 4449, 49 of which are presumed to be GCs. NGC 4214 (PGC 039225) is an irregular galaxy with active star formation. Twenty nine compact star clusters were identified around this galaxy; Hunter *et al.* (2001) believe these clusters to be young. Similarly, three clusters were found around the galaxy DDO 50 (PGC 023324), 16 clusters were found around NGC 1705 (PGC 016282), three clusters were found around DDO 168 (PGC 046039), four clusters were found around DDO 165 (PGC 045372), and one compact cluster was found in DDO 75 (PGC 029653) and in WLM. All of these compact objects have radii $r < 15$ pc and M_V brighter than -10^m. Most of the star clusters are young with sizes and masses comparable to those of GCs. Thus, the number of GCs in irregular galaxies may be much larger than is generally thought to be.

The irregular galaxies NGC 4214, DDO 50, NGC 1705, DDO 168, and DDO 75 with compact

---

[2]Table 3 is published in electronic form only and is accessible via ftp cdsarc.u-strasbg.fr/pub/cats/J (130.79.128.5) or http://cdsweb.ru-strasbg.fr/pub/cats/J.

Table 4. The spiral and elliptical galaxies in whose halos GCs and quasars were detected

1	2	3	4
PGC041297 (NGC4478)	E2	3	Present
PGC044324 (NGC4840)	SAB0	1	10^3-10^4
PGC043008 (NGC4673)	E1-2	1	10^3-10^4
PGC044553	E	1	10^3-10^4
PGC013344 (NGC1387)	S0	2	385 ± 80
PGC013418 (NGC1399)	E1pec	1	5340 ± 1780
PGC013433 (NGC1404)	E1	1	950 ± 140
PGC024930 (NGC2683)	Sb	4	310 ± 100
PGC032226 (NGC3377)	E5	1	210 ± 50
PGC032256 (NGC3379)	E1	6	260 ± 140
PGC036487 (NGC3842)	E3	2	$14\,000 \pm 2500$
PGC039764 (NGC4278)	E1	2	1050 ± 120
PGC039246 (NGC4216)	Sb	129	620 ± 310
PGC041327 (NGC4486)	E0	1	$13\,000 \pm 500$
PGC041968 (NGC4552)	E0	34	2400
PGC042051 (NGC4564)	E6	3	1000 ± 300
PGC042628 (NGC4621)	E5	7	1900 ± 400
PGC000218 (NGC7814)	Sab	1	500 ± 160

star clusters detected in their halos investigated by Hunter *et al.* (2001) are members of the quasar–galaxy associations found previously (Bukhmastova 2001); several quasars are projected onto each of them.

Table 4 lists the spiral and elliptical galaxies in whose halos GCs and quasars have been detected to date: column 1 gives the galaxy name; column 2, the galaxy type; column 3, the number of quasars projected onto the galaxy; and column 4, the number of GCs detected in the halo.

The data on the presence of GCs were taken from astro-ph/0206140 for the first galaxy and from astro-ph/0112209 for the second, third, and fourth galaxies; for the remaining galaxies, we took the data from Harris (1996).

Thus, the data of Tables 3 and 4 support the hypothesis that the associations of distant quasars with nearby galaxies can be provided by compact star clusters (GCs), which act as gravitational lenses. Therefore, a more thorough analysis of the properties of these associations may give an independent scenario for the formation and evolution of globular clusters or may shed light on the nature of the clustered hidden mass in galactic halos.

MAIN CONCLUSIONS AND OBSERVATIONAL TESTS

As a result of our analysis of the properties of quasar–galaxy associations and according to the assumptions about their possible nature made in the Introduction, we favor GCs as good candidates for gravitational lenses located in the halos of galaxies in associations. Apart from the aforesaid, the following arguments support this assumption:

(1) GCs are actually existing observed objects. Their number in the halos of some galaxies can reach several thousand (Harris 1996). This provides a good possibility of using them as good candidates for gravitational lenses, because the more potential candidates for gravitational lenses in the halo, the larger the total cross section of the lenses and, hence, the higher the probability of lensing distant sources.

(2) The possible manifestations of gravitational lensing include source image splitting. For GCs described by the King model, one, two, or three images of a distant source can appear. The characteristic linear separation between the possible images is defined via the Chwolson–Einstein ring radius $\xi = \sqrt{2r_g \dfrac{R_{\mathrm{OL}} R_{\mathrm{LS}}}{R_{\mathrm{OS}}}}$, where R_{OL}, R_{LS}, and R_{OS} are the

distances from the observer to the lens, from the lens to the source, and from the observer to the source, respectively, and r_g is the gravitational radius of the lens. If the GC mass is assumed to be $10^6 M_\odot$, then for $z_{OL} = 10^{-3}$–0.1 (the redshifts of the galaxies in associations), ξ varies between 10^{-2} and 0.4 pc. The quasar sizes are very small ($\sim 10^{-3}$ pc). Since the possible amplification factor for a source is equal to the ratio of the area of its image to the source area, GCs make an amplification by 3^m–10^m possible. This favorably distinguishes them from stars, which are often invoked as the lenses to explain quasar–galaxy associations. Stars described by the model of a point-like lens (microlensing) give only a weak amplification of the source and cannot explain the enhanced quasar luminosity in this way.

(3) The GC central surface densities closely correspond to the condition that the observer be near the lens focus. This gives large amplification factors 3^m–5^m and, hence, allows us to explain the enhanced quasar luminosity and their leftward displacement in the Hubble diagram $z = f(m)$ for galaxies and quasars (see Bukhmastova 2001).

(4) The preferential GC distances from the galactic centers are less than 50 kpc. This is in agreement with the fact that most of the nearby quasars from associations with $z_G/z_Q > 0.9$ are also projected onto galactic halos with radii up to 50 kpc (Bukhmastova 2001).

(5) Since the critical central surface densities of the putative lenses calculated for quasar–galaxy associations match the GC central surface densities, we can assume that the association between distant quasars and nearby galaxies is provided by halo GCs.

(6) The expected light variability due to the motion of a GC as a whole has a time scale of more than a thousand years, because the time scale estimate is given by the expression

$$t_{var} \sim \frac{R_{OL}\Theta_{Ch-E}}{v} \sim 10^3 \left(\frac{M}{10^6 M_\odot}\right)^{1/2}$$
$$\times \left(\frac{D'}{10 \text{ Mpc}}\right)^{1/2} \left(\frac{v}{10^3 \text{ km s}^{-1}}\right)^{-1} \text{ yr,}$$

where $D' = \dfrac{R_{LS}R_{OL}}{R_{OS}}$ is the reduced distance to the lenses and v is the lens velocity. However, light variability on time scales of less than one year is possible because of microlensing by individual GC stars.

(7) Since GCs belong to the galactic halos, absorption lines can emerge in the spectra of quasars with z_{abs} corresponding to this galaxy. In this case, an excess of systems with $z_{abs}/z_{em} > 0.9$ or $z_{abs}/z_{em} < 0.1$ must be observed. This theoretical assumption follows from the dependence (Baryshev and Ezova 1997) of the differential lensing probability on $a = z_L/z_S$, where z_L and z_S are the redshifts of the lens and the source, respectively. Note that if the distribution of galaxies on the line of sight is fractal, then this can also cause the number of systems with $a > 0.9$ and $a < 0.1$ to increase. Thus, gravitational lensing by objects with a King mass distribution and the fractal distribution of galaxies with their surrounding lenses can mutually enhance this effect. More specifically, in this case, galaxies in pairs with quasars will avoid the central position on the observer–quasar line of sight, as confirmed by analysis of the catalogs of associations (Burbidge *et al.* 1990; Bukhmastova 2001).

(8) Based on the observed widths of the 21-cm absorption lines in the spectra of some quasars and assuming the clouds responsible for their appearance to be gravitationally bound, Komberg (1986) estimated the optical luminosities and sizes of these clouds. The luminosities, sizes, and masses were found to be typical of GCs. These authors also estimated the H I mass to be only a few solar masses. Even in old GCs one might expect such an amount of neutral hydrogen. It thus follows that the 21-cm lines can be detected in nearby GCs. The coincidence of the 21-cm absorption lines and the systems with absorption lines of metal ions may also stem from the fact that GCs, which can be gravitational lenses (thus, can be responsible for the enhanced luminosities of some quasars), fall on the line of sight.

(9) An increased number of GCs, a factor of 6–7 larger than that for spiral and elliptical galaxies of each type (Table 2), must presumably be observed in irregular galaxies of types 9 and 10. These GCs are presumed to be located at distances of ~ 100 kpc from the galactic centers and to have central surface densities $\sigma_0 \sim 1$–200 g cm^{-2}. This follows from our analysis of the types of galaxies in pairs and from our calculations of the critical surface density of the putative lenses for each quasar–galaxy pair (formula (5)), as well as from the available HST observational data (Hunter *et al.* 2001).

(10) The Chandra observations revealed X-ray sources in six Galactic GCs. Thus, GCs can be detected in galactic halos by X-ray sources (Revnivtsev *et al.* 2002).

(11) Quasars in associations with $a > 0.9$, i.e., close to galaxies in redshift, are observable in the radio frequency range. This allows the possible splitting of the core structure in these quasars to be detected at the millisecond level. These quasars can be found at http://www.astro.spbu.ru/staff/baryshev/gl_dm.htm.

It follows from our analysis of the data on 135 Milky Way GCs given in Table 1 that the GC distribution in central surface density is lognormal.

I hope that all of the arguments listed above will help to explain the appearance of quasar−galaxy associations and to refine the nature of the candidates for gravitational lenses. Apart from GCs, the latter can include dwarf galaxies and clustered hidden mass objects.

ACKNOWLEDGMENTS

REFERENCES

1. H. Arp, Astrophys. J. **148**, 321 (1967).
2. H. Arp, *Quasars, Redshifts and Controversies, Interstellar Media* (Berkelley, Ca., 1987).
3. H. Arp, Astron. Astrophys. **229**, 93 (1990).
4. J. M. Barnothy, Astron. J. **70**, 666 (1965).
5. J. M. Barnothy, Bull. Am. Astron. Soc. **6**, 212 (1974).
6. M. Bartusiak, Astronomy, June, 42 (1998).
7. Yu. V. Baryshev and Yu. L. Ezova (Yu. L. Bukhmastova), Astron. Zh. **74**, 497 (1997) [Astron. Rep. **41**, 436 (1997)]; astro-ph/0206348.
8. P. V. Bliokh and A. A. Minakov, *Gravitational Lenses* (Naukova Dumka, Kiev, 1989).
9. Yu. L. Bukhmastova, Astron. Zh. **78**, 675 (2001) [Astron. Rep. **45**, 581 (2001)]; astro-ph/0206456.
10. G. Burbidge, A. Hewitt, J. V. Narlikar, and P. Das Gupta, Astrophys. J., Suppl. Ser. **74**, 675 (1990).
11. C. R. Canizares, Nature **291**, 620 (1981).
12. B. Carr, Ann. Rev. Astron. Astrophys. **32**, 531 (1994).
13. Y. Chu, X. Zhu, G. Burbidge, and A. Hewitt, Astron. Astrophys. **138**, 408 (1984).
14. N. Dalal and C. S. Kochanek, astro-ph/0111456 (2001).
15. H. C. Ferguson and B. Binggeli, Astron. Astrophys. Rev. **6**, 67 (1994).
16. J. S. Gallagher and R. F. G. Wyse, Publ. Astron. Soc. Pac. **106**, 1225 (1994).
17. W. E. Harris, Astron. J. **112**, 1487 (1996); http://www.physics.mcmaster.ca/Globular.html.
18. A. Hewitt and G. Burbidge, Astrophys. J., Suppl. Ser. **43**, 57 (1980).
19. L. C. Ho and J. Kormendy, astro-ph/0003267 (2000).
20. D. A. Hunter, O. H. Billet, and B. G. Elmegreen, astro-ph/0112260 (2001).
21. A. Klypin, S. Gottloeber, A. Kravtsov, and A. Khohlov, Astrophys. J. **516**, 516 (1999).
22. B. V. Komberg, Astrofizika **24**, 321 (1986).
23. J. Kormendy, astro-ph/0007400 (2000a).
24. J. Kormendy, astro-ph/0007401 (2000b).
25. J. Kormendy and L. C. Ho, astro-ph/0003268 (2000).
26. V. G. Kurt and O. S. Ugol'nikov, Kosm. Issled. **38**, 227 (2000).
27. G. Mandushev, N. Spassova, and A. Staneva, Astron. Astrophys. **252**, 94 (1991).
28. B. Moor, S. Ghigna, F. Governato, *et al.*, Astrophys. J. Lett. **524**, L19 (1999).
29. B. Morr, C. Calcaneo-Roldan, J. Stade, *et al.*, astro-ph/0106271 (2001).
30. J.-L. Nieto, Nature **270**, 411 (1977).
31. J.-L. Nieto, Astron. Astrophys. **70**, 219 (1978).
32. J.-L. Nieto, Astron. Astrophys. **74**, 152 (1979).
33. J.-L. Nieto and M. Seldner, Astron. Astrophys. **112**, 321 (1982).
34. F. De Paolis, G. Ingrosso, Ph. Jetzer, and M. Roncadelli, Astrophys. J. **500**, 59 (1998).
35. J. Paturel, Astrophys. J., Suppl. Ser. **124**, 109 (1997).
36. P. J. E. Peebles and B. Ratra, astro-ph/0207347 (2002).
37. M. G. Revnivtsev, S. P. Trudolyubov, and K. N. Borozdin, Pis'ma Astron. Zh. **28**, 276 (2002) [Astron. Lett. **28**, 237 (2002)].
38. M. Rózyczka, Acta Astron. **22**, 93 (1972).
39. V. Sahni and A. A. Starobinsky, Int. J. Mod. Phys. D **9**, 373 (2000); astro-ph/9904398.
40. P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses* (Spriger-Verlag, New York, 1992).
41. J. Stuart, B. Wyithe, and A. Loeb, astro-ph/0203116 (2002).
42. J. A. Tyson, Astron. J. **92**, 691 (1986).
43. O. S. Ugol'nikov, Preprint No. 17 (Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 2001).
44. M. P. Veron-Cetty and P. Veron, *A Catalogue of Quasars and Active Nuclei* (European Southern Observatory (ESO), Garching, 1998), 8th ed., ESO Scientific Report Series **18** (1998).
45. D. G. Yakovlev, L. G. Mitrofanov, S. A. Levshakov, and D. A. Varshalovich, Astrophys. Space Sci. **91**, 133 (1983).
46. A. V. Yushchenko, Yu. Baryshev, and A. A. Raikov, Astron. Astrophys. Trans. **17**, 9 (1998).
47. A. F. Zaharov, *Gravitational Lenses and Microlenses* (Yanus-K, Moscow, 1997).

Translated by V. Astakhov

Search for Groups of Extragalactic Radio Sources

A. G. Gubanov[1*], A. I. Kopylov[2], and A. S. Potapov[1]

[1]*Astronomical Institute, St. Petersburg State University, Universitetskii pr. 28, Petrodvorets, 198504 Russia*
[2]*Special Astrophysical Observatory, Russian Academy of Sciences,*
Nizhnii Arkhyz, Karachaevo-Cherkessian Republic, 369167 Russia
Received September 12, 2002

Abstract—We explore the possibility of searching for groups of radio sources from the FIRST catalog on angular scales $1'-5'$. We developed an efficient method of searching for such groups that takes into account the need for combining the components of extended sources represented in the catalog by separate objects. We found 31 groups of radio sources with angular sizes $<5'$ that contain no fewer than five sources with flux densities ≥ 3 mJy. This number is at least triple the expected number of such groups for a random Poisson distribution of radio sources in the sky. The prospects for using groups of radio sources to detect and study distant systems of galaxies are discussed. © *2003 MAIK "Nauka/Interperiodica"*.

Key words: *extragalactic radio sources, clusters of galaxies.*

INTRODUCTION

Interest in investigating the clustering of extragalactic radio sources arose immediately after the first radio-astronomical sky surveys and the appearance of catalogs with a large number of objects. However, the results obtained to date have pertained mainly to the study of the large-scale structure (see, e.g., Cress *et al.* 1996; Magliocchetti *et al.* 1998). This is because, on the one hand, many radio sources are observed at large distances (at redshifts $z \sim 1-2$) and, on the other hand, the occurrence of an intense radio emission is a rare event among galaxies even for such a recent deep radio sky survey as FIRST (Faint Images of the Radio Sky at Twenty Centimeters; Becker *et al.* 1995; White *et al.* 1997).

On small ($<5'$) angular scales, the clustering properties of radio sources have not yet been adequately studied. Apart from the relative rareness of extragalactic radio sources, this stems from the fact that, first, an appreciable fraction of radio sources on these angular scales have a complex, multicomponent structure and, second, the faint sources that determine these properties are difficult to distinguish against the background of the various kinds of noise in the available data (e.g., the radio telescope sidelobes). At the same time, investigating clustering on these angular scales is of particular interest, because they correspond to the scales of rich clusters (or protoclusters) at high ($z > 0.5$) redshifts and to the sizes of compact groups of galaxies in the local neighborhood ($z \sim 0.1$).

Studies of groups of physically associated galaxies with active nuclei (as, e.g., radio galaxies) are undoubtedly of value in the questions of their origin and evolution, as well as the connection to their environment—the central questions of modern astrophysics. The very presence of galaxy systems with an anomalously large number of radio sources (compared to their random number) will be a strong argument for the existence of an external factor responsible for the emergence of this activity, which is associated with the interaction of the galaxies themselves and their systems. Detailed studies of such groups will also provide valuable information on their origin, structure, and dynamics.

Note that the closest compact groups of galaxies can hardly be responsible for a significant clustering of radio sources from the FIRST catalog on small angular scales (if detected). A radio study of these objects at the FIRST sensitivity level indicates that if such groups contain a radio source, then it is generally one in a group and is identified with one of the massive galaxies, and that the currently high galaxy density has no significant effect on the occurrence of an intense radio emission (see, e.g., Menon 1999; Tovmassian 2000).

What the population of more distant groups and clusters of galaxies by radio sources is still remains an open question. However, intense distant extragalactic radio sources were found to be in a relatively rich environment (compared to local sources). There is also mounting evidence that mergers and close interactions of galaxies (more frequent at early cosmological epochs and in richer systems) play a key role

*E-mail: ag@astro.spbu.ru

in the emergence of activity in galactic nuclei and in intense starbursts. These two phenomena are known to be accompanied by an enhanced radio emission. Direct attempts have been made to find protoclusters of galaxies near distant radio galaxies and quasars (some of them have been successful; see, e.g., Kurk *et al.* 2001; Komberg *et al.* 1996). This serves as an additional reason for searching for relatively close groups of radio sources.

At present, about 10% of the rich galaxy clusters contain intense radio sources (as a rule, one source) whose luminosities correspond to the flux densities at the FIRST limit if they are placed at cosmological distances. It is also well known that the density of radio sources at $z \approx 1-3$ could be several hundred times higher than the local density (Rawlings 2000). Thus, there is a chance of detecting several comparatively intense sources in each distant protocluster (at the FIRST sensitivity level) and this can be a new method for their search (see, e.g., Croft *et al.* 2001). Recall that the cluster sizes at such distances correspond to angular scales of several arcminutes.

The radio emission from galaxies with intense star formation located in groups at $z \sim 0.1$ corresponds to these angular scales and radio flux densities. Therefore, searching for groups of radio sources on arcminute angular scales is of current interest.

Here, we give numerical estimates for the cosmological model with the Hubble constant $H_0 = 60$ km s^{-1} Mpc^{-1} and the density parameters $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

THE CHOICE OF SEARCHING PARAMETERS

A major goal of this study is to find relatively compact and populated groups of radio sources from the latest version of the FIRST catalog that will be candidates for searching for physically associated systems of active galactic nuclei (AGNs). Subsequently, we intend to investigate the basic properties of such systems.

The latest (October 2001) version of the FIRST catalog of radio sources contains 771 076 objects with a mean surface density of 92 sources per square degree (for a limiting flux density of ≈ 1 mJy). This corresponds to a survey area of about 8380 square degrees.

Let us choose the range of scales for groups of radio sources that are most suitable for our purposes.

Of course, the smallest scales ($1''-60''; 5-500$ kpc) are of considerable interest. There are impressive examples, as in the case of the radio galaxy 3C 75 in the cluster A400 with two close ($16''$) AGNs and a well-developed radio structure, whose study gives valuable information on the origin and state of the AGN system (Owen *et al.* 1985). However, there is the only search limitation from below that is attributable to the finite telescope beam ($\sim 5''$) in the FIRST survey, which by itself makes the search for populated groups on scales $\lesssim 30''$ pointless. Here, the telescope beam sidelobes and the structure of radio sources proper have the greatest effect, which will definitely lead to a dominance of false groups.

Large scales ($>5'-10'$) are being intensively studied, for instance, using the FIRST survey. Thus, for example, we investigate FIRST radio sources in Abell clusters (Abell *et al.* 1989) on these scales. We study rich clusters of galaxies with an anomalously large number of radio sources that we revealed when carrying out a program of identifications of FIRST and NVSS radio sources in Abell clusters (see, e.g., Gubanov and Reshetnikov 1999). Spectroscopic observations of several clusters that could contain an enhanced number of radio galaxies were performed with the 6-m telescope. A preliminary analysis of the observations indicates that such systems of galaxies actually exist. Thus, compact ($<5'$) configurations of radio galaxies (whose number is ≥ 4) were found in the clusters A1234 and A2224 and close pairs were found in a number of clusters (Gubanov and Kopylov (2003) in preparation). This gives hope that clusters will be detected at moderate $z \sim 0.3-0.6$, especially if an evolution effect similar to the optical Butcher−Oemler effect takes place in the radio frequency range. The detection of this effect or the proof of its absence will be of importance in studying the evolution of activity in galactic systems. An attempt can also be made to find forming clusters of galaxies at $z \sim 1-2$ or even farther. Of course, this will require a deep optical (or infrared) identification. Therefore, for the first search, emphasis can be placed on the scales $1'-5'$ ($100-3000$ kpc). In this case, both the range of still comparatively unexplored galaxy groups (a linear size of 130 kpc corresponds to an angular size of $1'$ at $z = 0.1$) and the range of distant clusters ($z = 1-5'-2.8$ Mpc) are covered.

To increase the reliability of the searching results and to ensure the proper statistical estimates in analyzing the detected groups, it should first be made sure that the initial data are complete (and clean). According to Magliocchetti *et al.* (1998), a limiting flux density of 3 mJy is a good choice for providing the completeness of the FIRST catalog. It would be also appropriate to additionally reduce the effect of telescope sidelobes on the identification of source groups. Indeed, although the VLA antenna gives high-quality images, our preliminary analysis indicates that intense radio sources in the FIRST catalog are often surrounded by an excess of faint (at a level $\lesssim 1\%$) objects, the clear remnants of the telescope sidelobes.

This interference can partially be eliminated by removing the objects with the flag W (i.e., with evidence of sidelobes) from the catalog. Running ahead, we note that this procedure actually reduces appreciably the occurrence of false groups and has no significant effect on the statistical inferences. (Although, of course, it can lead to the loss of a small number of real groups because of errors in the automatic classification of such W-objects in the catalog. This is undesirable, but a certain tradeoff between the losses and the volume of work is inevitable.)

We can also select only those groups where the flux density contrast for the sources in a group does not exceed 100. Note that this restriction cannot significantly affect the loss of groups. Indeed, intense radio sources are known to be comparatively rare and lie in the steep segment of the radio luminosity function (RLF) after its kink. Thus, if, for example, the faintest source with a flux density of 3 mJy were detected in a group at $z = 0.2$, then, at a contrast >100, the most intense source would have a radio luminosity (above $10^{25.7}$ W Hz^{-1} at a frequency of 1400 MHz) that falls on the steep segment of the RLF, indicating that its occurrence is relatively rare.

As for the choice of a population, experience in searching for radio sources, primarily in rich clusters and groups of galaxies, shows that three to five moderately intense radio galaxies per system are rarely encountered in a local volume at low z. Therefore, new results can be obtained in this range of populations (but in a much broader z range).

The main scale under study from the range $1'-5'$ was chosen to be $5'$, because many groups of the largest size from the chosen range must contain smaller groups (with a similar density but a lower population), which can spare efforts at the subsequent stages of the search for and analysis of groups with different parameters.

Thus, at the first stage, we will select groups of extragalactic radio sources from the FIRST catalog with flux densities ≥ 3 mJy (with the W-objects removed), a population of no fewer than five objects, and a flux density contrast of no more than 100 located within $5'$ areas.

Such groups with a high surface density of radio sources will be candidates for our subsequent search for physically (gravitationally or genetically) associated systems with active galaxies.

THE METHOD FOR SELECTING GROUPS

To ensure that the initial data were statistically homogeneous and to simplify the subsequent statistical estimates, we imposed constraints on the limits of the sky region in which these estimates would be correct.

For this purpose, we broke down the entire region of the FIRST survey into elementary areas $1° \times 1°$ in right ascension and declination. All incomplete areas (at the survey boundaries) were excluded from the subsequent analysis.

In addition, we constructed the distributions of the remained areas (whose number is 8181) in (mean and maximum) noise level, mean source flux density in the area, and number of sources in it. This information can be easily extracted from the catalog. The areas for a noise level and a mean flux density that clearly corresponded to outliers in the observations or contained the most intense radio sources (naturally, with all of their numerous sidelobes, as can be easily verified) were excluded from the subsequent analysis by using the 3σ criterion. The area being analyzed then decreased to $8096\square°$ (we call the corresponding area the A-region for short).

A standard method of searching for populated systems of astronomical objects is their separation by the surface density contrast. However, in contrast, for example, to the search for clusters or groups of galaxies in the optical wavelength range, where the characteristic densities of objects are high and the main interfering factor is the projection of foreground and background objects onto the system, in our case, the major problems of searching for groups of radio sources are associated with the character of the initial FIRST data.

The fact that extended radio sources are generally represented in the catalog by a set of separate close components indistinguishable in the catalog from single sources and that artificial (i.e., false) objects attributable to the influence of VLA beam sidelobes are present in the catalog severely complicate the search for groups of separate real radio sources.

As was noted above, we first searched for groups of only the FIRST objects whose total flux densities exceeded 3 mJy. In addition to throwing away a significant fraction of the faint FIRST objects attributable to the VLA sidelobes, this ensures the homogeneity of the catalog itself, because objects with a 5σ excess above the noise were selected for the catalog. The noise level, whose typical value is about 0.15 mJy, varies over the survey region but does not exceed 0.6 mJy (which we checked directly by using the FIRST data). The FIRST objects with the flag W were also removed.

As a result, after imposing all of the constraints listed at the end of the preceding section, the total number of the FIRST objects remained in the A-region was 349 329.

The components of structurally complex radio sources are actually responsible for a significant fraction (about a third; see Cress *et al.* 1996) of
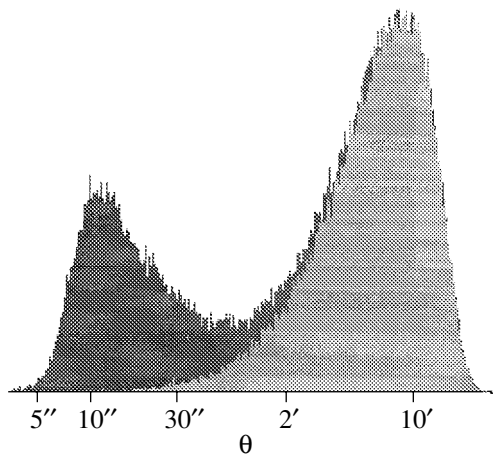
Fig. 1. The distribution of the number of closest pairs (histogram) as a function of the angular distance θ between the components. The light-gray region corresponds to the Poisson approximation for widely separated FIRST objects.

Fig. 2. Parameters of the orientation in the pairs of FIRST objects: θ is the angular distance between the objects; ϕ_1 and ϕ_2 are the angular sizes along the major axis; β_1 and β_2 are the angles of deviation of the objects' orientation from the line of their connection; and β is the alignment parameter, the mean of β_1 and β_2.

the FIRST objects. Since these components are generally close to each other, the number of false groups whose population is attributable to the separate components of these extended sources rather than to real sources significantly increases. This appreciably reduces the efficiency of the search for real groups. Thus, if the multicomponent structure is disregarded, then the number of groups with a population of 5 or more and with angular sizes of no larger than $5'$ is 2318. In this case, as the subsequent studies indicate, only a few percent of these groups could be formed by five or more real radio sources rather than the components of extended sources.

To reduce the influence of such extended sources, their separate components should be combined into a single source. The quality of this combination is determined by the number of the remaining uncombined components of real extended sources and by the number of single sources lost through an incorrect combination. The former determines the number of remaining false groups, while the latter determines the number of lost groups consisting of real objects.

The main parameter of the combination of two FIRST objects into a single source is the angular distance between them. Figure 1 shows the distribution of the number of closest pairs of objects (histogram) selected from the FIRST catalog in the A-region by the distance between them. This distribution is bimodal; for wide (with distances $\gtrsim 1'.5$) pairs, it corresponds to the distance distribution for the closest pairs in the "random" catalog (i.e., in the catalog compiled from fictitious objects uniformly thrown into the FIRST region). Clearly, the sharp peak at small angular distances is determined by the real extended
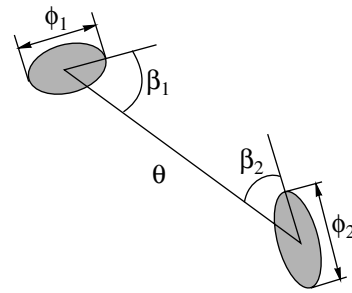
sources represented in the catalog by separate components. Thus, two populations of FIRST objects are naturally distinguished: multicomponent and single sources.

Trial experiments based on the simple technique of combining objects from the original catalog within $15''$, $30''$, and $60''$ radii show that $30''$ is better suited to the selection of groups on scales $2'-5'$. This value is close to the values for the combination that were used by Magliocchetti et al. (1998) when studying the correlation function of FIRST radio sources. However, this technique is far from being optimal for the problem in question.

It is well known that most of the extended radio sources are significantly elongated along some axis, the sizes of their components correlate with the distances between them, and the components are oriented along the source's major axis. There is also a certain correlation of these quantities with the flux density of the components. The FIRST catalog contains information on the structure of its objects, which we can attempt to use to optimize the procedure of combining the components of extended radio sources.

Figure 2 shows some of the geometric parameters for pairs of objects that can be easily extracted from the available catalog. We analyzed the distributions in various pairs of parameters (including the angular distance, the angular size, the alignment parameter, and the flux density) for the closest FIRST objects. Figures 3a–3c show the distributions of the closest pairs in the three pairs of parameters for which the separation of the two populations of FIRST objects is clearest.

Two classes of objects can be identified in all three distributions. The distributions have a simple but asymmetric shape. Since the two populations of objects overlap, they cannot be unambiguously separated by the corresponding parameters. We also clearly see that the dividing line between the classes
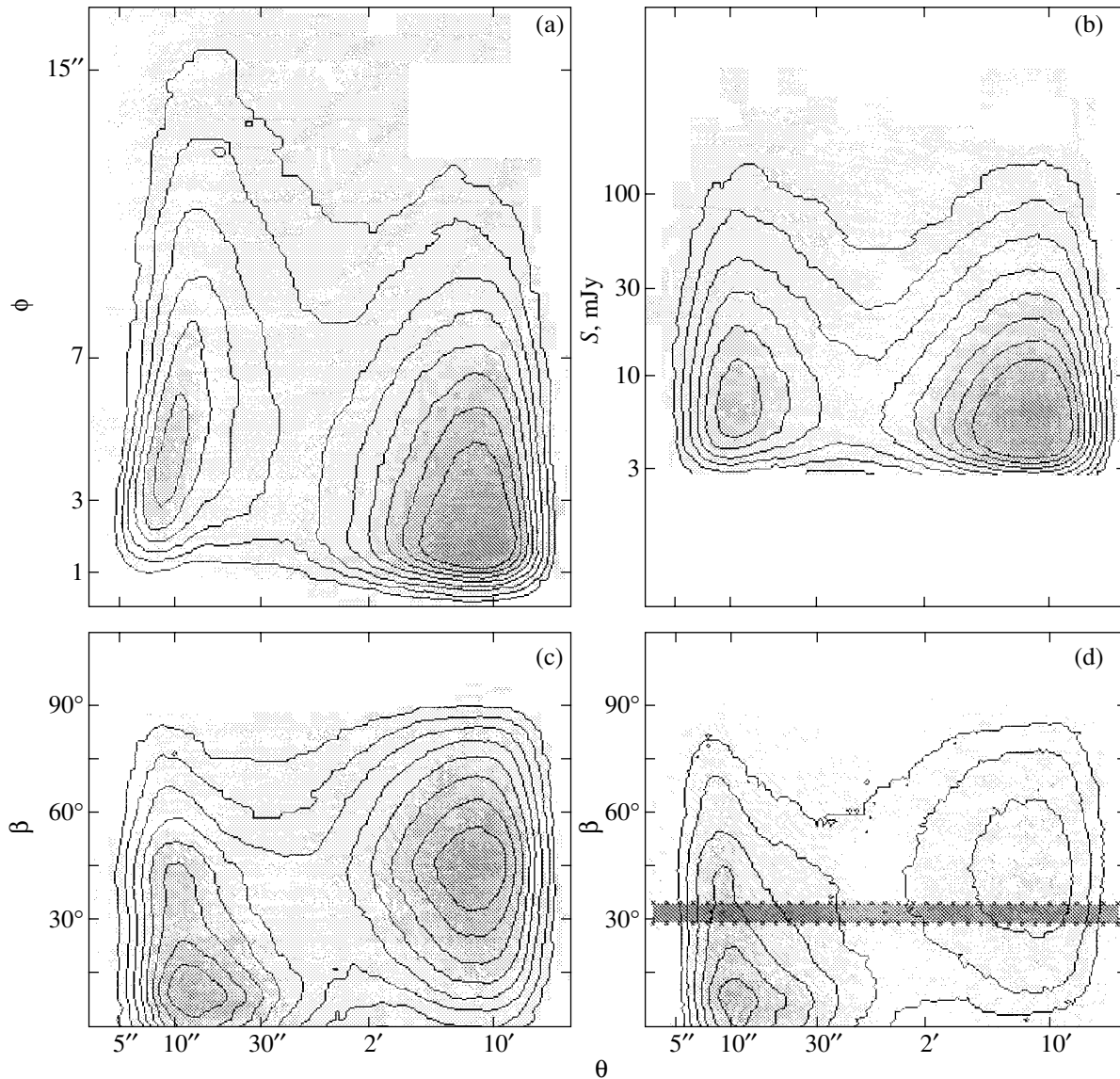
Fig. 3. The distributions of the closest FIRST objects: (a) in mean (in the pair) angular size ϕ, (b) mean flux density S, and (c) alignment parameter β as a function of the angular distance θ between the components; (d) the same as Fig. 3c for the pairs where one FIRST object has a size in the range $3''-6''.5$, while the size of the other object is $>6''.5$. The strip of objects for which the distribution was fitted (see Fig. 4) is highlighted.

cannot be described by a constant in a particular parameter. The figures show that an attempt can be made to separate the populations of FIRST objects by fitting their distributions in those segments where the corresponding population dominates.

For the corresponding fit, we constructed histograms as a function of the mutual angular distance in sections in the second pair parameter. The parts of the histograms between the peaks were fitted by the sum of two functions,

$$f(r) = \frac{A}{(r - r_0)^2 - a^2},$$

where r is the logarithm of the angular distance between the components and the free parameters r_0, A, and a were determined iteratively.

Next, using our fits, we can estimate the number of objects that belong to each class in order to choose criteria for separating objects into classes with admissible classification errors; i.e., we can estimate the number of the uncombined components of real multicomponent sources and the number of unjustifiably combined single sources (their loss) from the tails of the corresponding distributions, as illustrated by Fig. 4.

The corresponding estimates for the number of

losses when separating the two populations of FIRST objects in the flux density–distance plane indicate that using the flux density improves the separation quality only slightly compared to the separation by using the criterion of the distance between the pair components alone. Therefore, in combining FIRST objects, we used only the angular distance between the objects, their mutual orientation (alignment), and the angular sizes along the object's major axis (see Fig. 2).

As a result, the procedure for combining FIRST objects reduced to the following. Two FIRST objects were assumed to belong to the same source if they were separated by a distance smaller than the combination radius, which can be determined from the sizes of the objects and the angle of their mutual orientation.

Given the ultimate goal, the preferred values of the combination radius were such that the number of lost groups was not significant. Thus, we admitted an additional "contamination" of the selection with the false groups attributable to the remaining relatively close uncombined components of real extended sources. The ratio of the number of sources after the combination to the number of uncombined components of extended sources was chosen to be $1:3$. Nevertheless, subsequent analysis showed a significant improvement in the efficiency of this selection compared not only to the selection procedure without combination but also to the method of combination within a $30''$ radius. Thus, the selection procedures without combination and with combination within a $30''$ radius give, respectively, 2318 and 168 groups with a population $\geqslant 5$, while our technique reveals 125 groups. Only one or two actually populated groups from the 168 were not among the 125.

To independently check the estimated fraction of missed groups when using our combination algorithm, we simulated its operation for artificial groups. To this end, we took isolated FIRST objects (here, these are objects without neighbors within a $5'$ circle) and formed random groups from them with a population of five objects each. For this purpose, we randomly chose a FIRST object with its parameters (size, orientation, flux density) and uniformly threw it on a $5'$ circle at random in position. For such groups, we then performed the combination procedure and determined their relative decrease after applying the combination algorithm.

Since extended radio sources account for an appreciable fraction of the sources (the combination procedure described above gives an estimate of 17.2%) and since the overwhelming majority of them are double, we also chose real two-component sources from the catalog for our simulations. Subsequently, we formed groups of four or three single
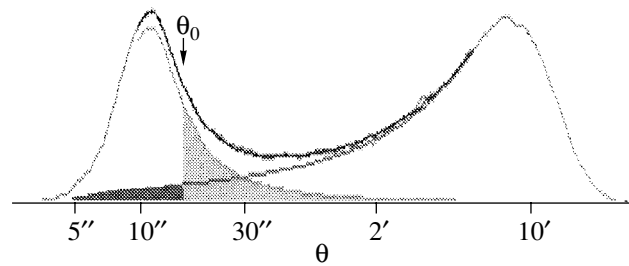

Fig. 4. An illustration of choosing the parameters of the combination of FIRST objects into a single multicomponent source and estimating the number of sources lost after the combination. One of the sections for the distribution of pairs of closest objects in the angular size–alignment plane and its fit are shown; θ_0 is the chosen boundary value of the combination parameter. The shaded region in the left tail of the fitted distribution of actually distinct sources corresponds to the estimated number of sources lost after the combination.

sources in a similar way and added one or two double sources, respectively, to the group. Since the faction of triple and other multicomponent sources was relatively small, we restricted our simulations of groups to this approximation.

The losses are estimated to be (9.9 ± 3.3), (11.3 ± 2.6), and $(12.9 \pm 2.9)\%$ for the three types of groups, respectively. These estimates closely match the estimates, 10.6%, obtained by fitting the distributions for the populations of single and extended sources when analyzing the pairs of closest FIRST objects.

After defining the combination criterion, we developed a program that automatically combined FIRST objects and compiled a new catalog of sources. The parameters for such combined sources were chosen as follows. The flux density was the sum of the flux densities for the combined sources and the coordinates corresponded to the weighted mean (in flux) position.

By applying this combination algorithm, we compiled a new catalog of 298 111 sources from 349 329 FIRST objects (with the sidelobe and W-objects removed) with flux densities $\geqslant 3$ mJy located in the survey A-region. The density of these sources is 36.8 per square degree. Thus, the mean number of sources within a $5'$ circle is 0.2007. The corresponding density contrast for a group with a population of five sources is approximately equal to 25.

The algorithm of searching for groups itself was specially chosen to be as simple as possible. We constructed a circle of a given diameter around each source and formed a group from the sources that fell within the circle. All of the group overlaps were taken into account in such a way that the final list included only the most populated groups.

Recall that when selecting groups, we additionally constrained the flux density contrast for their objects. We selected only those groups where the contrast (the ratio of the flux densities for the brightest and faintest sources) did not exceed 100. This constraint has no significant effect on the selection of real groups but eliminates an appreciable fraction of definitely false groups (which was checked directly).

RESULTS

We applied the procedure of searching for groups located within a 5′ circle with a population of five or more sources (there were 125 of them) to the new catalog obtained after the combination.

For all of these selected candidates for groups, we constructed the images of $10′ \times 10′$ optical fields centered on the group and with superimposed radio contours by using the Digital Sky Survey, DSS-II (R band), and FIRST digital images. We used a slightly modified service of the galaxy database (http://www.astro.spbu.ru/CLUSTERS/) to request the corresponding data and the MIDAS и PDL (Perl Data Language) software packages to construct the corresponding images.

The constructed field maps and images were examined to visually classify the groups of radio sources by types. To this end, we also used the NRAO VLA Sky Survey (NVSS) carried out at the same frequency but with a different antenna configuration. In general, the arrangement of the telescope sidelobes for NVSS differs from that for FIRST and this survey is more sensitive to low-surface-brightness features. In several cases, this allowed the actual situation with the sources to be refined. We also used information provided by the NED (NASA Extragalactic Database) service.

First, we broke down the groups into three main types. The tentatively false groups whose population was actually less than five were assigned to the first two types. They were selected because either a radio source (or radio sources) with a complex structure, where the combination procedure did not work properly, fell within the corresponding circle or there were objects within the circle that were actually attributable to the VLA sidelobes. The groups for which there was no reason to believe that they were less populated constituted the third type.

The type for which the main selection criterion was a multicomponent structure of one (or more) radio sources that fell within the corresponding circle proved to be most numerous (79 groups, 63.2%). This was the result of the "mildness" of the chosen criterion for the combination of FIRST objects, because we sought to avoid excessive losses of real groups in this procedure.

Fifteen groups (12.0%) proved to be so dense because some of the FIRST objects were sidelobe ones. Note that removing the objects with the flag W from the catalog, increasing the limiting flux density to 3 mJy, and eliminating the noise areas significantly (severalfold) decreases the fraction of this type of groups in the selection procedure. Thus, for example, if no W-objects are removed, then the number of groups increases from 125 to 158; i.e., 33 more candidates for groups are added. Only one of them is most likely a real group with a population of five, in which one object was erroneously flagged as W. We reached this conclusion based on the reliable identification of this source with a galaxy and on its presence in the NVSS catalog. Hence, 32 more groups are added to the 15 false groups; i.e., their number triples. In general (but not always), the groups of this type were near intense radio sources.

Finally, 31 groups (24.8%) listed in the table were assigned to the third type. Of course, a group cannot always be unambiguously classified. Therefore, this classification has a certain element of subjectivity. Still, the number of arguable cases is relatively small (2−3).

Both real groups of radio sources located in physically small volumes and sources projected onto the line of sight that are not associated with each other at all or those in projected less populated groups can be among the selected 31 groups (distinguished by a high surface density of objects). These groups were split into two parts. One part (list B) clearly shows evidence of a chance projection of some group objects. A projection can be determined with confidence by the presence or absence of optical counterparts on our maps. However, the DSS sensitivity limit is not enough to identify most of the objects in the groups found. The other part (list A) shows no clear evidence of a projection. This includes mostly groups without identifications or with a small number of source identifications with optically faint objects. While list B so far contains 13 groups, list A includes 18 groups. Examples of the images of groups from the two lists are shown in Fig. 5.

Apart from the groups of lists A and B, we added four more groups to the table that could consist of five separate sources (list C). These groups were missed by the selection procedure described above, but we found them using other procedures. One of these groups (5GR0126.05) contains a FIRST object with flag W as a source. That is why it was not assigned to the populated groups by the selection procedure in question. However, this object has an optical counterpart, indicating that it is real as a radio source. In addition, the NVSS data we analyzed also argue for this interpretation. The remaining groups (5GR0127.05, 5GR0128.05, and 5GR0129.05) were

Lists of groups of FIRST radio sources with a population of at least five within a 5′ circle

Group	RA(2000)	DEC(2000)	IAU designation*
List A. Groups with no clear evidence of projection			
5GR0007.06	09^{h}13^{m}17^{s}29	+33°47′29″9	CGR_F J091317.2+334729
5GR0010.06	09 54 58.79	+47 17 09.6	CGR_F J095458.7+471709
5GR0016.06	15 31 00.81	−05 08 13.7	CGR_F J153100.8−050813
5GR0048.05	10 13 59.87	+12 28 10.9	CGR_F J101359.8+122810
5GR0052.05	10 48 40.57	+04 27 06.4	CGR_F J104840.5+042706
5GR0056.05	11 21 12.73	+57 54 30.8	CGR_F J112112.7+575430
5GR0060.05	12 13 19.63	+51 56 09.0	CGR_F J121319.6+515609
5GR0067.05	13 03 33.16	+41 02 54.2	CGR_F J130333.1+410254
5GR0073.05	13 25 30.65	−01 47 30.9	CGR_F J132530.6−014730
5GR0082.05	14 01 58.58	+23 39 58.9	CGR_F J140158.5+233958
5GR0084.05	14 08 23.32	+18 22 35.0	CGR_F J140823.3+182235
5GR0089.05	14 20 30.82	+12 13 34.7	CGR_F J142030.8+121334
5GR0103.05	15 16 36.22	+22 20 54.9	CGR_F J151636.2+222054
5GR0105.05	15 19 21.43	+18 36 08.3	CGR_F J151921.4+183608
5GR0114.05	16 31 29.44	+32 23 49.4	CGR_F J163129.4+322349
5GR0116.05	16 39 17.73	+11 28 46.7	CGR_F J163917.7+112846
5GR0117.05	16 52 04.15	+34 01 22.6	CGR_F J165204.1+340122
5GR0125.05	23 14 41.56	+00 39 43.3	CGR_F J231441.5+003943
List B. Groups with evidence of projection			
5GR0008.06	09 46 17.52	+47 37 20.9	CGR_F J094617.5+473720
5GR0011.06	10 32 19.90	+31 53 14.5	CGR_F J103219.9+315314
5GR0014.06	13 42 55.76	+47 48 04.8	CGR_F J134255.7+474804
5GR0022.05	07 40 49.60	+56 16 11.7	CGR_F J074049.6+561611
5GR0033.05	08 47 54.14	+12 40 47.5	CGR_F J084754.1+124047
5GR0053.05	10 51 40.58	+10 59 47.6	CGR_F J105140.5+105947
5GR0080.05	13 58 25.80	+33 18 05.5	CGR_F J135825.8+331805
5GR0092.05	14 35 21.94	+11 50 52.7	CGR_F J143521.9+115052
5GR0095.05	14 43 09.50	+31 50 33.2	CGR_F J144309.5+315033
5GR0096.05	14 47 01.33	+44 56 28.5	CGR_F J144701.3+445628
5GR0101.05	15 05 23.43	+47 06 25.6	CGR_F J150523.4+470625
5GR0110.05	16 26 18.60	+37 13 22.1	CGR_F J162618.6+371322
5GR0113.05	16 29 50.09	+45 12 43.2	CGR_F J162950.0+451243
List C. Groups missed by selection procedure			
5GR0126.05	07 42 25.47	+44 46 19.5	CGR_F J074225.4+444619
5GR0127.05	08 14 29.75	+21 48 51.8	CGR_F J081429.7+214851
5GR0128.05	11 23 21.78	+30 43 38.3	CGR_F J112321.7+304338
5GR0129.05	16 40 09.62	+19 01 58.4	CGR_F J164009.6+190158

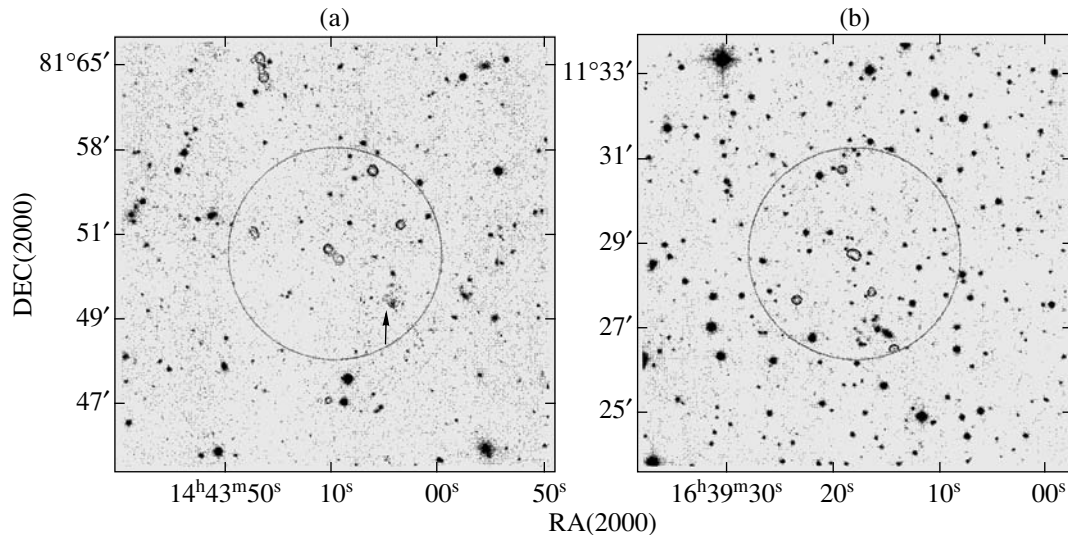* IAU designation—Compact Groups of Radio Sources in FIRST, epoch J2000—CGR_F J.

Fig. 5. Examples of groups: (a) from list B (5GR0095.05) and (b) from list A (5GR0116.05). The radio source (marked by the arrow) in the southwestern part of the circle for 5GR0095.05 is identified with a bright galaxy in a distant cluster, whereas the other four radio sources have no optical counterparts.

missed due to the combination of, probably, separate but close sources. In one case (5GR0127.05), the FIRST objects combined by the procedure have clear optical counterparts; i.e., they are actually separate radio galaxies. In another case (5GR0128.05), one of the combined objects has an optical counterpart, while in the group 5GR0129.05, the extended structure of the FIRST objects combined by the procedure also allows their interpretation as separate sources. Note that the number of these missed groups is in close agreement with our estimated number of groups lost after the combination (about 10%).

Thus, we found 31 groups with a population of no fewer than five and angular sizes of no larger than $5'$ in the $8096\square°$ A-region of the FIRST survey among 298 111 (with combination). Given that groups could be missed by our searching procedure, their number in this region must be about 34.

Of course, there are also close groups of sources of certain interest among the less populated groups. A list of all of the groups we found, radio-optical images of the corresponding field, and detailed comments can be found at http://www.astro.spbu.ru/ag/ FSTgr/. The properties of these groups will be considered in more detail in a separate paper.

DISCUSSION

To estimate how significant the number of groups found is, we simulated a uniform random distribution of radio sources in the sky and determined the number of groups for this distribution. To this end, we compiled a catalog that contained the expected number of objects with random coordinates in the A-region. The catalog compiled in this way was subjected to the procedure of searching for groups (about a hundred trials were performed).

In this case, the mean number of groups with a population of five or more was found to be 10 ± 2.5, which is determined by a chance projection of objects not associated with each other. Groups with a population larger than 5 have virtually no chance of appearing in the random catalog. This number is most likely slightly overestimated, because it does not include the contribution from the undetected W-objects in the FIRST catalog. However, even without an allowance for this possibility, there is a threefold excess of detected groups above their number found in the simulation.

Thus, there is a significant excess of groups of FIRST radio sources with a high (a factor of 25 higher than the mean) surface density compared to such groups for the Poisson distribution (of course, only for the special case of $5'$ circles considered). This excess suggests the existence of locations of preferential formation of active galaxies and that using groups of radio sources to search for the corresponding systems, groups and clusters of galaxies, holds much promise.

It should be noted, however, that the total number of such structures is comparatively small. We found only 31 (34) such systems among almost three hundred thousand sources. In addition, our identifications of the radio sources with optical DSS objects revealed no groups where the existence of more than three sources in one (physically associated) system of galaxies could be assumed with confidence. What can this mean?
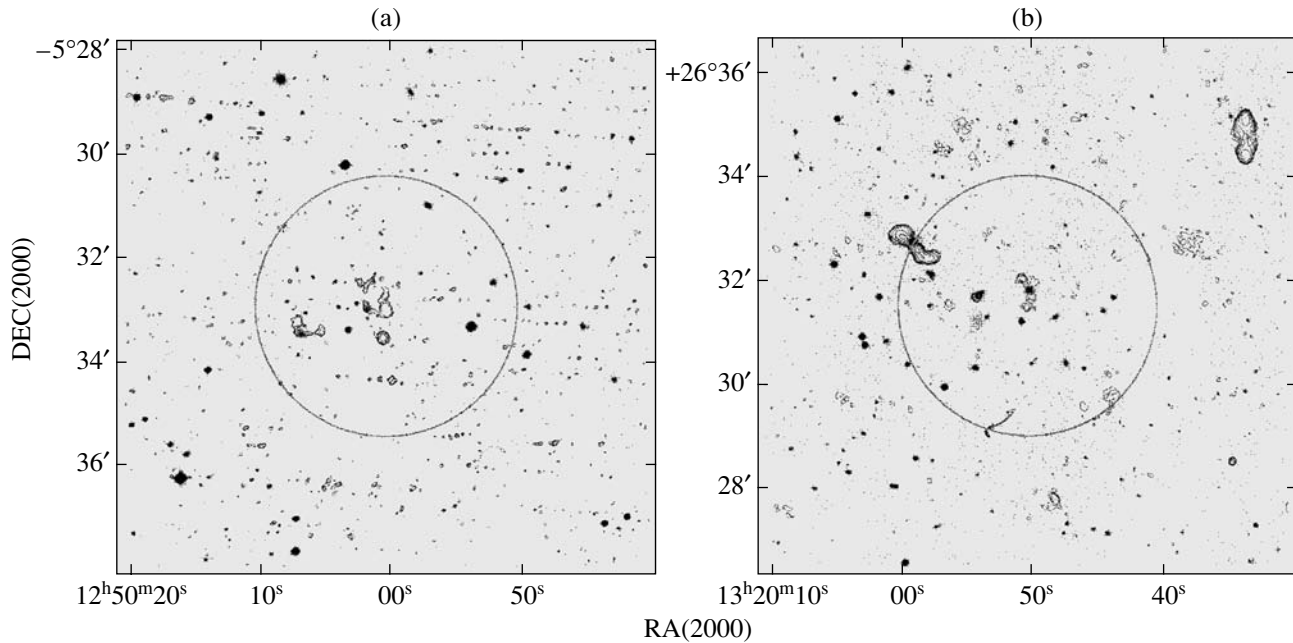
Fig. 6. (a) A poor cluster at $z = 0.19$ with the two brightest galaxies that are WAT radio sources (group 5GR0063.05). (b) The group 5GR0070.05, in which three radio galaxies, probably, belong to a poor cluster with detected X-ray radiation.

Our search for dense groups corresponds to a kind of space probing in a cone with an opening angle of $5'$ for the presence of active galaxies and galactic systems containing them that can lie at almost any cosmological distance. It is clear, however, that during this probing, different characteristic radio luminosities of objects (because the flux density is limited) and different linear sizes of galaxy systems (because the cone opening angle is finite) correspond to different ranges of redshifts z.

We can arbitrarily separate three characteristic redshift ranges. Up to $z = 0.1$, where a radio intensity of 10^{23} W Hz^{-1} corresponds to the 3-mJy limit, starburst galaxies account for a significant fraction of the FIRST sources, because here they dominate in the local radio luminosity function (RLF) (see, e.g., Sadler *et al.* 2002). Their number decreases sharply with increasing intensity in accordance with the steep RLF slope. At this boundary, the linear size that corresponds to 5 arcmin is about 650 kpc, the size of galaxy groups. Thus, we found no close (with scales of several hundred kpc) groups of galaxies densely populated by regions with intense star formation (an/or weak AGNs) in the local neighborhood.

In the range of redshifts z from 0.1 to 1, moderate-intensity radio galaxies constitute the main population of the FIRST sources up to the kink in the RLF (about $10^{25.7}$ W Hz^{-1}). Here, the 3-mJy limit at the far boundary corresponds to an intensity of $10^{25.3}$ W Hz^{-1} and the linear size subtended by the

cone changes from 650 kpc to 2.8 Mpc, which corresponds to the scales of the central regions in rich clusters of galaxies. Our search revealed no reliable case where a cluster could contain four or five radio galaxies within the radius bounded by the cone. However, groups pointing to systems with two or three such radio galaxies are commonly encountered. Figure 6, with the two clusters revealed by our search, serves as an illustration of such systems. More populated clusters but with a larger angular separation between radio sources (as, e.g., the cluster A1095 at $z = 0.21$, which may contain seven radio galaxies) are also encountered more rarely.

The properties of the galaxy systems detected in our search for groups will be analyzed in a separate paper. Here, we note that for the DSS sensitivity level (about 21^m in R) and typical absolute magnitudes of bright cluster galaxies (which radio galaxies mostly become) from $M^* \approx -22.5$ (where M^* is the characteristic kink in the optical luminosity function) to $M_{BCG} \approx -24.5$ (the absolute magnitude of the brightest galaxy), the typical redshift z lies within the range $0.5-0.7$. Only cD-galaxies (which are brighter by another $0.5-1$ magnitude) can be detected at $z \approx 1$ or slightly farther. Therefore, the possibility of detecting comparatively compact systems populated by radio galaxies cannot yet be ruled out.

Finally, comparatively intense radio sources are the main population among the FIRST objects in the z range from 1 to 6. The radio luminosity limit at the far boundary for a 3-mJy flux is about 10^{27} W Hz^{-1}.

In this case, since the spatial density of intense radio sources at high z increases by two or three orders of magnitude because of strong cosmological evolution (see, e.g., Willott *et al.* 2001), this population of FIRST objects can still be numerous, despite the changes in RLF slope. The characteristic linear sizes here change only slightly and, as previously, the diameter of our cone spans the scales that correspond to the central regions of rich clusters or protoclusters.

The detection of a few densely populated groups of radio sources implies that, on the probed scales, we failed to find any clear evidence for the existence of "nests" of radio sources containing $>4-5$ objects in one system. Thus, we found no close (with scales of several hundred kpc) local groups of galaxies densely populated by regions of intense star formation. Nor did we find any manifestations of activity accompanied by a comparatively intense radio emission simultaneously in several galactic nuclei in the central regions of galaxy clusters at moderate z and protoclusters at cosmological z. In principle, since an excess of populated groups, compared to their number for the Poisson distribution, was detected, this can be used to impose constraints both on the conditions for the emergence and evolution of activity in galaxies and on the formation of galaxy systems and the possible relationship between these two processes.

In conclusion, we point out the three most promising possibilities of using groups of radio sources.

First, it is quite clear that they can be used to search for rich systems of galaxies in a wide redshift range and this detection method finds increasingly wide applications (see, e.g., Croft *et al.* 2001; Buttery *et al.* 2001). This study shows the possibilities of making such a search for the FIRST survey. Our simple estimates using numerical simulations indicate that the probability of finding a radio source of a group in a system of galaxies in a pair with another source (or in a triple with other two sources) can be tens (hundreds) of times higher than this probability for an arbitrary FIRST source.

The appearance of SDSS (Sloan Digital Sky Survey; see, e.g., Loveday 2002) optical data significantly enhances the potential of this search. Recall that the two surveys are carried out in the same sky region with an area of about 10 000 square degrees. The simultaneous use of electronic data from these surveys will also make it possible to realize and develop the models in terms of the concepts of the Virtual Observatory, which is a pressing problem of modern astronomy.

Second, the excess of populated groups that we found admits the existence of galaxy systems containing not only two or three comparatively intense radio source each (which we clearly see in several cases) but also four or five such radio galaxies. As

long as the existence of an appreciable number of systems populated by radio galaxies will be proven, this will raise the question of the simultaneous emergence of activity in the nuclei of several cluster galaxies, because the occurrence of even one radio source is a rare event. Recall that, on average, one radio source with an intensity of about 10^{26} W Hz^{-1} is encountered in a volume of about 10^6 Mpc^3. Therefore, the detection of even one system in a limited z range with five intense sources can come into conflict with the number of (most likely rich and distant) clusters if the occurrence of radio sources is independent and random. Another possibility can be an appreciable increase in the characteristic lifetime for radio galaxies in some systems. In any case, this will argue for a common external factor responsible for the simultaneous emergence and/or maintenance of activity in the nuclei of several galaxies.

Third, we cannot rule out the possibility that the first radio galaxies in different directions emerged in some time interval and, possibly, not in isolation but in close neighborhood with other radio galaxies. Indeed, current views of the hierarchical formation of galaxies suggest this possibility. Since our method of selecting groups is based on a multifold excess of the surface density of objects in a given region above its typical value, the probability of detecting such groups in the direction where the formation epoch of the first radio galaxies (or systems of galaxies with several sources) was most distant in z can be higher than its mean value. As long as many FIRST radio sources can sound the Universe almost all the way through, it is possible to effectively detect radio galaxies at cosmological distances and galactic systems containing them by using groups.

What objects can be found in groups actually depends both on the pattern of the cosmological evolution of radio galaxies and galaxy clusters and the choice of searching parameters. Even when the selected groups are projections of independent sources, there is a chance of detecting interesting objects in them.

Of course, other things being equal, the chance of finding a distant object in such a group, as, for example, that selected here, can differ little from the probability of its detection for five radio sources taken at random. In this case, however, the efficiency of the following observations significantly increases, because simultaneously all of the group objects can fall within the telescope field of view. This saves the observational and reduction time.

CONCLUSIONS

(1) We developed a technique of searching for groups of extragalactic radio sources in the FIRST

catalog with a high surface density of the objects. The use of such groups to detect systems populated by radio galaxies (groups and distant clusters of galaxies) was shown to be promising.

(2) We detected 31 groups of radio sources with angular sizes $<5'$ that contain at least five sources with flux densities ≥ 3 mJy among the 125 candidates for groups found. The population of the remaining 94 groups proved to be lower. We showed that, for the searching method used, the number of missed groups is no more than 10%.

(3) We detected at least a threefold excess of the groups found above their expected number for a random Poisson distribution in the sky region under study. This argues for the existence of not only locations where activity preferentially appears (clusters and groups of galaxies) but, possibly, also for the simultaneous emergence of this activity in the galaxies of these systems.

ACKNOWLEDGMENTS

REFERENCES

1. G. O. Abell, H. G. Corwin, and R. P. Olowin, Astrophys. J., Suppl. Ser. **70**, 1 (1989).
2. R. H. Becker, R. L. White, and D. J. Helfand, Astrophys. J. **450**, 559 (1995).
3. H. J. Buttery, G. Cotter, R. W. Hunsted, and E. M. Sadler, astro-ph/0112129 (2001).
4. C. M. Cress, D. J. Helfand, R. H. Backer, *et al.*, Astrophys. J. **473**, 7 (1996).
5. S. Croft, S. Rawlings, G. J. Hill, *et al.*, astro-ph/0110119 (2001).
6. A. G. Gubanov and V. P. Reshetnikov, Pis'ma Astron. Zh. **25**, 447 (1999) [Astron. Lett. **25**, 380 (1999)].
7. B. V. Komberg, A. V. Kravtsov, and V. N. Lukash, Mon. Not. R. Astron. Soc. **282**, 713 (1996).
8. J. D. Kurk, B. P. Venemans, H. J. A. Rottgering, *et al.*, astro-ph/0110131 (2001).
9. J. Loveday, astro-ph/0207189 (2002).
10. M. Magliocchetti, S. J. Maddox, O. Lahav, and J. V. Wall, Mon. Not. R. Astron. Soc. **300**, 257 (1998).
11. T. K. Menon, *Galaxy Interactions at Low and High Redshift, Proceeding of IAU Symposium No. 186, Kyoto, Japan, 1997*, Ed. by J. E. Barnes and D. B. Sanders (1999), p. 383.
12. F. N. Owen, C. P. O'Dea, M. Inoue, and J. A. Eilek, Astrophys. J. Lett. **294**, L85 (1985).
13. S. Rawlings, Astron. Soc. Pac. Conf. Ser. **199**, 2000, astro-ph/0008067.
14. E. M. Sadler, C. A. Jackson, R. D. Cannon, *et al.*, Mon. Not. R. Astron. Soc. **329**, 227 (2002).
15. H. M. Tovmassian, *Cluster Mergers and Their Connection to Radio Sources, 24th Meeting of the IAU, Joint Discussion 10, Manchester, England, 2000* (2000).
16. R. L. White, R. H. Becker, D. J. Helfand, and M. D. Gregg, Astrophys. J. **475**, 479 (1997).
17. C. J. Willott, S. Rawlings, K. M. Blundell, *et al.*, Mon. Not. R. Astron. Soc. **324**, 1 (2001).

Translated by V. Astakhov

Observations of the Gamma-Ray Flux
from the X-ray Source Cyg X-3 in 1994−1995

Yu. I. Neshpor, O. R. Kalekin, A. A. Stepanian[*], V. P. Fomin, and N. N. Chalenko

Crimean Astrophysical Observatory, p/o Nauchnyi, Crimea, 334413 Ukraine
Received May 29, 2002

Abstract—We present the observations of Cygnus X-3 carried out with the GT-48 gamma-ray telescope at the Crimean Astrophysical Observatory in 1994−1995. The mean gamma-ray flux at energy $E > 10^{12}$ eV is shown to be approximately equal to 1.3×10^{-11} cm^{-2} s^{-1}. The flux in 1994 was much lower than that in 1995, being $(6.2 \pm 2.6) \times 10^{-12}$ cm^{-2} s^{-1}; i.e., it was statistically insignificant. The flux in 1995 was $(2.7 \pm 0.7) \times 10^{-11}$ cm^{-2} s^{-1}. Thus, the very high energy gamma-ray emission from Cyg X-3 is variable. These measurement results can be used to obtain upper limits on the flux from Cyg X-3 in 1994−1995. © 2003 MAIK "Nauka/Interperiodica".

Key words: *X-ray and gamma-ray sources.*

INTRODUCTION

Cyg X-3 has been known as an X-ray source since 1966 (Giacconi *et al.* 1967). It attracted the attention of many researchers after September 2, 1972, when Canadian radio astronomers detected an intense outburst followed by a second, even more intense radio outburst two weeks later (Gregory *et al.* 1972). After the detection of its outbursts, Cyg X-3 was extensively observed over the entire electromagnetic spectrum: from radio waves to very high energy (VHE) gamma rays. A detailed overview of the results can be found in Vladimirsky *et al.* (1985). Below, we briefly consider these and more recent results of observations of Cyg X-3.

The observations of Cyg X-3 in the VHE range, $>10^{12}$ eV, by recording Cherenkov radiation at the Crimean Astrophysical Observatory (CrAO) were started on September 5, 1972, with Cherenkov detectors and continued until 1980. The VHE gamma-ray flux was detected during the observations of September 1972 (Vladimirsky *et al.* 1972, 1973, 1974). In the 1980s, the presence of a VHE gamma-ray flux from Cyg X-3 was confirmed at many observatories that carried out observations in this spectral range (Danaher *et al.* 1981; Lamb *et al.* 1982; Dawthwaite *et al.* 1983; Mukanov 1981; Cawley *et al.* 1985; Chadwick *et al.* 1982).

X-ray studies of the object revealed periodic variations (Canizares *et al.* 1973), suggesting that Cyg X-3 is a close binary system (with a period of 4.8 h). A

(4.8-h) periodicity was also detected in the infrared (Becklin *et al.* 1973). The same period was also found in the VHE gamma-ray emission (Stepanian *et al.* 1975), which allowed the detected object to be confidently identified with the X-ray source Cyg X-3. The following ephemeris was obtained from eight-year-long VHE gamma-ray observations: $\mathrm{JD}_0 = 2441550.542$, $T_0 = 0.199680$ day, and rate of change of the period $\dot{T}_0 = 3 \times 10^{-9}$ s s^{-1} (Neshpor and Zyskin 1982). Long-term observations (1972−1980) of Cyg X-3 revealed long-period variations in the intensity of the VHE gamma-ray (Neshpor and Zyskin 1988) and X-ray (3−12 keV) emissions with a period of (327.9 ± 2.0) days (Neshpor and Zyskin 1986). In recent years, the conclusion that the object is a binary system has received serious confirmation: the velocity of the gas outflow from one of the binary components and the periodic variation in the helium line wavelength have been determined (Van Kerkwijk *et al.* 1992).

The studies performed have established that a distinctive feature of Cyg X-3 is the high intensity of its radiation. Its total X-ray and gamma-ray luminosity is $(10^{38}-10^{37})$ erg s^{-1} or even slightly higher (Vladimirsky *et al.* 1985). For comparison, the total intensity of the electromagnetic radiation from the Crab Nebula, including the pulsar radiation, is $S = 5 \times 10^{37}$ erg s^{-1}. At the same time, the electromagnetic spectra of the Crab Nebula and the X-ray source Cyg X-3 are similar (Stepanian 1978).

Contradictory data were obtained on the high-energy emission. Gal'per *et al.* (1973) and Lamb

---

[*]E-mail: `arnold@crao.crimea.ua`

et al. (1977) obtained received statistically significant fluxes from Cyg X-3. The COS-B satellite observations in 1975−1980 did not reveal any gamma-ray flux (Hermsen 1983). The EGRET measurements on the CGRO satellite (Thompson *et al.* 1995) showed the presence of a flux from a region close to Cyg X-3. However, no periodic signal was detected.

In the late 1980s and early 1990s, a search for VHE gamma-ray sources was initiated with second-generation ground-based telescopes. These telescopes imaged Cherenkov flashes and were more sensitive to gamma-ray flux. However, virtually no positive results on the VHE gamma-ray flux from Cyg X-3 have been published. Analysis of the observations carried out in 1988−1990 at the Smithsonian Observatory with a second-generation telescope (O'Flaherty *et al.* 1992) allowed only an upper limit on the flux to be determined, 3.5×10^{-12} cm^{-2} s^{-1} for photon energies above 300 GeV. At the same time, according to the observations at the CrAO in 1989, Cyg X-3 was a GT-48 gamma-ray telescope VHE gamma-ray source (Vladimirsky *et al.* 1991).

Thus, the results of the observations of gamma rays from the X-ray source Cyg X-3 in the late 1980s are contradictory, probably due to the variability of the source. Therefore, further observations of the VHE gamma-ray flux from this object seem appropriate. In 1994−1995, we carried out observations with the GT-48 gamma-ray telescope at the CrAO.

INSTRUMENTATION FOR THE GAMMA-RAY OBSERVATIONS

The observations of Cyg X-3 in 1994−1995 were carried out with a second-generation gamma-ray telescope. This facility was described in detail by Vladimirsky *et al.* (1994). Its brief description is given below.

The gamma-ray telescope with a multichannel GT-48 camera (a second-generation telescope) began its operation in 1989. The facility consists of two altazimuth mountings (sections), northern and southern, spaced 20 m apart at an altitude of 600 m above the sea level. Six telescopes were mounted on each section in parallel. The optics of each telescope consists of four 1.2-m mirrors with a common focus. The mirrors of the three telescopes had a focal length of 5 m. Cameras composed of 37 photomultipliers were placed in their focal plane. They image Cherenkov flashes in the visible range (300−600 nm). A conical optical fiber was placed in front of each photomultiplier. The mean diameter of the fiber entrance window is $0°.4$. The field of view of the entire detector is $2°.6$. The signals from the identically arranged cells of the three detectors were linearly added and fed through 37 channels to a charge−digit

converter (CDC). Thus, we obtain a digitized image of a Cherenkov flash. Flashes are recorded only when the amplitudes of time-coincident signals in any two of the 37 channels exceed a specified threshold. The resolving time of the coincidence circuit is 15 ns. The other three telescopes have a focal length of 3.2 m and are designed to record ultraviolet radiation in the range 200−300 nm. The total area of the mirrors on both mountings is 54 m^2; the effective gamma-ray detection threshold energy is 1.0 TeV.

OBSERVATIONS

Cyg X-3 (its coordinates for 1995 are $\alpha = 20^h32^m26^s$ and $\delta = 40°57'$) was observed in 1994 for 19 moonless nights from August 10 through October 9 and in 1995 for 14 moonless nights from July 27 through October 18. The observations were carried out in source−background mode; i.e., the observing session included an observation of the source for 35 min when the object was at the center of the field of view of the telescope camera and of the background with the same duration; the field of view of the gamma-ray telescope was shifted by 40 min in right ascension in order that its observation be performed at the same zenith and azimuth angles as those of the source. In 1994, we conducted 46 sessions, 21 of which were excluded from the subsequent reduction because of the low count rate of flashes or a great nonuniformity on the source or the background due to poor weather conditions and poor telescope guiding. In 1995, we conducted 18 sessions, seven of which were excluded from the subsequent reduction for the same reasons. The total duration of the source observations in the two years was 19 h 15 min. At the preliminary data reduction stage, we excluded from our analysis the flashes for which a maximum of CDC counts was reached at least in one channel and for which the channel with the largest amplitude was in the outer ring of the detector camera. As a result of the primary data reduction, 19 266 flashes in the source observations and 19 460 flashes in the background observations remained for the subsequent analysis in three years of observations. Thus, the difference between the numbers of flashes detected in the source (N_s) and background (N_b) observations is $N_\gamma = N_s - N_b = -194 \pm 197$, where 197 is the statistical error. To determine the possible gamma-ray flux requires selecting them by eliminating the flashes produced by the charged cosmic-ray component. The parameters of VHE gamma rays differ only slightly from the parameters of the flashes produced by charged cosmic-ray particles. Nevertheless, using this difference and eliminating the flashes produced mostly by cosmic rays, we can significantly increase the signal-to-noise ratio. The flash parameters can
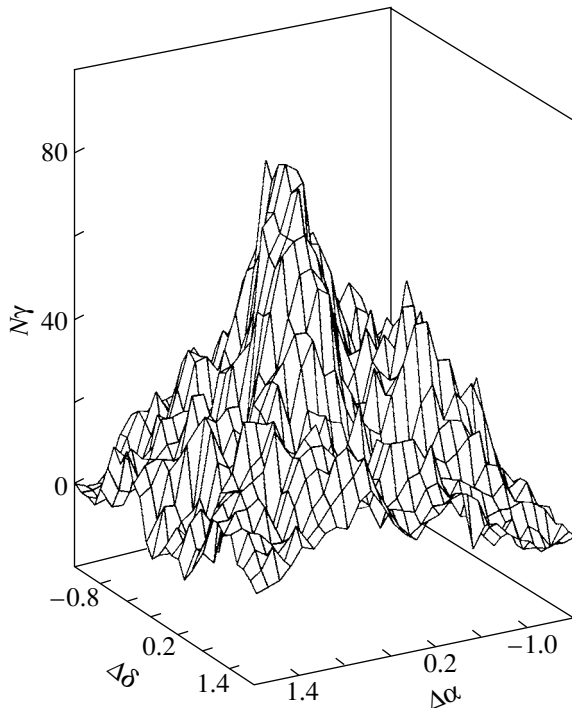
Fig. 1. The distribution of the number of selected gamma-ray photons relative to the source direction; $\Delta\alpha$ is the angle between the source and the specified direction in right ascension (in degrees); $\Delta\delta$ is the same angle for the declination; and $N\gamma$ is the number of selected gamma-ray photons.

Fig. 2. The mean count rate per minute for the gamma-ray photons selected in lunar months (MJD = JD−2400000.5).

be determined from the flash brightness distributions. The maximum and minimum second moments of the brightness distribution are commonly designated as Length and Width, respectively; the angle between the major axis of the brightness distribution ellipse and the source direction is designated as ALPHA. The position-dependent parameter is MISS = DIST sin(ALPHA), where DIST is the angular distance between the centroid of the brightness distribution and the source position.

In this case, the boundary values of the selection parameters must be properly chosen to obtain a maximum signal-to-noise ratio $Q = (N0_{\mathrm{s}} - N0_{\mathrm{b}})/\sqrt{N0_{\mathrm{s}} + N0_{\mathrm{b}}}$, where $N0_{\mathrm{s}}$ and $N0_{\mathrm{b}}$ are the numbers of gamma-ray-like flashes selected in the source and background observational data, respectively. The difference $N0_{\mathrm{s}} - N0_{\mathrm{b}} = N_\gamma$ is the selected number of gamma-ray-like flashes detected during the observations and $\sqrt{N0_{\mathrm{s}} + N0_{\mathrm{b}}}$ is the statistical error in the signal level after the selection. If the selection is made by several parameters, then approximately 99% of the flashes from the charged component can be eliminated (for the selection results, see below). Naturally, in this case, the number of flashes
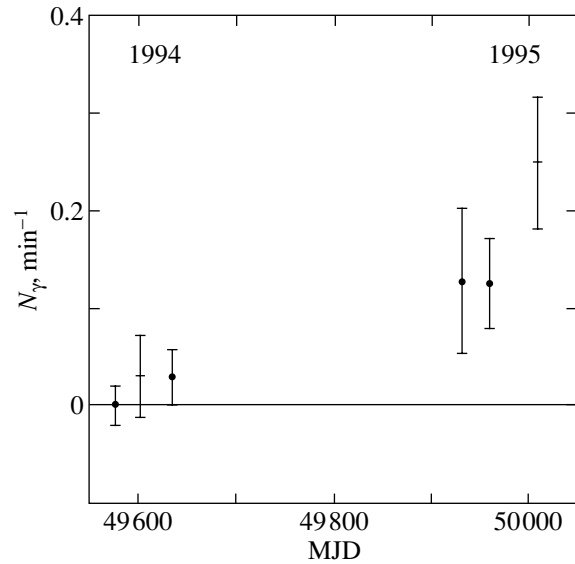
from gamma rays also decreases by approximately 30%.

The trial source method (Akerlof *et al.* 1991; Neshpor *et al.* 1994; Fomin *et al.* 1994) is used to determine the direction of the gamma-ray flux. The method is based on the fact that the major axes of the image ellipses for flashes from gamma rays in the telescope focal plane are oriented toward the source, whereas the major axes of the image ellipses for p-showers (showers from charged particles) are, to a first approximation, uniformly oriented in all directions. Therefore, if we select flashes by taking an arbitrary point in the focal plane with coordinates X_i and Y_j as the source direction and use the selection by position-dependent parameters, then the number of remained p-showers will not depend on the position of the putative source. In contrast, the number of images from gamma-ray showers will significantly depend on the position of the putative source and will have a maximum in the direction of the true source. In our case, the source direction coincided with the camera center.

We can construct the distribution of the number of selected flashes over the detector field of view as a function of the position of the putative source, $N(X_i, Y_j)$; i.e., a histogram map with the Cartesian coordinates of the trial source relative to the center of the detector field of view plotted along its X and Y axes and the number of gamma-ray-like flashes selected by parameters plotted along its Z axis. Thus, we can determine the position of the true gamma-ray source. Such histograms were constructed for the data from the northern and

southern sections. The selection was made both by the above position-independent parameters and by the position-dependent parameters DIST and MISS (see Fig. 1). The following limits were placed on the parameters: Length $< 0°30$, Width $< 0°17$, DIST $< 1°10$, and MISS $< 0°20$.

Having the maps for the two sections, we can construct a combined map. To correctly determine the statistical error, we must take into account the fact that some of the flashes selected from each section $Nc(X_i, Y_j)$ can coincide (in time). Such flashes must be treated as one flash. Therefore, for each section, we found gamma-ray-like flashes $Nc(X_i, Y_j)$ and gamma-ray-like flashes that remained in the data of only the northern ($N1(X_i, Y_j)$) or southern ($N2(X_i, Y_j)$) section after the selection by position-dependent parameters. In this case, the number of detected flashes identified as gamma-ray photons in the source observational data will be $N_s(X_i, Y_j) = Nc_s(X_i, Y_j) + N1_s(X_i, Y_j) + N2_s(X_i, Y_j)$. Similarly, for the background data, this quantity is $N_b(X_i, Y_j) = Nc_b(X_i, Y_j) + N1_b(X_i, Y_j) + N2_b(X_i, Y_j)$. Constructing a map from the background data is necessary to eliminate the instrumental and methodical effects (see, e.g., Neshpor *et al.* 1994).

The histogram of the difference $N_s(X_i, Y_j) - N_b(X_i, Y_j) = N_\gamma(X_i, Y_j)$ makes it possible to determine the position of the gamma-ray source with an accuracy of a few tenths of a degree. The histogram obtained in this way is shown in Fig. 1. The expression for the statistical error of N_γ is written as

$$\sigma = \sqrt{(Nc_s + N1_s + N2_s) + (Nc_b + N1_b + N2_b)}.$$

After the selection by the above parameters, we have $N_s(X_i, Y_j) = 162$, $N_b(X_i, Y_j) = 78$. In this case, we find that $N_\gamma = 84 \pm 15.5$. The N_γ peak has the coordinates $\alpha = 20^h 32^m 48^s$ and $\delta = 41°04'$. Thus, the N_γ peak coincides, to within $0°1$, with the position of Cyg X-3. We cannot assert with certainty that the N_γ peak coincides with the direction of Cyg X-3. However, taking into account the statistical accuracy of the data, we can assume this difference ($0°1$) to be negligible.

DISCUSSION

Our results suggest that, to all appearances, Cyg X-3 continues to be a VHE gamma-ray source. The mean gamma-ray flux from Cyg X-3 for 1972–1980 was 1.6×10^{-11} cm^{-2} s^{-1} for energy $> 2 \times 10^{12}$ eV (Stepanian 1984). We estimated the mean flux for 1994–1995 to be 1.3×10^{-11} cm^{-2} s^{-1}. However, our analysis of the observational data indicates that the gamma-ray fluxes are nonuniformly distributed in years. Figure 2 shows the flux distributions in lunar months. We see that in 1994, the fluxes were low or even statistically insignificant (the mean flux was $(6.2 \pm 2.6) \times 10^{-12}$ cm^{-2} s^{-1}). In 1995, the mean flux was $(2.7 \pm 0.7) \times 10^{-11}$ cm^{-2} s^{-1}. It should be noted that the measurements carried out at the Whipple Observatory (USA) in 1989–1992 yielded only an upper limit for the flux, 3.5×10^{-12} cm^{-2} s^{-1}. Thus, the VHE gamma-ray radiation from Cyg X-3 is clearly variable, as we pointed out previously (Neshpor *et al.* 1981).

REFERENCES

1. C. W. Akerlof, M. F. Cawley, M. Chantell, *et al.*, Astrophys. J. Lett. **377**, L97 (1991).
2. E. E. Becklin, G. Neugebauer, F. J. Hawkins, *et al.*, Nature **245**, 302 (1973).
3. C. R. Canizares, J. E. McClintock, G. W. Clark, *et al.*, Nature Phys. Sci. **241**, 28 (1973).
4. M. F. Cawley, D. J. Fegan, K. Gibbs, *et al.*, Astrophys. J. **296**, 185 (1985).
5. P. M. Chadwick, N. A. Dipper, J. C. Dawthwaite, *et al.*, Nature **318**, 642 (1982).
6. S. Danaher, D. J. Fegan, N. A. Porter, and T. C. Weekes, Nature **289**, 568 (1981).
7. J. C. Dawthwaite, A. I. Gibson, A. B. Harrison, *et al.*, Astron. Astrophys. **126**, 1 (1983).
8. V. P. Fomin, S. Fennell, R. C. Lamb, *et al.*, Astropart. Phys. **2**, 151 (1994).
9. A. M. Gal'per, V. G. Kirillov-Ugryumov, V. A. Kurochkin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 217 (1973) [JETP Lett. **18**, 169 (1973)].
10. R. Giacconi, P. Gorestein, H. Gursky, *et al.*, Astrophys. J. Lett. **148**, L119 (1967).
11. P. C. Gregory, P. P. Kronberg, E. R. Seaquist, *et al.*, Nature **239**, 440 (1972).
12. W. Hermsen, Space Sci. Rev. **36**, 61 (1983).
13. R. C. Lamb, C. F. Fichtel, R. C. Hartman, *et al.*, Astrophys. J. **212**, L63 (1977).
14. R. C. Lamb, C. P. Godfrey, W. A. Weaton, and T. Tumer, Nature **296**, 543 (1982).
15. D. B. Mukanov, Izv. Krym. Astrofiz. Obs. **63**, 151 (1981).
16. Yu. I. Neshpor and Yu. L. Zyskin, Perem. Zvezdy **21**, 709 (1982).
17. Yu. I. Neshpor and Yu. L. Zyskin, Pis'ma Astron. Zh. **12**, 452 (1986) [Sov. Astron. Lett. **12**, 189 (1986)].
18. Yu. I. Neshpor and Yu. L. Zyskin, Izv. Krym. Astrofiz. Obs. **78**, 122 (1988).
19. Yu. I. Neshpor, Yu. L. Zyskin, J. B. Mukanov, *et al.*, Phil. Trans. R. Soc. Lond. A **301**, 633 (1981).
20. Yu. I. Neshpor, A. P. Kornienko, A. A. Stepanian, and Yu. L. Zyskin, Exp. Astron. **5**, 405 (1994).

21. K. S. O'Flaherty, M. F. Cawley, D. J. Fegan, *et al.*, Astrophys. J. **396**, 674 (1992).
22. A. A. Stepanian, Adv. Space Res. **3**, 123 (1984).
23. A. A. Stepanian, Izv. Krym. Astrofiz. Obs. **58**, 51 (1978).
24. L. L. Stepanian, T. M. Vladimirsky, Yu. I. Neshpor, and V. P. Fomin, Astrophys. Space Sci. **38**, 267 (1975).
25. D. J. Thompson, D. L. Bertsch, B. L. Dingus, *et al.*, Astrophys. J., Suppl. Ser. **101**, 259 (1995).
26. M. H. van Kerkwijk, P. A. Charies, T. Geballe, *et al.*, Nature **355**, 703 (1992).
27. B. M. Vladimirsky, I. V. Pavlov, A. A. Stepanian, and V. P. Fomin, Izv. Akad. Nauk SSSR, Ser. Fiz. **36**, 2281 (1972).
28. B. M. Vladimirsky, A. A. Stepanian, and V. P. Fomin, *Proc. of the 13th Int. Conf. on Cosmic Rays, Denver, USA, 1973*, Vol. 1, p. 456.
29. B. M. Vladimirsky, A. A. Stepanian, and V. P. Fomin, Izv. Krym. Astrofiz. Obs. **51**, 3 (1974).
30. B. M. Vladimirsky, A. M. Gal'per, B. I. Luchkov, and A. A. Stepanian, Usp. Fiz. Nauk **145**, 255 (1985).
31. B. M. Vladimirsky, Yu. L. Zyskin, A. P. Kornienko, *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **55**, 2045 (1991).
32. B. M. Vladimirsky, Yu. L. Zyskin, A. P. Kornienko, *et al.*, Izv. Krym. Astrofiz. Obs. **91**, 74 (1994).

Translated by G. Rudnitskii

Timing Results for the Binary Millisecond Pulsar J1640+2224 Obtained on the RT-64 Radio Telescope in Kalyazin

V. A. Potapov[*], Yu. P. Ilyasov[**], V. V. Oreshko[***], and A. E. Rodin[****]

Pushchino Radio Astronomy Observatory, Astrospace Center, Lebedev Institute of Physics,
Russian Academy of Sciences, Pushchino, Moscow oblast, 142292 Russia
Received October 31, 2002

Abstract—We present the timing results for the binary millisecond pulsar J1640+2224 obtained with the RT-64 radio telescope (TNA-1500, Special Design Bureau, Moscow Power Engineering Institute) at the Kalyazin Observatory (Astrospace Center of the Lebedev Institute of Physics) in 1997−2002. We obtained Keplerian and post-Keplerian parameters of the binary system, which allowed us to estimate an upper limit for the energy density of the stochastic gravitational-wave background radiation at very low frequencies. © 2003 MAIK "Nauka/Interperiodica".

Key words: *pulsars, timing, gravitational waves.*

INTRODUCTION

Millisecond pulsars are neutron stars (NSs) with rotation periods from 1.5 to 60 ms that are generally members of binary systems. The high rotation frequency of the millisecond pulsars can be explained by the NS spinup due to the angular momentum transfer by accreting matter during the evolution of a binary system.

The pulsar pulse repetition period is known to be highly stable, which allows new time scales to be formed: the PT pulsar scale based on the intrinsic NS rotation period (Il'in *et al.* 1984) and the BPT scale based on the orbital period of a binary pulsar (Ilyasov *et al.* 1998; Petit and Tavella 1996).

The pulsar timing method that underlies the formation of pulsar time scales requires a regular recording of the times of arrival (TOAs) of pulsar pulses with a high accuracy over a long time interval (one year or longer). In general, the so-called residuals of the barycentric TOAs, i.e., the deviations of the observed TOAs to the Solar system barycenter from the predicted ones, are the subject of research. Analysis of long-term series of residuals yielded interesting results on general-relativity effects in binary systems (Damour and Taylor 1991), upper limits for the energy density of the stochastic gravitational-wave background (GWB) radiation (Lommen 2002; Kopeikin 1997, 1999; McHugh *et al.*

1996; Sazhin 1978; Thorsett and Dewey 1996), and the search for planetary systems (Wolszczan 1994; Rodin 2000).

The Kalyazin Radio Astronomy Observatory (KRAO, Astrospace Center, Lebedev Institute of Physics, Russian Academy of Sciences) is the only observatory in Russia at which millisecond and binary pulsars have been monitored as part of a timing program for several years. The rms deviations of the TOA residuals on long time scales for the various pulsars observed on the Kalyazin facility at a frequency of 0.6 GHz range from 1.7 to 13 μs, which corresponds to a relative rotation-period instability of 10^{-14}–10^{-13}.

Here, we analyze the timing series for the binary millisecond pulsar J1640+2224. This pulsar has a low level of the so-called intrinsic timing noise. A logarithmic measure of the timing noise is the activity parameter, which can be estimated from the formula $A = -1.31 + 0.71 \log \dot{P}$ (Cordes and Downs 1985). For the pulsar J1640+2224 with one of the lowest values of $\dot{P} = 2.8 \times 10^{-21}$ s s^{-1}, this parameter is $A = -5.3$. For comparison, the activity parameter for the single, highly stable pulsar B1937+21 is $A = -4.1$. In addition, as was shown by Ilyasov *et al.* (1998), the dispersion of the observed orbital parameters for a binary pulsar is sensitive to "red" noise of various astrophysical origins (the GWB manifests itself in the TOA residuals as noise with a spectrum $\propto 1/f^5$) and is proportional to the ratio P_b^2/x; for the pulsar J1640+2224, this ratio is relatively large, 4.153×10^{12} s. The low eccentricity of J1640+2224 allows the noise estimation formalism developed for

[*]E-mail: potap@prao.psn.ru
[**]E-mail: ilyasov@prao.psn.ru
[***]E-mail: oreshko@prao.psn.ru
[****]E-mail: rodin@prao.psn.ru

binary systems with circular orbits (Kopeikin 1997, 1999; Kopeikin and Potapov 1998, 2000) to be used. All of the aforesaid makes J1640+2224 a good candidate for GWB detectors.

Here, we briefly describe the receivers and the data acquisition system and the observing technique and present refined astrometric, spin, and orbital parameters of the pulsar. Based on the the observations, we estimated an upper limit for the GWB in the very-low frequency range. The results and prospects are briefly discussed.

THE OBSERVING TECHNIQUE. DATA ANALYSIS

The pulsar J1640+2224 was observed with the fully steerable RT-64 (TNA-1500) KRAO radio telescope. The AS-0.6 instrumental complex of the Pushchino Radio Astronomy Observatory (Astrospace Center, Lebedev Institute of Physics, Russian Academy of Sciences) (Ilyasov *et al.* 2000a, 2000b) was used for the reception. The pulsar pulses were accumulated by a spectrum analyzer in two circular polarizations, 80 channels in each, with a channel frequency band of 40 kHz. The observing sessions were conducted, on average, once every two weeks. The total signal integration time in each session is about 2 h. The TOAs were determined on a local time scale with an accuracy no less than 100 ns by referencing it to UTC(SU) via a TV channel and to UTC(USNO) via a GPS receiver. The pulse integration synchronous with the pulsar period was made in 3-min cycles followed by the system restart with a new value of the observed pulsar period. The data of each 3-min cycle were stored in separate files. The pulse delay in channels due to frequency dispersion was offset in the subsequent data reduction. This observing procedure makes it possible to check the quality of the observations during the integration, to reduce the phase shift error due to the Doppler shift of the observed pulsar frequency caused by the relative motion of the pulsar and the observer to the required value, and to filter out the noise during the data reduction.

The topocentric TOAs were determined by fitting the session-summed pulsar pulse profiles into a reference profile (template) with a high signal-to-noise ratio. We computed the barycentric TOAs, determined the TOA residuals, and refined the pulsar timing parameters by minimizing the residuals by the least-squares method using the TIMAPR (ftp://psun32.prao.psn.ru/pub/timapr) (Doroshenko and Kopeikin 1990) and TEMPO (http://pulsar. princeton.edu/tempo) (Taylor and Weisberg 1989) software. The DD model (Damour and Deruelle 1986)

was used to refine and compute the pulsar orbital parameters.

The table gives the least-squares estimates of the timing parameters for the pulsar J1640+2224. The values from the catalog of Taylor *et al.* (1993) were used as the initial values. The table gives the following data: α(J2000) and δ(J2000) are the pulsar right ascension and declination at epoch 2000; μ_α and μ_δ are the proper motions in right ascension and declination; P and $\dot{P}$ are the pulsar period and its first derivative; DM is the dispersion measure; d(DM) is the distance determined from the dispersion measure by using the model of the Galactic electron distribution (Taylor and Cordes 1993); x is the projection of the semimajor axis of the pulsar orbit onto the line of sight, in light seconds; e is the orbital eccentricity; T_p is the epoch of periastron passage; ω is the periastron longitude of the binary system; and σ_{res} is the rms deviation of the residuals calculated from the least-square estimates of the parameters for the pulsar model. Since the Kalyazin observations were carried out in single-frequency mode, DM and d(DM) were not calculated but were taken as the constant model parameters. For comparison, the second column of the table gives previously published data of the simultaneous reduction of the timing series obtained in 1993–1994 in Arecibo and in 1994–1998 in Effelsberg (Wolszczan *et al.* 2000). The timing results at the Kalyazin Observatory were reduced with the TIMAPR and TEMPO software and are presented in the third and fourth columns of the table. The formal least-squares estimate of the parameter dispersion, which refers to the last significant figures, is given in parentheses. The figure shows the variation of the pulsar TOA residuals.

ESTIMATING AN UPPER LIMIT FOR THE ENERGY DENSITY OF THE STOCHASTIC GRAVITATIONAL-WAVE BACKGROUND

One of the most interesting problems that can be solved by using pulsar timing is to estimate an upper limit for the energy density of the low-frequency gravitational-wave background (GWB), in particular, the relic background that was generated at early evolutionary stages of the Universe. It should be noted that, while the observations of single pulsars allow an upper limit to be estimated for frequencies 10^{-7}–10^{-9} Hz (Romani and Taylor 1983), the observations of binary pulsars potentially make it possible to detect radiation at much lower frequencies, 10^{-9}–10^{-12} Hz (Bertotti *et al.* 1983). At these frequencies, the effect of the relic stochastic gravitational-wave background becomes significant. In our work, we use the method of Kopeikin (1997),
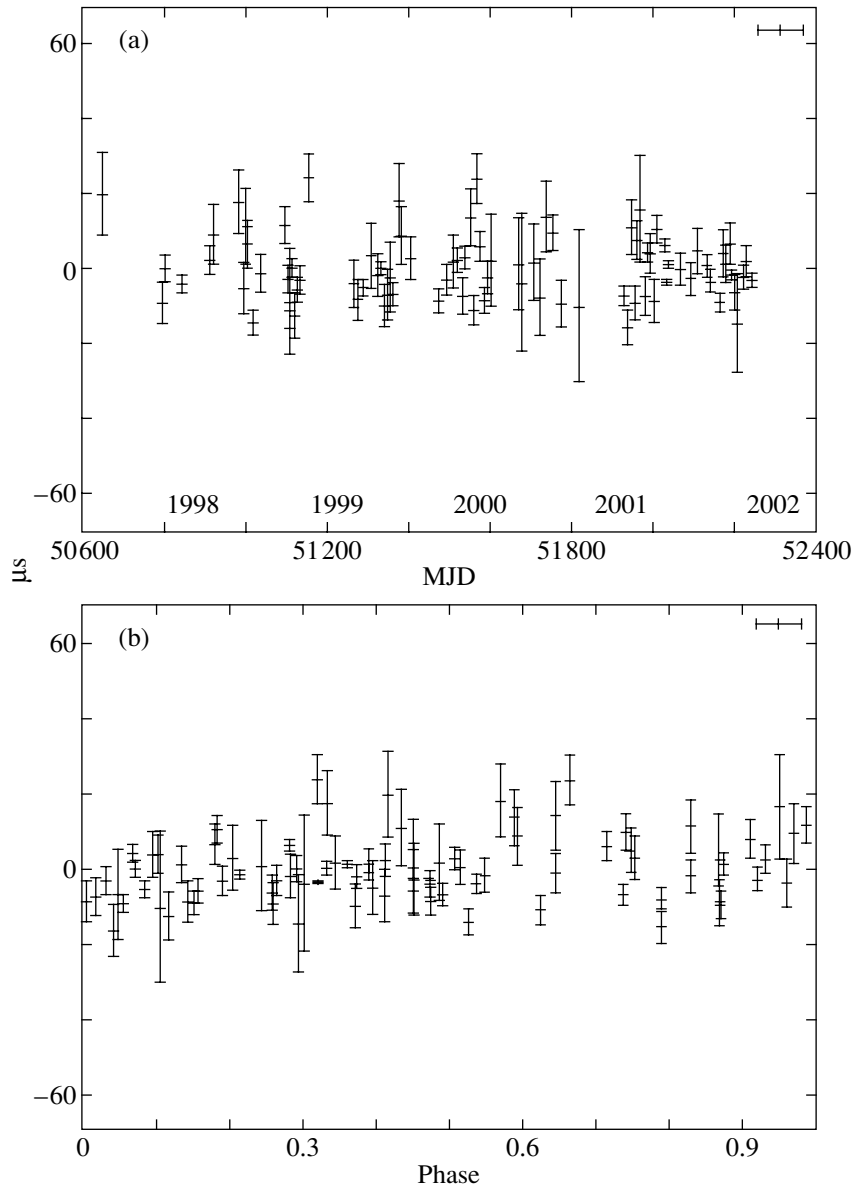
Fig. 1. (a) TOA residuals for the pulsar J1640+2224, as inferred from the data reduction with the TIMAPR software and (b) residuals *versus* pulsar orbital phase.

which allows an upper limit on the GWB logarithmic density Ω_g to be obtained from the variation in the relativistic parameter of the derivative of the orbital period of a binary pulsar with a circular orbit:

$$\Omega_g(f) = \frac{f}{\rho_c}\frac{d\rho_g(f)}{df}, \qquad (1)$$

where ρ_c is the critical energy density of the Universe and ρ_g is the energy density of the stochastic gravitational-wave background. The estimate is given in the frequency range $1/L < f < 1/T$, where L is the light travel time from the pulsar to the observer and T is the total observing time. In our case, $10^{-11} < f < 7.1 \times 10^{-9}$ Hz.

If we assume that the power spectrum of the noise of the TOA residuals produced by the GWB is given by the expression (Bertotti *et al.* 1983)

$$S(f) = \frac{H_0^2\Omega_g(f)}{16\pi^4}f^{-5}B(2\pi f L), \qquad (2)$$

$$B(\zeta) = 1 - \frac{3}{4}\left(\frac{2\zeta - \sin(2\zeta)}{\zeta^3}\right)$$

and that Ω_g is constant in the frequency range $10^{-9}-10^{-12}$ Hz concerned, then the estimate of the

Least-squares estimates of the timing parameters for the pulsar J1640+2224

Parameter	Arecibo−Bonn	Kalyazin	
		TIMAPR	TEMPO
α(J2000)	16:40:16.74173(3)	16:40:16.7417(2)	16:40:16.74064(5)
δ(J2000)	22:24:09.0113(6)	22:24:09.019(3)	22:24:08.9927(8)
μ_α, $0.001''$/year	−0.1(3)	1.2(5)	2.2(1)
μ_δ, $0.001''$/year	−13.3(7)	−12.1(5)	−8.2(1)
P, ms	3.1633158173430(2)	3.163315817342(1)	3.163315817345(1)
$\dot{P}/10^{-20}$	0.2827(2)	0.2831(5)	0.2822(3)
DM, cm^{-3} pc	18.4245(1)	−	−
d(DM), kpc	1.2(2)	−	−
x, s	55.3297195(8)	55.329728(6)	55.329730(1)
e	0.00079739(2)	0.00079737(4)	0.00079716(5)
T_p, MJD	49345.1877(7)	49345.190(2)	49345.157389(1)
P_b, day	175.4606613(2)	175.460662(3)	175.46256911(6)
ω, deg	50.726(2)	50.731(4)	50.664177(2)
σ_res, ms	8.7	8.8*	15.8**

* The rms deviations were obtained by reducing the series averaged over a 1-day interval.

** The rms deviations were obtained by reducing the series averaged over a half-hour interval.

GWB upper limit takes the form (Kopeikin 1997)

$$\Omega_g \leq \frac{56\pi^4}{5}\left[\sin^2\sigma + \frac{578}{\pi^2}\left(\frac{P_b}{T}\right)^2\cos^2\sigma\right]^{-1} \quad (3)$$

$$\times \left(\frac{T}{P_b}\right)^2\left(\frac{x}{P_b}\right)^2\left[\frac{\Delta\dot{P}_b/P_b}{10^{-10}yr^{-1}}\right]^2 h^{-2},$$

where h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}), P_b is the pulsar orbital period, x is the projection of the orbital semimajor axis, T is the total observing time (duration of the timing series), and σ is the pulsar orbital phase at the beginning of the observations.

We can estimate an upper limit for the variation $\Delta\dot{P}_b = \dot{P}_b^{\mathrm{obs}} - \dot{P}_b^{\mathrm{pred}}$ in formula (3):

$$\Delta\dot{P}_b \leq \max\left[\dot{P}_b^{\mathrm{obs}} - \dot{P}_b^{\mathrm{pred}}, \delta\dot{P}_b^{\mathrm{obs}}, \delta\dot{P}_b^{\mathrm{pred}}\right], \quad (4)$$

$$\dot{P}_b^{\mathrm{pred}} = \dot{P}_b^{\mathrm{gr}} + \dot{P}_b^k, \quad (5)$$

where $\dot{P}_b^{\mathrm{obs}}$ and $\delta\dot{P}_b^{\mathrm{obs}}$ are the observed derivative of the orbital period and its dispersion, $\dot{P}_b^{\mathrm{pred}}$ and $\delta\dot{P}_b^{\mathrm{pred}}$ are their theoretical values, $\dot{P}_b^k$ is the kinematic correction caused by the pulsar proper motion and its acceleration in the Galactic field, and $\dot{P}_b^{\mathrm{gr}}$ is the rate of change of the orbital period when gravitational waves are emitted by the binary system.

Let us estimate each of the quantities on the right-hand side of Eq. (4).

The values $\dot{P}_b^{\mathrm{obs}} \approx -1.4 \times 10^{-9}$ and $\delta\dot{P}_b^{\mathrm{obs}} \approx 0.6 \times 10^{-9}$ were obtained from the observational data after their reduction with the TEMPO software and are, strictly speaking, least-square estimates.

The eccentricity and orbital period required to estimate $\dot{P}_b^{\mathrm{gr}}$ were also obtained from the observations (table). The mass of the pulsar companion M_c was estimated by substituting reasonable values of the sine of the orbital inclination $\sin i$ and the pulsar mass M_p into the observed mass function $f = M_\mathrm{c}^3\sin^3(i)/(M_\mathrm{p} + M_\mathrm{c})^2$. Note that this is a standard method of estimation if the masses and the orbital inclination cannot be determined directly; in this case, the orbital inclination $i = 30°$ or $i = 60°$ is usually taken and the pulsar mass is limited astrophysically: $1M_\odot < M_\mathrm{p} < 4M_\odot$.

Assuming that $\sin i = 1/2$ and $M_\mathrm{p} = 1.4M_\odot$, we obtain $M_\mathrm{c} = 0.5M_\odot$, which gives $\dot{P}_b^{\mathrm{gr}} \approx -2.3 \times 10^{-18}$ in general relativity . The kinematic correction was calculated by using the data from Wolszczan *et al.* (2000), $\dot{P}_b^k \approx 6.0 \times 10^{-12}$.

Thus, it is clear that $\dot{P}_b^{\mathrm{pred}} \ll \dot{P}_b^{\mathrm{obs}}$. Since $\delta\dot{P}_b^{\mathrm{pred}} \sim \dot{P}_b^{\mathrm{pred}}$ and $|\delta\dot{P}_b^{\mathrm{obs}}| < |\dot{P}_b^{\mathrm{obs}}|$, $\dot{P}_b^{\mathrm{obs}}$ must be chosen as a majorizing estimate for the variation in the derivative of the orbital period:

$$\Delta\dot{P}_b \leq \dot{P}_b^{\mathrm{obs}} - \dot{P}_b^{\mathrm{pred}} \approx \dot{P}_b^{\mathrm{obs}} = -1.4 \times 10^{-9}.$$

Substituting the observed pulsar parameters, the total timing time, and the estimate of $\Delta\dot{P}_b$ into for-

mula (3) finally yields $\Omega_g h^2 \leq 8.5 \times 10^{-4}$ in the frequency range $10^{-11} < f < 7.1 \times 10^{-9}$ Hz.

CONCLUSIONS

The observations of the binary millisecond pulsar J1640+2224 with the PRAO AS-0.6 pulsar complex and the RT-64 KRAO (TNA-1500) radio telescope showed that the operation of the complex was highly stable in a long (more than five years) interval. The spin, astrometric, and orbital parameters of the pulsar were refined. The results are in good agreement with the previous Arecibo and Effelsberg observations (Wolszczan *et al.* 2000). The series of TOA residuals obtained from the Kalyazin observations is more accurate (has a lower dispersion of the residuals) than the series from Wolszczan *et al.* (2000). This fact can be explained by the regularity of the Kalyazin observations and by the use of the same pulsar instrumental complex throughout them, which allowed us to avoid the joining of heterogeneous data.

We estimated an upper limit for the stochastic GWB,

$$\Omega_g h^2 \leq 8.5 \times 10^{-4},$$

from the variation $\dot{P}_b$ by a direct method (Kopeikin 1997, 1999) at frequencies $10^{-11}-7.1 \times 10^{-9}$ Hz.

Note that the GWB energy density was estimated using pulsar timing by different methods at frequencies higher by two or three orders of magnitude (McHugh *et al.* 1996; Thorsett and Dewey 1996). A comparative analysis of the various approaches to solving this problem is an independent theoretical problem that is beyond the scope of this paper.

A further development of our work is the analysis of long timing series for the four millisecond pulsars that are being currently observed at the Kalyazin Radio Astronomy Observatory in an effort to refine their parameters and obtain an independent estimate for the GWB at very low frequencies $10^{-12}-10^{-9}$ Hz and in the range $10^{-9}-10^{-8}$ Hz by various methods.

ACKNOWLEDGMENTS

REFERENCES

1. B. Bertotti, B. J. Carr, and M. J. Rees, Mon. Not. R. Astron. Soc. **203**, 945 (1983).
2. J. M. Cordes and G. S. Downs, Astrophys. J., Suppl. Ser. **59**, 343 (1985).
3. T. Damour and N. Deruelle, Ann. Inst. Henri Poincaré, Phys. Theor. **44**, 263 (1986).
4. T. Damour and J. H. Taylor, Astrophys. J. **366**, 501 (1991).
5. O. V. Doroshenko and S. M. Kopeikin, Astron. Zh. **67**, 986 (1990) [Sov. Astron. **34**, 496 (1990)].
6. V. G. Il'in, Yu. P. Ilyasov, A. D. Kuzmin, *et al.*, Dokl. Akad. Nauk SSSR **275**, 835 (1984).
7. Yu. P. Ilyasov, S. M. Kopeikin, and A. E. Rodin, Pis'ma Astron. Zh. **24**, 275 (1998) [Astron. Lett. **24**, 228 (1998)].
8. Yu. P. Ilyasov, B. A. Poporechenko, and V. V. Oreshko, in *Radiostronomical Technics and Methods,* Ed. by O. N. Krokhin and N. S. Kardashev (Lebedev Phys. Inst., Moscow, 2000a), Part 2; Tr. Fiz. Inst. im. P.N. Lebedeva, Ross. Akad. Nauk **229**, 44 (2000a).
9. Yu. P. Ilyasov, V. V. Oreshko, and O. V. Dorochenko, in *Radiostronomical Technics and Methods,* Ed. by O. N. Krokhin and N. S. Kardashev (Lebedev Phys. Inst., Moscow, 2000b), Part 2; Tr. Fiz. Inst. im. P.N. Lebedeva, Ross. Akad. Nauk **229**, 95 (2000b).
10. S. M. Kopeikin, Phys. Rev. D **56**, 4455 (1997).
11. S. M. Kopeikin, Mon. Not. R. Astron. Soc. **305**, 563K (1999).
12. S. Kopeikin and V. Potapov, physics/9812050 (1998).
13. S. Kopeikin and V. Potapov, in *Proceedings of the 177th Colloquium of the IAU "Pulsar Astronomy-2000 and Beyond," Bonn, 1999,* Ed. by M. Kramer, N. Wex, and R. Wielebinski (Astronomical Society of the Pacific, San Francisco, 2000), Astron. Soc. Pac. Conf. Ser. **202**, 117 (2000).
14. A. N. Lommen, astro-ph/0208572 (2002).
15. M. P. McHugh, G. Zalamansky, F. Vernotte, and E. Lantz, Phys. Rev. D **54**, 5993 (1996).
16. G. Petit and P. Tavella, Astron. Astrophys. **308**, 290 (1996).
17. A. E. Rodin, in *Radiostronomical Technics and Methods,* Ed. by O. N. Krokhin and N. S. Kardashev (Lebedev Phys. Inst., Moscow, 2000), Part 2; Tr. Fiz. Inst. im. P.N. Lebedeva, Ross. Akad. Nauk **229**, 119 (2000).
18. R. W. Romani and J. H. Taylor, Astrophys. J. **265**, L35 (1983).
19. M. V. Sazhin, Astron. Zh. **55**, 65 (1978) [Sov. Astron. **22**, 36 (1978)].
20. J. H. Taylor and J. M. Cordes, Astrophys. J. **411**, 674 (1993).
21. J. H. Taylor, R. N. Manchester, and A. G. Lyne, Astrophys. J., Suppl. Ser. **88**, 529 (1993).
22. J. H. Taylor and J. M. Weisberg, Astrophys. J. **345**, 434 (1989).
23. S. E. Thorsett and R. J. Dewey, Phys. Rev. D **53**, 3468 (1996).
24. A. Wolszczan, Science **264**, 538 (1994).
25. A. Wolszczan, O. Doroshenko, M. Konacki, *et al.*, Astrophys. J. **528**, 907 (2000).

Translated by G. Rudnitskii

Dissipative-Acoustic Instability in Accretion Disks at a Nonlinear Stage

A. V. Khoperskov[*], S. S. Khrapov, and E. A. Nedugova

Volgograd State University, Vtoraya Prodol'naya ul. 30, Volgograd, 400062 Russia
Received July 10, 2002

Abstract—Nonstationary hydrodynamic models of a viscous accretion disk around a central compact object were constructed. Two different numerical methods (TVD and SPH) are used to study the dynamics of dissipatively unstable acoustic perturbations at the nonlinear stage in terms of the standard α-disk model. The standard disk accretion in the Shakura–Sunyaev model is unstable against acoustic waves for various parameters of the system. If the α parameter, which specifies the level of turbulent viscosity, exceeds $\alpha \gtrsim 0.03$, then a complex nonstationary system of small-scale weak shock waves is formed. The growth rate of the perturbations is higher in the central disk region. For $\alpha \gtrsim 0.2$, the relative shock amplitude can exceed 50% of the equilibrium disk parameters. The reflection of waves from the disk boundaries and their nonlinear interaction are important factors that can produce unsteady accretion. The luminosity of such a disk undergoes quasi-periodic oscillations at a level of several percent ($\lesssim 5\%$) of the equilibrium level. © 2003 MAIK "Nauka/Interperiodica".

Key words: *accretion disks, hydrodynamic models*.

INTRODUCTION

One of the fundamental problems of accretion disk flows is the question of the angular-momentum removal mechanisms that provide the fall of matter to a gravitating center. The progress made is largely based on viscous accretion-disk (AD) models (Shakura and Sunyaev 1973; Paczynski and Bisnovatyi-Kogan 1981; Bath and Pringle 1982; Smak 1984b; Artemova *et al.* 1996, 2001; Szuszkiewicz and Miller 1997; and others). An alternative possibility is the removal of angular momentum by large-scale spiral shocks produced during mass transfer in close binaries with a nonaxisymmetric potential (Spruit *et al.* 1987; Bisikalo *et al.* 1998, 2001).

Of particular interest in explaining the broad spectrum of nonstationary manifestations in accreting systems are unsteady flows attributable to the development of hydrodynamic instabilities. Quasi-periodic light curves with time scales longer than 10^5 s are a characteristic feature of dwarf novae. In particular, vertical convection in ADs was invoked to explain these light curves. The nonstationary models constructed by Smak (1984b), Mineshige and Shields (1995), and Agol *et al.* (2001) allow the viscosity level in ADs to be estimated. Comparison with observations indicates that the α parameter, which characterizes the viscosity, can reach $\alpha \simeq 0.2$ (Smak 1999) or even $\alpha \simeq 0.45$–0.65 in the binary

system GS 1124–683 (Lipunova and Shakura 2002). This result points to a possible high level of turbulent viscosity, at least in some binary systems. The observational data of Smak (1984a), O'Donoghue (1986), and Wood and Marsh (1989) on the dynamics of the outer disk radius in some dwarf novae (e.g., Z Cham and IP Peg) also argue for an extremely high $\alpha \gtrsim 0.1$ (Livio and Verbunt 1988; Ichikawa and Osaki 1992).

Here, we consider much more rapid processes that can be responsible, in particular, for quasi-periodic oscillations (QPOs) in X-ray sources without outburst activity with characteristic frequencies $\simeq 1$–40 Hz and lower-frequency ($\simeq 0.05$ Hz) QPOs in X-ray pulsars (van der Klis 2000). QPOs with even lower frequencies ($\simeq 0.008$ Hz) and with amplitudes of several percent were detected in a number of X-ray bursters (Revnivtsev *et al.* 2001). Millisecond X-ray flux variations are a glaring example of the nonstationarity of accreting systems (van der Klis 2000). It may well be that perturbations in the central disk regions during the nonlinear development of various instabilities are responsible for some of these QPOs[1] (Revnivtsev *et al.* 2001).

The stability of disk accretion onto a stellar-mass black hole was first considered by Lightman and Eardley (1974). A detailed analysis of the stability against axisymmetric perturbations with

---

[*]E-mail: Khoperskov@sai.msu.ru

[1] Note that in several cases, magnetic fields can be responsible for the variability, as in the case of magnetorotational instability (Reynolds and Armitage 2001).

radiation pressure was performed by Shakura and Sunyaev (1976). It became clear that the viscous models were unstable against thermal and viscous oscillation modes. Subsequently, stabilizing general-relativity effects (Abramowicz 1981), general polytropic models (Camenzind *et al.* 1986), and non-axisymmetric perturbations (Mc Kee 1990) were considered. Since the increment of the thermal mode is much larger than that of the viscous mode under typical conditions, the thermal oscillation branch can be responsible for the nonstationary processes in accreting systems (Szuszkiewicz and Miller 1997). Favorable conditions for the development of thermal instability exist only in the innermost AD regions, where the radiation pressure dominates over the gas pressure ($p_{rad} > 1.5p_{gas}$). In the inverse limit ($p_{rad} < 1.5p_{gas}$), the growth of thermal perturbations is possible only for certain specific viscosity and opacity laws (Khoperskov and Khrapov 1999a).

Szuszkiewicz (1990), Wallinder (1990, 1995), Khoperskov and Khrapov (1995), and Wu *et al.* (1995) showed that acoustic waves could, in principle, be unstable in the linear short-wavelength approximation in a viscous AD under a wide variety of assumptions about the viscous and optical disk properties. Significantly, acoustic perturbations can be unstable both in a radiation-dominated region (near a black hole) and in outer disk regions. The necessary conditions for the growth of acoustic perturbations include a strongly differential rotation, $d\ln\Omega/d\ln r \lesssim -1.3$, and the dependence of the coefficient of dynamic viscosity η on thermodynamic parameters (e.g., density and/or temperature) (Khoperskov and Khrapov 1999a). We know of other examples where dissipation (viscosity) is responsible for the instability of gaseous disks (Lynden-Bell and Pringle 1974; Morozov *et al.* 1985).

Here, our goal is to consider the dynamics of dissipatively unstable acoustic perturbations in ADs at the nonlinear stage. The algorithms used to solve the hydrodynamic equations are described in Section 1. For the TVD-E algorithm, we propose a version of computations convenient for studying the dynamics of perturbations in a viscous gaseous disk. In Section 2, the TVD algorithm is used to study the initial development of dissipative-acoustic instability. We show that the acoustic oscillation amplitude can, in principle, increase until the formation of a system of small-scale shock waves. We investigated a nonadiabatic model in Section 3, in which we obtained QPOs in the AD luminosity by the SPH method.

1. THE MODEL AND TECHNIQUE FOR NUMERICAL SIMULATIONS

1.1. Basic Equations

We consider a thin, $H \ll r$ (H is the vertical disk half-thickness), axisymmetric gaseous disk in the gravitational field of a central object of mass M ($\Phi(r) = -GM/r$). In the thin-disk model (TDM), the evolution of an AD with turbulent viscosity is described by the system of equations

$$\frac{\partial\sigma}{\partial t} + \frac{\partial}{r\partial r}(ru\sigma) = 0, \qquad (1)$$

$$\frac{du}{dt} - \frac{v^2}{r} = -\frac{1}{\sigma}\frac{\partial p}{\partial r} + \frac{1}{\sigma}\frac{\partial(rW_{rr})}{r\partial r} - \frac{W_{\varphi\varphi}}{\sigma r} - g, \quad (2)$$

$$\frac{dv}{dt} + \frac{uv}{r} = \frac{1}{\sigma}\frac{\partial(r^2 W_{r\varphi})}{r^2\partial r}, \qquad (3)$$

where $\frac{d}{dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial r}$; $g = \frac{\partial\Phi}{\partial r}$; u and v are the radial and azimuthal velocities in the disk plane, respectively; σ is the surface density; p is the surface pressure; $W_{r\varphi} = \eta\left(\frac{\partial v}{\partial r} - \frac{v}{r}\right)$, $W_{rr} = 2\eta\frac{\partial u}{\partial r} - \eta\frac{\partial(ru)}{r\partial r}$, and $W_{\varphi\varphi} = 2\eta\frac{u}{r} - \eta\frac{\partial(ru)}{r\partial r}$ are the surface viscous stress tensor components; and η is the coefficient of dynamic viscosity. We write the law of variations in specific internal energy ε when the gas pressure dominates ($p \equiv p_{gas} \gg p_{rad}$) as

$$\frac{d\varepsilon}{dt} + \frac{p}{\sigma}\frac{\partial(ru)}{r\partial r} = \frac{1}{\sigma}\left(Q^+ - Q^-\right). \qquad (4)$$

Here, $\varepsilon = \frac{p}{(\Gamma - 1)\sigma}$; $\Gamma = 1 + 2\frac{\gamma - 1}{\gamma + 1}$ is the flat adiabatic index; and Q^+ and Q^- are the energy released per unit time per unit area from the AD surface through viscous dissipation and the energy carried away by radiation, respectively. For Q^+, we have

$$Q^+ = W_{r\varphi}\left(\frac{\partial v}{\partial r} - \frac{v}{r}\right) + W_{rr}\frac{\partial u}{\partial r} + W_{\varphi\varphi}\frac{u}{r}. \quad (5)$$

For an optically thick disk, Q^- is given by the expression

$$Q^- = 2\frac{acT^4}{\bar{\kappa}\sigma}, \qquad (6)$$

where a is the radiation constant, c is the speed of light, $\bar{\kappa}$ is the Rosseland averaged opacity, and T is the mean AD temperature in the vertical direction. For a plasma of completely ionized hydrogen in the case of local thermodynamic equilibrium and when the gas pressure dominates over the radiation pressure, $T = \frac{m_p}{2k_B}\frac{p}{\sigma}$ (here, k_B is the Boltzmann constant and m_p is the proton mass).

The applicability of the TDM in this case was studied in detail by Khoperskov and Khrapov (1999b). They compared the dynamics of linear perturbations in the TDM and in the model with vertical motions.

For the acoustic perturbations to be unstable, the dynamic viscosity must be assumed to depend on the surface pressure p, the surface density σ, or both of these quantities. A linear analysis indicates that the models of ADs with different turbulent viscosity laws $\eta(\sigma, p)$ can admit the growth of acoustic waves (Khoperskov and Khrapov 1995, 1999a). Consider a simple model of turbulent viscosity that leads to the instability of acoustic perturbations. Following Shakura and Sunyaev (1973), we assume for the kinematic viscosity $\nu = \eta/\sigma$ that

$$\nu = \frac{2}{3}\alpha\frac{p}{\sigma\Omega_K},\qquad(7)$$

where $\Omega_K = \sqrt{GM/r^3}$ is the Keplerian angular velocity. Relation (7) corresponds to the expression for the tensor component $W_{r\varphi} = -\alpha p$ in the case of the azimuthal velocity $v = r\Omega_K$. For the characteristic disk half-thickness, we have $H \sim c_s/\Omega_K$ ($c_s = \sqrt{\Gamma p/\sigma}$ is the adiabatic speed of sound), which gives an estimate of the viscosity coefficient $\nu \sim \alpha c_s H$.

As the characteristic distance, time, and mass, we take, respectively, $r_* = 3r_G = 6GM/c^2$ (r_G is the gravitational radius); $t_* = 2\pi\sqrt{r_*^3/GM}$, the disk rotation period at radius r_*; and $m_* = \dot{m}t_*$, where $\dot{m} = \dot{M}/\dot{M}_{crit}$ characterizes the rate of accretion onto the central compact object, $\dot{M}_{crit} = 3 \times 10^{-8}mM_\odot/$yr is the critical accretion rate, and m is the central mass measured in solar masses $M_\odot$. We then have $u_* = r_*/t_*$ for the unit of linear velocity, $\Omega_* = 2\pi/t_*$ for the unit of angular velocity, $\mathfrak{L}_* = r_*u_*$ for the unit of specific angular momentum, and $\sigma_* = m_*/r_*^2$ for the unit of surface density. Given these quantities, the basic system of equations can be easily made dimensionless. We denote the steady-state solutions of the system of equations (1)–(4) by the subscript 0.

1.2. Numerical Schemes

The study of the dynamics of small-scale acoustic waves that evolve into a system of shock waves imposes special requirements on the numerical scheme. We use two numerical algorithms to solve the system of equations (1)–(4).

Algorithm I. To analyze the detailed structure of nonlinear waves, we use the TVD-E numerical scheme (Harten 1983; Ryu *et al.* 1993), which we write in a form convenient for computing the perturbation dynamics against the background of the equilibrium state. In this model, we take into account only

those factors that are minimally required to describe the dissipative instability of acoustic perturbations. We restrict our analysis to the adiabatic case by assuming that the condition $Q^+(t) = Q^-(t)$ is always satisfied. Khoperskov and Khrapov (1999a) showed that nonadiabaticity weakly affects the dynamics of linear acoustic waves, at least. Since the equilibrium radial velocity $u_0(r)$ in the models of thin axisymmetric ADs is low compared to the speed of sound $c_{s0}(r)$ and the azimuthal velocity $v_0(r)$ ($u_0 \sim (H/r)c_{s0}$, $c_{s0} \sim (H/r)v_0$), the relation $|W_{rr}| \sim |W_{\varphi\varphi}| \ll |W_{r\varphi}|$ holds for long-wavelength perturbations, $kH \lesssim 1$ (k is the radial wave number). Therefore, we take into account only the $W_{r\varphi}$ viscous stress tensor component, for which we assume, in this case, that $W_{r\varphi} = -\alpha p$.

Under the above assumptions, the following relations hold for $\gamma = 5/3$ for the steady-state solution of the system of equations (1)–(4): $\sigma_0 \propto r^{-1/2}$, $p_0 \propto r^{-3/2}$, $v_0 \propto r^{-1/2}$, $\Omega_0 = v_0/r \propto r^{-3/2}$, $u_0 \propto r^{-1/2}$, and $c_{s0} \propto r^{-1/2}$. In this case, the Mach number does not depend on the radial coordinate, $\mathfrak{M} \equiv v_0/c_{s0} = $ const.

It follows from a linear analysis of the stability (see Khoperskov and Khrapov 1999a) that the wave number is $k^2(r) \simeq \dfrac{\omega^2 - \Omega_0^2(r)}{c_{s0}^2(r)}$ (ω is the acoustic wave frequency). For the equilibrium distributions under consideration for $\omega \gg \Omega_0(r)$, we obtain $k \propto r^{1/2}$; i.e., the perturbation wavelength decreases with the radius ($\lambda \propto r^{-1/2}$). Therefore, it is convenient to change to a new variable x, which is related to r by $x = r^{3/2}$. After this change, the perturbation wavelength along the x coordinate remains constant, because the perturbed quantities in the WKB approximation are $\propto \exp\left\{i\int k(r)dr\right\} = \exp\left\{i\hat{k}x\right\}$ ($\hat{k} = $ const).

Substituting the solution $f(r,t) = \hat{f}(x,t)r^{\delta_f}$ ($f = \{\sigma, p, u, v\}$, δ_f is the exponent of the equilibrium distributions for f) into the system of equations (1)–(4) and changing from the variable r to x, we obtain

$$\frac{\partial\mathbf{q}}{\partial t} + \frac{3}{2}\frac{\partial\mathbf{F}}{\partial x} = \mathbf{S},$$

$$\mathbf{q} = \begin{pmatrix} \hat{\sigma} \\ \hat{\sigma}\hat{u} \\ \hat{\sigma}\hat{v} \\ \hat{E} \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \hat{\sigma}\hat{u} \\ \hat{\sigma}\hat{u}^2 + \hat{p} \\ \hat{\sigma}\hat{u}\hat{v} \\ \hat{u}(\hat{E} + \hat{p}) \end{pmatrix}, \qquad(8)$$

$$\mathbf{S} = \frac{1}{x} \begin{pmatrix} 0 \\ \frac{3}{2}\hat{p} + \hat{\sigma}\left(\frac{1}{2}\hat{u}^2 + \hat{v}^2 - \hat{g}\right) \\ -\frac{1}{2}\hat{\sigma}\hat{u}\hat{v} - D \\ \hat{u}(\hat{E} + \hat{p} - \hat{\sigma}\hat{g}) - \hat{v}D \end{pmatrix},$$

where $\hat{E} = \hat{\sigma}\dfrac{\hat{u}^2 + \hat{v}^2}{2} + \hat{\varepsilon}$, $D = \alpha x\left(\dfrac{3}{2}\dfrac{\partial\hat{p}}{\partial x} + \dfrac{\hat{p}}{2x}\right)$.

According to the TVD-E algorithm, the integration of the system of equations (8) breaks up into three stages in this case. At the first stage, intermediate values of $\mathbf{q}^{n*}$ are calculated at time t^n by taking into account the acting sources $\mathbf{S}$:

$$\mathbf{q}_i^{n*} = \mathbf{q}_i^n + \frac{\Delta t^n}{2}\mathbf{S}(\mathbf{q}_i^n).$$

At the second stage, the changes in the parameters attributable to the mass, momentum, and energy fluxes through the Eulerian cell boundaries are calculated:

$$\mathbf{q}_i^{nx} = \mathbf{q}_i^n - \frac{3}{2}\frac{\Delta t^n}{\Delta x}\left(\mathbf{f}_{i+1/2}^* - \mathbf{f}_{i-1/2}^*\right).$$

Here, $\mathbf{f}_{i\pm1/2}^* = \dfrac{1}{2}\left[\mathbf{F}(\mathbf{q}_i^{n*}) + \mathbf{F}(\mathbf{q}_{i\pm1}^{n*})\right] - \dfrac{\Delta x}{2\Delta t^n} \times \sum_{k=1}^4 \beta_{k,i\pm1/2}^* \mathbf{R}_{k,i\pm1/2}^{n*}$; $\Delta x = (x_{\text{out}} - x_{\text{in}})/N_x$ is the cell size; x_{in} and x_{out} specify the inner and outer boundaries of the computational region, respectively; and N_x is the number of cells.

At the third, final stage, the flow parameters are determined at the following time $t^{n+1} = t^n + \Delta t^n$;

$$\mathbf{q}_i^{n+1} = \mathbf{q}_i^{nx} + \Delta t^n\mathbf{S}(\bar{\mathbf{q}}_i^{n+1/2}),$$

where $\bar{\mathbf{q}}^{n+1/2} = \dfrac{1}{2}[\mathbf{q}_i^n + \mathbf{q}_i^{nx} + \Delta t^n\mathbf{S}(\mathbf{q}_i^n)]$. The quantities $\beta_{k,i\pm1/2}^*$, $\bar{\mathbf{R}}_{k,i\pm1/2}^{n*}$, and Δt^n are determined in the standard way in terms of the TVD-E algorithm (see, e.g., Harten 1983; Ryu et al. 1993).

In the basic model, the inner and outer boundaries of the computational region are $x_{\text{in}} = 30$ ($r_{\text{in}} \simeq 10$) and $x_{\text{out}} = 6030$ ($r_{\text{out}} \simeq 330$), respectively. We performed our computations for grids with various resolutions, $N_x = 3 \times 10^2 - 3 \times 10^4$. The boundary conditions at the inner and outer boundaries of the computational region were specified in the following form: $\mathbf{q}^{n+1}(x_{\text{in}} - \Delta x) = \mathbf{q}^{n+1}(x_{\text{in}} + \Delta x)$, $\mathbf{q}^{n+1}(x_{\text{out}} + \Delta x) = \mathbf{q}^{n+1}(x_{\text{out}} - \Delta x)$.

Our test computations show that an equilibrium solution in the numerical algorithm under consideration ensures that the unperturbed disk is stationary with an accuracy higher than $10^{-10}\%$ during at least 10^5 rotations at radius $3r_G$ ($t \sim 10^5$). Therefore,

this method of solution is convenient for studying the dynamics of unstable perturbations against the background of the disk equilibrium state. A drawback is the nonphysical influence of the disk boundaries when the waves reach the radius r_{in} or r_{out}. Therefore, our computations stopped at this time.

Algorithm II. To study the dynamics of unstable perturbations with an allowance made for the wave reflection from the disk boundaries, which is needed for describing the long-term disk evolution, it is more preferable to use schemes that take into account the surfaces on which the gas density tends to zero. To this end, we integrated the hydrodynamic equations in Lagrangian form by the SPH method. In this case, we consider a nonadiabatic AD model; i.e., we integrate Eq. (4) for $Q^+(t) \neq Q^-(t)$. This allows us, on the one hand, to investigate the effect of nonadiabaticity on the dynamics of nonlinear acoustic waves and, on the other hand, to derive the time dependence of the AD luminosity. Note that in the nonadiabatic AD model, using the expression $W_{r\varphi} = -\alpha p$ can lead to the violation of the condition $Q^+ \geq 0$. Therefore, in this case, we consider the more general expression

$$W_{r\varphi} = \eta r\frac{\partial\Omega(r,t)}{\partial r}.$$

According to the SPH method (Kim et al. 1994; Simpson 1995), any parameter of the medium $A(r,t)$ can be substituted with a smoothed value $\langle A(r)\rangle = \int A(r')\mathfrak{W}(|r - r'|, h)dr'$. For a finite number of particles N, the integration can be substituted with the summation

$$A(r) = \sum_{j=1}^N \frac{\mu_j}{r_j\sigma_j}A(r_j)\mathfrak{W}(|r - r_j|, h), \qquad (9)$$

where μ_j is the mass of particle j whose density is distributed in the annulus $r_j - 2h < r < r_j + 2h$. The kernel $\mathfrak{W}(s, h)$ that smoothes the local fluctuations and satisfies the normalization condition $2\pi\int_0^\infty \mathfrak{W}(s, h)ds = 1$ was chosen in the form

$$\mathfrak{W}(s, h) = \frac{1}{3h\pi} \qquad (10)$$
$$\times \begin{cases} 1 - \frac{3}{2}\left(\frac{s}{h}\right)^2 + \frac{3}{4}\left(\frac{s}{h}\right)^3, & 0 \leq \frac{s}{h} \leq 1 \\ \frac{1}{4}\left(2 - \frac{s}{h}\right)^3, & 1 \leq \frac{s}{h} \leq 2 \\ 0, & \frac{s}{h} > 2, \end{cases}$$

where $s = |r - r_j|$. All particles are assumed to have equal masses $\mu \equiv \mu_j$. For the derivative, we have

$$\frac{\partial A(r)}{\partial r} = \sum_{j=1}^N \frac{\mu}{r_j\sigma_j}A(r_j)\frac{\partial\mathfrak{W}(|r - r_j|, h)}{\partial r}. \qquad (11)$$
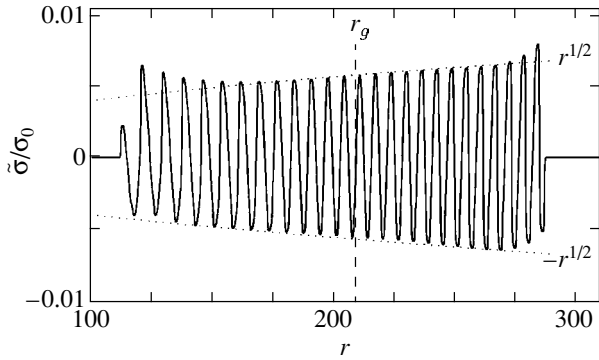
Fig. 1. Surface-density perturbation $\tilde{\sigma}$ normalized to the equilibrium value σ_0 versus radial coordinate r for $\alpha = 0$ at time $t = 4000$. The perturbation generator is located at the radius $r = r_g$. The dotted lines indicate the asymptotics for the wave amplitude.

Using (9), the surface density of particle i can be represented as

$$\sigma_i = \sum_{j=1}^{N} \frac{\mu}{r_j} \mathfrak{W}(|r_i - r_j|, h). \quad (12)$$

Using (9), (11), and (12), Eqs. (2)–(4) can be written as

$$\frac{du_i}{dt} = -\sum_{j=1}^{N} \frac{\mu}{r_j} \left(\frac{p_j}{\sigma_j^2} + \frac{p_i}{\sigma_i^2} + \Pi_{ij} \right) \frac{\partial \mathfrak{W}_{ij}}{\partial r_i}$$
$$+ \frac{\mathfrak{L}_i^2}{r_i^3} - \frac{4\pi^2}{r^2}, \quad (13)$$

$$\frac{d\mathfrak{L}_i}{dt} = \sum_{j=1}^{N} \mu \left[\frac{(W_{r\varphi})_j}{\sigma_j^2} + \frac{(W_{r\varphi})_i}{\sigma_i^2} \right] \frac{\partial \mathfrak{W}_{ij}}{\partial r_i}, \quad (14)$$

$$\frac{d\varepsilon_i}{dt} = \frac{1}{2} \sum_{j=1}^{N} \mu \left(\frac{p_j}{\sigma_j^2} + \frac{p_i}{\sigma_i^2} + \Pi_{ij} \right) \quad (15)$$

$$\times \left(\frac{u_i}{r_j} - \frac{u_j}{r_i} \right) \frac{\partial \mathfrak{W}_{ij}}{\partial r_i} + \frac{Q_i^+}{\sigma_i} - \frac{Q_i^-}{\sigma_i},$$

where $\mathfrak{W}_{ij} \equiv \mathfrak{W}(|r_i - r_j|, h)$, $(W_{r\varphi})_i = \nu_i \sum_{j=1}^{N} \mu \times$ $\left(\frac{\mathfrak{L}_j}{r_j^2} - \frac{\mathfrak{L}_i}{r_i^2} \right) \frac{\partial \mathfrak{W}_{ij}}{\partial r_i}$, $\mathfrak{L}_i = r_i v_i$ is the specific angular momentum of particle i, $Q_i^+ = \sigma_i \sum_{j=1}^{N} \mu \frac{(W_{r\varphi})_j}{\sigma_j^2} \times$ $\left(\frac{\mathfrak{L}_j}{r_j^2} - \frac{\mathfrak{L}_i}{r_i^2} \right) \frac{\partial \mathfrak{W}_{ij}}{\partial r_i}$, and Π_{ij} is the numerical viscosity that simulates the kinematic scattering (Kim *et al.* 1994; Simpson 1995). The position of particle i can be determined by integrating the equation $dr_i/dt =$

u_i. Note that Eqs. (12)–(15) identically satisfy the integral laws of conservation of mass, angular momentum, and total energy.

To numerically integrate the system of equations (13)–(15), we used the predictor–corrector method with a second order of accuracy in time (see, e.g., Kim *et al.* 1994). An efficient algorithm of detecting the nearest particles is needed to implement the SPH method; it allows us to significantly reduce the computational time and optimize the code. When calculating the parameters of the medium and the forces acting on any particle, it is necessary that all of the particles within a given region be identified, which is achieved by sorting the particles (Simpson 1995).

The initial state for the nonadiabatic AD model is specified by the condition $Q_0^+ \simeq Q_0^-$. For a Keplerian disk $\Omega_0 \propto r^{-3/2}$ at $r \gg 1$, this condition yields $p_0 \propto r^{-3/2}, \sigma_0 \propto r^{-3/5}, v_0 \propto r^{-1/2}$, and $u_0 \propto r^{-2/5}$. These solutions correspond to the standard α-AD model of Shakura and Sunyaev (1973) when the gas pressure and Thomson scattering dominate.

The computations were performed with various numbers of particles, $N = 2500$–$10\,000$. The disk luminosity $L(t)$ was determined by integrating $Q^-(r, t)$: $L(t) = 2\pi \int rQ^- dr$.

2. THE DYNAMICS OF NONLINEAR WAVES

In this section, we consider the dynamics of dissipative-unstable acoustic perturbations based on the numerical integration of Eqs. (1)–(4) by using algorithm I (see Subsect. 1.2).

2.1. The Dynamics of Perturbations from a Stationary Source

Consider the dynamics of perturbations from a periodic source of the form

$$\hat{g}(x, t) = 4\pi^2$$

$$\times \left(1 + A_g \sin(\omega_g t) \exp\left\{ -\frac{(x - x_g)^2}{\Delta_g^2} \right\} \right), \quad (16)$$

where A_g and ω_g are the amplitude and frequency of the source, respectively. In order to get rid of the influence of the boundaries, the perturbation source in the numerical model is localized within the computational region near the radius $r_g = x_g^{2/3}$ far from the outer and inner disk boundaries. The characteristic scale Δ_g is chosen in such a way that the conditions $\Delta x \ll \Delta_g \ll x_{\text{out}} - x_{\text{in}}$ are satisfied. Let us define the basic model of the perturbation source: $x_g = 3000$, $\omega_g = 2\pi/300$, $\Delta_g = 10$, and $A_g = 0.005$. Unless otherwise specified, the parameters take these values.
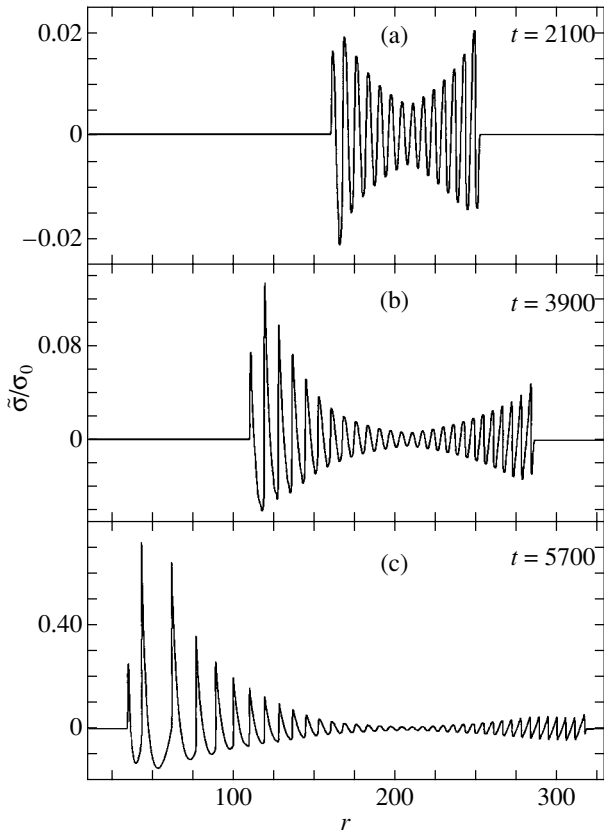
Fig. 2. The radial dependences of relative surface-density perturbations $\tilde{\sigma}/\sigma_0$ at various times for the basic model.
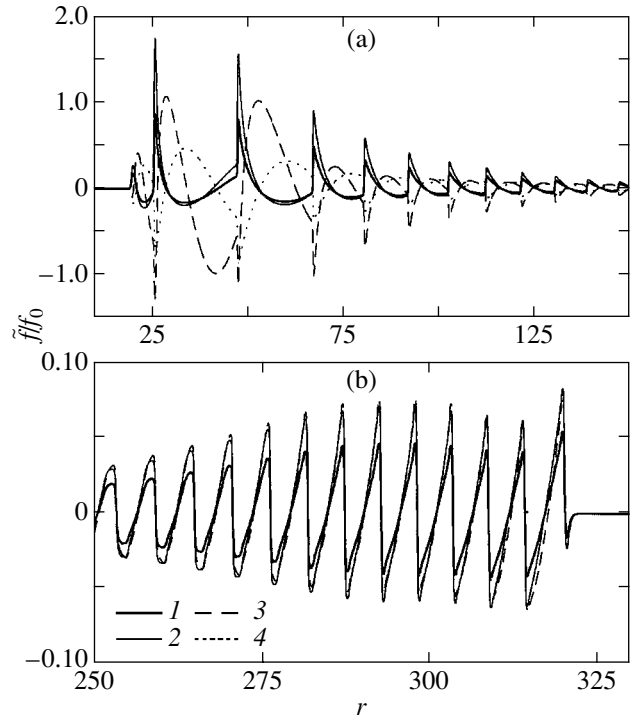

Fig. 3. The radial structure of the perturbed quantities for (a) inwardly and (b) outwardly propagating waves: *1*, the relative surface-density perturbation $\tilde{\sigma}/\sigma_0$; *2*, the relative pressure perturbation $\tilde{p}/p_0$; *3*, the relative azimuthal-velocity perturbation $\tilde{v}/v_0 \times 10^2$; and *4*, the relative radial-velocity perturbation $\tilde{u}/|u_0| \times 10^{-2}$.

Figure 1 shows the radial distribution of perturbations from a stationary source with $A_g = 0.005$ in a nondissipative disk ($\alpha = 0$). This disk is stable and the variations in wave amplitude are attributable only to the geometry of the problem and the radial disk inhomogeneity. As we see from Fig. 1, the relative amplitude of the surface-density perturbation increases with the radius as $\tilde{\sigma}/\sigma_0 \propto r^{1/2}$ ($\tilde{\sigma} = \sigma - \sigma_0$), while the wavelength decreases with the radius as $\lambda \propto r^{-1/2}$. Such radial dependences are also obtained for the pressure, radial-velocity, and azimuthal-velocity perturbations. Thus, for $|\tilde{f}/f_0| \ll 1$, $\omega_g \gg \Omega(r)$, and $r \gg 1$, the perturbed quantities can be represented as

$$\tilde{f}(r,t) = A_f r^{\tilde{\delta}_f} \cos\left(\hat{k}r^{3/2} - \omega_g t\right), \qquad (17)$$

where A_f is the perturbation amplitude and $\tilde{\delta}_f = \delta_f + \delta$. For arbitrary power-law equilibrium models, δ can be determined from the constancy condition for the mean acoustic-wave energy flux density

$$\langle r\tilde{u}\tilde{p}\rangle = \langle rc_s\varepsilon_v\rangle = \text{const}, \qquad (18)$$

where ε_v is the acoustic-wave energy density. Condition (18) leads to $\delta = -(1 + \delta_c + \delta_p)/2$, and we have

$\delta = 1/2$ for the equilibrium model under consideration ($\delta_c = -1/2$, $\delta_p = -3/2$).

The dynamics of waves in a viscous disk ($\alpha > 0$) changes qualitatively. Figure 2 shows the radial distributions of the surface density normalized to its equilibrium value at various times. As the waves propagate from the source inward and outward, the amplitude of these perturbations increases due to dissipative-acoustic instability (DAI). At $t = 2100$ (Fig. 2a), the relative surface-density amplitude does not exceed 2% and there are no shock waves (SWs). Starting from $t \gtrsim 3900$, conditions for the SW formation in the perturbed regions farthest from the source arise (Fig. 2b). By $t \simeq 6000$, the perturbations reach the inner (r_{in}) and outer (r_{out}) boundaries. In the central zone, the SW amplitudes are much larger than those in the outer disk region. This effect primarily stems from the fact that the instability increment is proportional to the angular velocity of the disk ($\Omega \propto r^{-3/2}$). This is also partly responsible for the saturation of the amplitudes of the outward propagating SWs (see Fig. 2c).

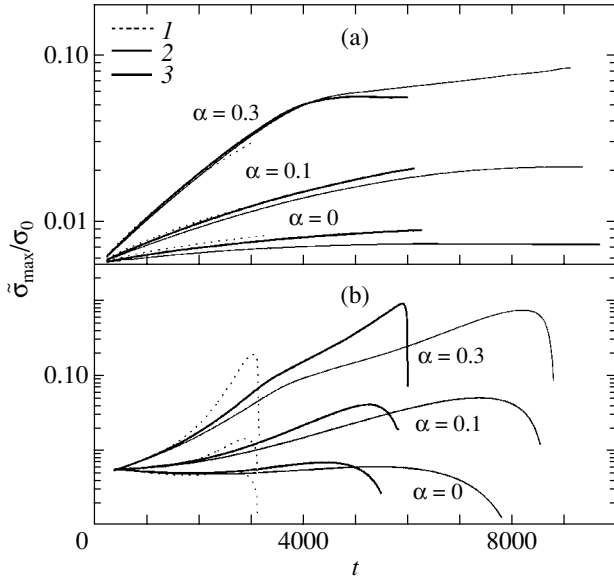Perturbations with a given frequency, which is determined by source (16), can propagate without

Fig. 4. Relative amplitude of the surface-density perturbation $\tilde{\sigma}_{max}/\sigma_0$ versus time t for (a) outwardly and (b) inwardly propagating waves at various values of α and the Mach number ($\mathfrak{M}$): *1*, $\mathfrak{M} = 10$; *2*, $\mathfrak{M} = 20$; *3*, $\mathfrak{M} = 30$.

damping only to the radius r_c, at which $\Omega(r_c) = \omega_g$ (for the computation shown in Fig. 2, we have $r_c \simeq 45$). At $r < r_c$, the SW amplitude decreases and the perturbations are spatially restructured. Only the small-scale component penetrates the $r > r_c$ region. Thus, the higher the frequency ω_g and, accordingly, the shorter the perturbation wavelength, the closer to the inner disk edge the SWs penetrate from the disk periphery.

Figure 3 shows the radial structures of the surface-density, radial-velocity, azimuthal-velocity, and pressure perturbations. For inwardly propagating waves (U^-), the radial- and azimuthal-velocity perturbations are out of phase with the surface-density and pressure perturbations (Fig. 3a), while for outwardly propagating waves (U^+), all of the perturbations are in phase (Fig. 3b). As we see from Fig. 3b, an additional modulation of the wave envelope arises for the U^+ perturbations that reached nonlinear saturation. This effect is enhanced with increasing Mach number.

The AD model considered here is described by two parameters—the α parameter and the Mach number ($\mathfrak{M}$). Since the DAI increment is proportional to α, the growth rate of the perturbations decreases with decreasing α. The wavelength and velocity of the acoustic perturbations decrease with increasing Mach number. In Fig. 4, the maximum relative amplitudes of the surface-density perturbations are plotted against time for U^+ (see Fig. 4a) and U^- (see Fig. 4b) waves at various values of α and $\mathfrak{M}$. As we see from

Fig. 4, the DAI growth rate at the linear stage is virtually independent of $\mathfrak{M}$. The model with $\alpha = 0$ is stable and the radial variations of the relative perturbation amplitude $\tilde{\sigma}/\sigma_0$ are attributable only to the geometric factor and the radial equilibrium-disk inhomogeneity.

At the nonlinear stage, the growth rate of the acoustic perturbations decreases with increasing $\mathfrak{M}$. Since the travel time of the acoustic waves from the source to the boundaries increases (with increasing $\mathfrak{M}$), the conditions for the SW formation through nonlinear steepening improve. Thus, as the Mach number increases, SWs can be formed even at lower values of α. The presence of maxima for the U^- waves (see Fig. 4b) is related to the SW passage through the radius r_c at which $\Omega(r_c) = \omega_g$. Note also that the U^+ waves are saturated on large time scales (at large radii) with decreasing source amplitude A_g. For the U^- waves, this effect is much less pronounced, because the nonlinear stage is reached faster than for U^+.

Consider the structure of the shock front. For weak SWs, the front width is $\ell \sim \dfrac{\nu p}{c_s \Delta p}$ in order of magnitude. Since the SWs are weak, $|\Delta p/p| \lesssim 1.5$, and since the coefficient of turbulent viscosity is large, we obtain an estimate of $\ell \sim \alpha H$ and the shock front width can be comparable to the characteristic disk thickness H. To avoid the influence of numerical effects, we performed our computations for detailed grids with the number of cells N_x ranging from 1000 to 30 000. Starting from some values of N_x, which depend primarily on α and the Mach number $\mathfrak{M}$, the shock front width ceases to depend on N_x. For example, at $\alpha = 0.3$ and $\mathfrak{M} = 20$, the critical value of N_x is 12 000. At small $N_x \lesssim 300$, we were unable to obtain a system of SWs. Figure 5 shows the structure of the SW fronts with $N_x = 30\,000$ cells. The dots denote the pressures in the computational cells. As we see, the SW front width is several times larger than the cell size.

It should be noted that the growth rate of the perturbations increases with increasing N_x but ceases to depend on N_x at some values of $N_x \geq N_x^{crit}$. The critical value N_x^{crit} is primarily determined by the Mach number and can assume $N_x^{crit} \sim 300 \cdot \mathfrak{M}$ for $\alpha \sim 0.1$.

2.2. The Dynamics of an Initial Perturbation

Consider the dynamics of the waves attributable to the presence of a local AD perturbation at the initial time in the absence of a source. We specify the perturbation in the range $x_1 \leq x \leq x_2$ in the form

$$\hat{v}(x, t = 0) = \hat{v}_0(x) \left[1 + A_v \sin \left(2\pi n \frac{x - x_1}{x_2 - x_1} \right) \right],$$
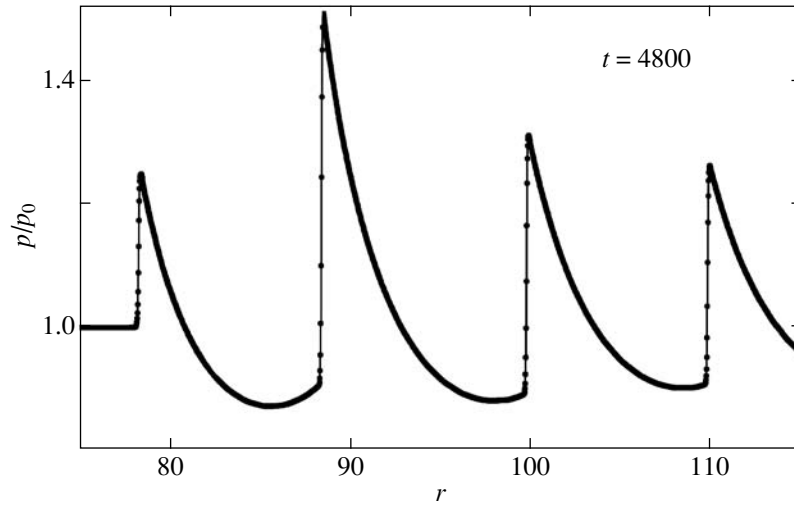
$$(19)$$

Fig. 5. The structure of shock waves. The distribution of pressure $p(r)$ normalized to its equilibrium value p_0 in the shock region at time $t = 4800$. The dots indicate the positions of the computational cells ($N_x = 30\,000$).
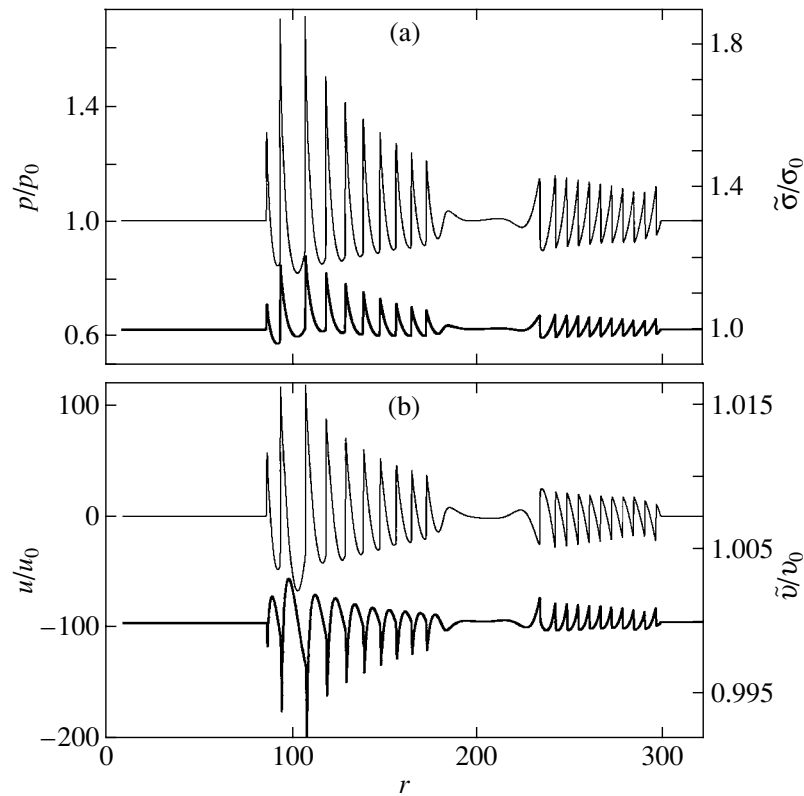

Fig. 6. The radial distributions of flow parameters at time $t = 5900$ if the initial density perturbation was located near the radius $r = 200$: (a) $\tilde{p}/p_0$ (thin line), $\tilde{\sigma}/\sigma_0$ (heavy line) and (b) u/u_0 (thin line), $\tilde{v}/v_0$ (heavy line).

where A_v is the perturbation amplitude and n is the number of wavelengths in the range from x_1 to x_2. In our computations, these parameters took the following values: $A_v = 0.005$, $n = 10$, $x_1 = 2250$, and $x_2 = 3750$. At $x < x_1$ and $x > x_2$, there are no perturbations at the initial time.

In this case, two wave packets propagating inward and outward are formed (Fig. 6). An equilibrium unperturbed flow is established behind these packets after their passage. The perturbation structure proves to be similar to the models considered in Subsect. 2.1. In particular, the wave amplitude increases until the
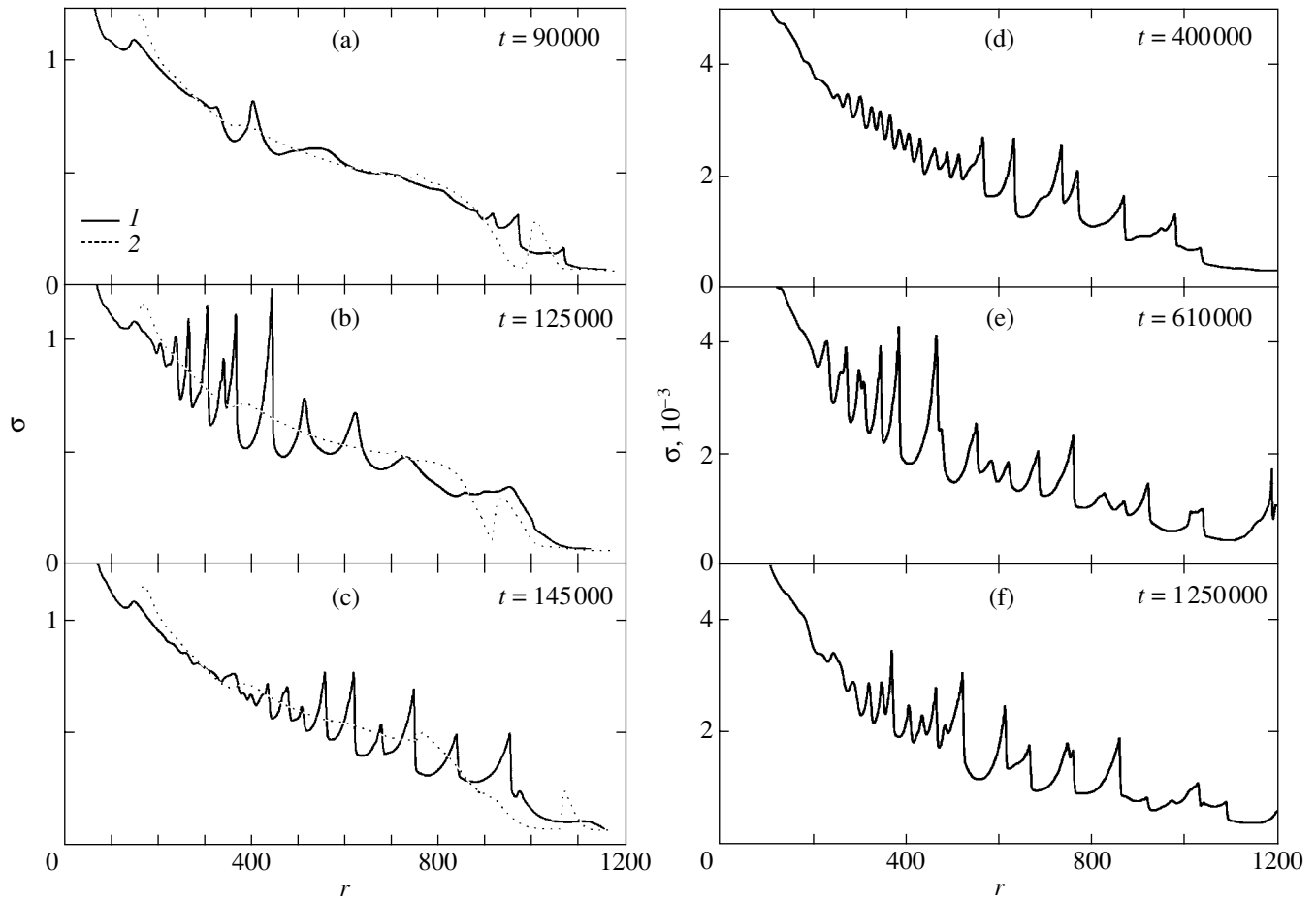
Fig. 7. The radial dependences of surface density σ at various times ($\alpha = 0.3$, $\mathfrak{M} = 40$): *1* for viscosity $\eta \propto \alpha p$ (unstable model), *2* for $\eta \propto \alpha p_0$ (stable model).
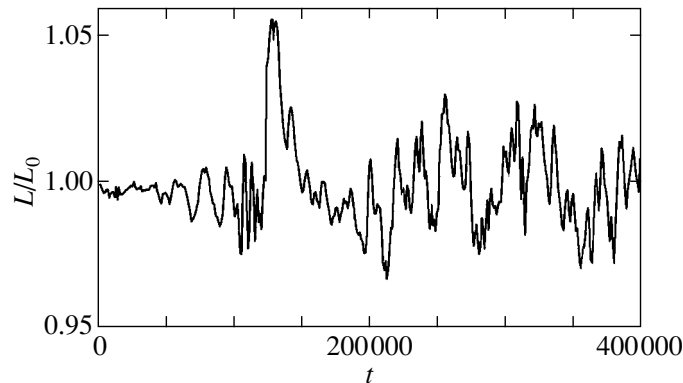

Fig. 8. Total AD luminosity (L) normalized to its equilibrium value L_0 versus time for $\alpha = 0.3$, $\mathfrak{M} = 40$.

shock formation, although the energy of the initial perturbation is low. This indicates that the increase in the perturbation amplitude is determined not by the properties of source (16) but by the instability of the initial flow. At the same time, an external impact is required for the formation of appreciable perturbations in an initially stationary disk.

The waves whose amplitude further increases as they pass through an unstable disk are partially re-

flected at the outer and inner disk boundaries.[2] This mechanism of multiple wave amplification cannot be considered in terms of the TVD method, and we study the long-term evolution of the perturbations in the next section.

3. THE DYNAMICS OF ACOUSTIC WAVES IN A NONADIABATIC AD

The models considered above do not admit any generalization to a nonadiabatic flow and do not allow the dynamics of the perturbations that approach the boundaries of the computational region to be accurately described. To study these factors, let us consider the evolution of an AD in terms of the SPH model (see Subsect. 1.2).

At the initial time, matter occupies the region $100 < r < 1000$ with a power-law distribution of the parameters inside the disk: $\sigma(r, 0) \propto r^{-3/5}$, $p(r, 0) \propto r^{-3/2}$, $v(r, 0) \propto r^{-1/2}$. Since the gaseous disk is viscous, a radial flow immediately emerges in it. One part of the disk matter moves inward, while its other part moves outward. A viscous spreading of the disk takes place, with the total angular momentum of the AD being constant.

Since in this model, in contrast to the model considered in Sect. 2, perturbations are always present at a level $\lesssim 1\%$, and since the standard α-AD model is unstable against acoustic oscillation modes (Khoperskov and Khrapov 1995), a system of shock waves is gradually formed in the disk because of the growth of these perturbations before the nonlinear stage. In Fig. 7, the surface density σ is plotted against the radial coordinate r at various times. As in the adiabatic model, the perturbations grow most rapidly in the inner AD regions (Fig. 7b); they evolve into outwardly propagating shock waves (Fig. 7c). The perturbations from outer disk regions propagate inward but their amplitude is smaller, because the DAI instability is proportional to the angular velocity $\Omega \propto r^{-3/2}$. As a result, the perturbation structure in the disk becomes complex on long time scales (see Fig. 7c). This is attributable to a nonlinear interaction of the oppositely propagating waves and to the reflection of these waves from the disk boundaries. These effects cause the shock waves to be intensified. Thus, since the perturbations are reflected from the disk boundaries, a system of high-amplitude shock waves can be formed even at low α.

As in the TVD-E model (see Sect. 3), the formation rate of the structures considered above decreases with decreasing α and/or increasing Mach number $\mathfrak{M}$.

Our computations with the model of turbulent viscosity $\eta \propto \alpha p_0(r)$, which does not lead to any dissipative-acoustic instability (because $\partial \eta / \partial t = 0$), show that no shock waves are formed (see the dotted lines in Fig. 7) and that quasi-steady accretion is established.

In Fig. 8, the relative disk luminosity is plotted against time. Intense wave motions in the disk produce QPOs of the AD luminosity with characteristic frequencies $(0.1-2)/m$ Hz for the specified disk parameters. It should be emphasized that the nonstationary luminosity component originates mainly in the central disk region.

Thus, based on two different numerical methods of integrating the hydrodynamic equations, we showed that the amplitude of initially small perturbations could increase until the formation of a system of small-scale shock waves through dissipative-acoustic instability.

CONCLUSIONS AND DISCUSSION

We constructed nonstationary axisymmetric hydrodynamic models of disk accretion onto a compact stellar-mass object. A viscous quasi-Keplerian disk with the dynamic viscosity dependent on thermodynamic parameters of the gas is unstable against acoustic perturbations. We investigated the dynamics of the dissipative-acoustic instability (DAI) at the nonlinear stage. In our numerical gas-dynamical simulations of a viscous α disk in the presence of initial perturbations with an amplitude $\gtrsim 0.1\%$ of the equilibrium values, the acoustic oscillations can grow to weak shock waves in the time of acoustic-wave passage through the disk. As a result, a nonstationary system of small-scale shock waves, whose relative amplitude can exceed 50% for $\alpha \gtrsim 0.2$, is formed in the disk. These wave motions in the AD give rise to a nonstationary disk luminosity component with an amplitude of several percent. The efficiency of these processes (shock formation and quasi-periodic oscillations) strongly depends on the level of turbulent viscosity and, without considering the effects of reflection from the boundaries, a system of large-amplitude waves can be formed only for $\alpha \gtrsim 0.05$. At low $\alpha \lesssim 0.05$, the wave amplitude is small and the periodic variations in the accretion rate and luminosity are negligible. It should be noted, however, that because of the wave reflection from the outer and inner disk boundaries on long time scales $t \gtrsim 10^6 t_*$ (t_* is the Keplerian disk rotation period at radius $3r_G$), the SW amplitude can also be large at low $\alpha \lesssim 0.05$.

---

[2]The stellar surface, the magnetospheric boundary, or the region of a highly nonuniform matter density as $3r_G$ is approached in the case of a black hole can act as the inner boundary. At the outer boundary, for some scenarios of mass transfer in a close binary, the region of a sharp change in density can also be small.

The acoustic perturbations are also dissipatively unstable when the radiation pressure dominates, $p_{rad} \gg p_{gas}$, i.e., at the innermost region of the AD around a black hole. In this region, the dynamic time is short ($\lesssim 10^{-3}$ s) and the characteristic times of the nonstationary processes because of the development of DAI can be $\ll 1$ s. Therefore, the formation mechanism of the nonstationary radiation component under consideration can also give a contribution in the X-ray range; in particular, it can be responsible for the millisecond oscillations in X-ray binaries (van der Klis 2000). However, a proper study of the QPO structure that would allow comparison with observational data requires using a two-dimensional model with a given mass inflow in the AD, which is planned to be done in the future. Note that the conditions for the growth of acoustic perturbations are more favorable in a nonaxisymmetric disk.

The standard α disk of Shakura and Sunyaev (1973) underlies our model. In the linear approximation, however, Khoperskov and Khrapov (1999b) showed that the conditions for the growth of DAI are retained when the dependence of the coefficient of dynamic viscosity on thermodynamic parameters is varied widely. In some viscosity models, the instability increment can be larger than that in the standard α disk.

The medium behind the shock front is additionally turbulized, which increases the turbulent viscosity. As a result, the conditions for the development of DAI improve. The system of small-scale shock waves is additionally self-sustained.

Since the most unstable wavelengths are $\lambda \simeq 10H$, a detailed numerical grid with a number of cells no fewer than $N \gtrsim 10^3$ for an annulus of width $r_{out} - r_{in} \sim 300 \cdot 3r_G$ is required to accurately describe these perturbations. If the outer disk radius in a typical binary is assumed to be 7×10^{10} cm, then for $3r_G = 10^6$ cm, we obtain a constraint on the number of cells in the radial direction, $N \gtrsim 200\,000$. Therefore, for $N \simeq 100$, which is typical in the simulations of two- and three-dimensional hydrodynamic flows in a close binary, the DAI cannot be studied.

ACKNOWLEDGMENTS

REFERENCES

1. M. A. Abramowicz, Nature **294**, 235 (1981).
2. E. Agol, J. Krolik, N. J. Turner, and J. M. Stone, Astrophys. J. **558**, 543 (2001).
3. I. V. Artemova, G. S. Bisnovatyi-Kogan, G. Bjoernsson, and I. D. Novikov, Astrophys. J. **456**, 119 (1996).
4. I. V. Artemova, G. S. Bisnovatyi-Kogan, I. V. Igumenshchev, and I. D. Novikov, Astrophys. J. **549**, 1050 (2001).
5. G. T. Bath and J. E. Pringle, Mon. Not. R. Astron. Soc. **199**, 267 (1982).
6. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).
7. D. V. Bisikalo, A. A. Boyarchuk, A. A. Kilpio, *et al.*, Astron. Zh. **78**, 707 (2001) [Astron. Rep. **45**, 611 (2001)].
8. M. Camenzind, F. Demole, and N. Straumann, Astron. Astrophys. **158**, 212 (1986).
9. D. O'Donoqhue, Mon. Not. R. Astron. Soc. **220**, L23 (1986).
10. A. Harten, J. Comput. Phys. **49**, 357 (1983).
11. S. Ichikawa and Y. Osaki, Publ. Astron. Soc. Jpn. **44**, 15 (1992).
12. A. V. Khoperskov and S. S. Khrapov, Pis'ma Astron. Zh. **21**, 388 (1995) [Astron. Lett. **21**, 347 (1995)].
13. A. V. Khoperskov and S. S. Khrapov, Astron. Zh. **76**, 256 (1999a) [Astron. Rep. **43**, 216 (1999a)].
14. A. V. Khoperskov and S. S. Khrapov, Astron. Astrophys. **345**, 307 (1999b).
15. W.-T. Kim, S. S. Hong, and H. S. Yun, J. Kor. Astron. Soc. **27**, 13 (1994).
16. A. P. Lightman and D. M. Eardley, Astrophys. J. Lett. **187**, L1 (1974).
17. G. V. Lipunova and N. I. Shakura, Astron. Zh. **79**, 407 (2002) [Astron. Rep. **46**, 366 (2002)].
18. M. Livio and F. Verbunt, Mon. Not. R. Astron. Soc. **232**, L1 (1988).
19. D. Lynden-Bell and J. E. Pringle, Mon. Not. R. Astron. Soc. **168**, 603 (1974).
20. M. R. McKee, Astron. Astrophys. **235**, 521 (1990).
21. S. Mineshige and G. A. Shields, Astrophys. J. **351**, 47 (1990).
22. A. G. Morozov, Yu. M. Torgashin, and A. M. Fridman, Pis'ma Astron. Zh. **11**, 231 (1985) [Sov. Astron. Lett. **11**, 94 (1985)].
23. B. Paczynski and G. Bisnovatyi-Kogan, Acta Astron. **31**, 283 (1981).
24. M. Revnivtsev, E. Churazov, M. Gilfanov, and R. Sunyaev, Astron. Astrophys. **372**, 138 (2001).
25. C. S. Reynolds and P. J. Armitage, Astrophys. J. Lett. **561**, L81 (2001).
26. S. Ryu, J. P. Ostriker, H. Kang, and R. Cen, Astrophys. J. **414**, 1 (1993).
27. N. I. Shakura and R. A. Sunyaev, Astron. Astrophys. **34**, 337 (1973).
28. N. I. Shakura and R. A. Sunyaev, Mon. Not. R. Astron. Soc. **175**, 613 (1976).
29. J. S. Simpson, Astrophys. J. **448**, 822 (1995).
30. J. Smak, Acta Astron. **34**, 93 (1984a).
31. J. Smak, Acta Astron. **34**, 161 (1984b).

32. J. Smak, Acta Astron. **49**, 391 (1999).

33. H. C. Spruit, T. Matsuda, M. Inoue, and K. Sawada, Mon. Not. R. Astron. Soc. **229**, 517 (1987).

34. E. Szuszkiewicz, Mon. Not. R. Astron. Soc. **244**, 377 (1990).

35. E. Szuszkiewicz and J. C. Miller, Mon. Not. R. Astron. Soc. **287**, 165 (1997).

36. M. van der Klis, Ann. Rev. Astron. Astrophys. **38**, 717 (2000).

37. F. H. Wallinder, Astron. Astrophys. **237**, 270 (1990).

38. F. H. Wallinder, Mon. Not. R. Astron. Soc. **273**, 1133 (1995).

39. J. H. Wood and T. R. Marsh, Mon. Not. R. Astron. Soc. **239**, 809 (1989).

40. X.-B. Wu, Q.-B. Li, Y.-H. Zhao, and L.-T. Yang, Astrophys. J. **442**, 736 (1995).

Translated by V. Astakhov

Measurement of the Longitudinal Magnetic-Field Component for FU Ori

D. A. Smirnov[1*], S. A. Lamzin[1], and S. N. Fabrika[2]

[1]*Sternberg Astronomical Institute, Moscow State University, Universitetskii pr. 13, Moscow, 119992 Russia*
[2]*Special Astrophysical Observatory, Russian Academy of Sciences,*
Nizhnii Arkhyz, Karachai-Cherkessian Republic, 357147 Russia
Received November 4, 2002

Abstract—The Main Stellar Spectrograph of the 6-m Special Astrophysical Observatory telescope equipped with a polarimetric analyzer was used to measure the longitudinal magnetic-field component of FU Ori on January 24, 2002. The following (3σ) upper limits were obtained for the magnetic field $B_{||}$: $B_{||} < 350-400$ G in the formation region of Fe I, Ni I, and Ca I absorption lines (disk + wind), and $B_{||} < 200$ G in the formation region of the absorption component of the Hα line with a P Cyg profile. We conclude that the strength of a large-scale magnetic field capable of collimating the disk wind does not exceed 300 G. For the region where the emission component of the Hα line is formed, we found that $B_{||} < 100$ G. Such a low value may have been obtained because the magnetic field lines in this region were almost perpendicular to the line of sight at the time of our observations. © *2003 MAIK "Nauka/Interperiodica".*

Key words: *stars, FU Ori stars, magnetic field, disk accretion, stellar wind.*

INTRODUCTION

FU Ori is the prototype of the subclass of young stars in which a brightening by 4^m-6^m in less than a decade was observed (Herbig 1966). A large increase in the rate of accretion from the disk surrounding a T Tauri star is currently believed to be responsible for the outburst. As a result, the accretion luminosity is hundreds of times higher than the luminosity of the central star. Therefore, the observed absorption spectrum of the FU Ori objects (fuors) is thought to be the spectrum of an accretion disk (see Hartmann and Kenyon (1996) and references therein).

The Balmer lines, as well as the Na I D and Ca II K lines, in the spectrum of FU Ori are known to exhibit deep and broad P Cyg profiles, suggesting an intense outflow of matter from the neighborhood of the star (see, e.g., Fig. 2). Hartmann and Calvet (1995) gave convincing arguments that the outflow originates from the surface of a Keplerian disk. The wind observed in optical lines "blows" from the inner ($R \leq 3 \times 10^{12}$ cm) accretion-disk regions; the outflowing gas has a velocity $V_w \sim 300$ km s^{-1}, a low temperature ($T_w < 6000$ K), and a high density: $N \sim 10^{12}-10^{13}$ cm^{-3} (Calvet *et al.* 1993). In this case, the mass loss rate is high: $\sim 10^{-5}$ $M_\odot$ yr^{-1}.

To explain the presence of emission components in the Hα, Na I D, Ca II K, and Mg II h, k lines,

D'Angelo *et al.* (2000) assumed that there were regions with temperature inversion (the chromosphere) at the wind base. By analyzing the pattern of profile variations for the Hα, Na I D, and Mg II h, k lines in the spectrum of FU Ori, Errico *et al.* (2003) concluded that the profile variations are attributable to the axial rotation of the inner accretion-disk and disk-wind regions where the temperature distribution and the wind streamline orientation are not axisymmetric. The interaction of circumstellar matter with the stellar magnetic field, provided that the magnetic axis is appreciably inclined to the rotation axis, may be responsible for the asymmetry.

Thus, the role of the magnetic field for FU Ori may be twofold: first, the magnetic field lines may act as the guides along which the disk-wind matter moves and, second, the stellar magnetic field may be responsible for the appearance of an asymmetry in the structure of the disk and the wind. Therefore, it is of current interest to detect the magnetic field of FU Ori and to study its parameters.

The method of magnetic-field measurements used here is based on the fact that, for the Zeeman splitting of the so-called σ-component, the lines are circularly polarized, with the oppositely polarized components being located on both sides of the central wavelength λ_0. If the magnetic field in the line formation region has a longitudinal (directed along the line of sight) component $B_{||}$, then, during observations in right- and left-hand circularly polarized light,

the lines will be shifted from each other by (Babcock 1958)

$$\Delta\lambda_{rl} \simeq 9.3 \times 10^{-10} g\lambda_0^2 B_{||} \text{ mÅ}, \qquad (1)$$

where g is the Lande factor of the line under consideration. The wavelength in relation (1) is in Å and $B_{||}$ is in G. This relation allows the mean longitudinal magnetic-field component in the line formation region to be measured by measuring $\Delta\lambda_{rl}$ from two spectra obtained in right- and left-hand circularly polarized light.

OBSERVATIONAL DATA

The spectra of FU Ori under study were obtained on January 24, 2002, with the Main Stellar Spectrograph of the 6-m Special Astrophysical Observatory telescope equipped with a circular polarization analyzer (Chuntonov 1997). The 6450–6700 Å spectra were taken with a 1160×1024 CCD detector. The spectrograph slit width, $0\overset{''}{.}5$, provided a spectral resolution $R \simeq 15\,000$. The weather conditions on the night of our observations were stable, with the seeing being within $1\overset{''}{.}5$ most of the time.

We reduced the spectra using the MIDAS software package. To measure the difference between the line positions in the spectra with opposite polarizations, we used a cross-correlation technique. This technique allows the shift of the line as a whole to be measured and has a low sensitivity to errors in the continuum placement (Klochkova *et al.* 1996).

To take into account the systematic instrumental errors, we organized our observations as follows. Between exposures, the phase compensator was rotated in such a way that the right- and left-hand circularly polarized spectra were interchanged on the CCD array. Let $\Delta\lambda_{rl}^{(1)}$ be the difference between the positions of some line in the spectra with opposite polarizations measured for the initial position of the phase compensator and $\Delta\lambda_{rl}^{(2)}$ be the same difference measured after the compensator rotation. According to Johns-Krull *et al.* (1999), the quantity

$$\Delta\lambda_{rl} = \frac{\Delta\lambda_{rl}^{(2)} - \Delta\lambda_{rl}^{(1)}}{2} \qquad (2)$$

will then be free from systematic errors, with the principal error being the tilt of the spectrograph slit. The same quantity was calculated for all measurable lines. Thus, to measure the line shift needed to calculate $B_{||}$ from relation (1) requires two exposures of the star.

On January 24, 2002, in addition to FU Ori, we observed the magnetic star HD 30466 and 109 Tau, which was used as a zero-field standard. Table 1

Table 1. The observed stars

Object	V	n	t_e	S/N
109 Tau	4.9	3	20	300
FU Ori	10.0	6	240	100
HD 30466	7.3	1	24	200

gives basic information on our observations: the object's apparent magnitude V, the number of observed pairs of spectra n, the total exposure time for the star in minutes t_e, and the signal-to-noise ratio S/N per pixel for each of the spectra obtained.

RESULTS

To check the reliability of the procedure for measuring the magnetic field, apart from FU Ori, we observed the magnetic star HD 30466 on the same night. For this star, Babcock (1958) found $B_{||} = +2320 \pm 340$ G and $B_{||} = +1890 \pm 130$ G from two independent observations. As an illustration, Fig. 1 shows our observed profile of the Cr I 6661.08 line in two polarizations. We see that the profiles are significantly shifted relative to each other: $\Delta\lambda_{rl} \simeq 126$ mÅ, which, according to relation (1), corresponds to $B_{||} \simeq +2.0$ kG with $g = 1.48$ (Kupka *et al.* 1999).

Figure 2 shows the spectrum of FU Ori that is the sum of the left- and right-hand circularly polarized spectra taken in two exposure times. In addition to a strong Hα line with a P Cyg profile, the spectrum exhibits a number of absorption features, which we used along with Hα to measure the magnetic field. The spectrum of FU Ori near Hα is similar to the spectrum of a G2 Ib supergiant (Herbig 1966). Therefore, we identified these absorption features with Fe I, Ni I, and Ca I lines by using the VALD database (Kupka
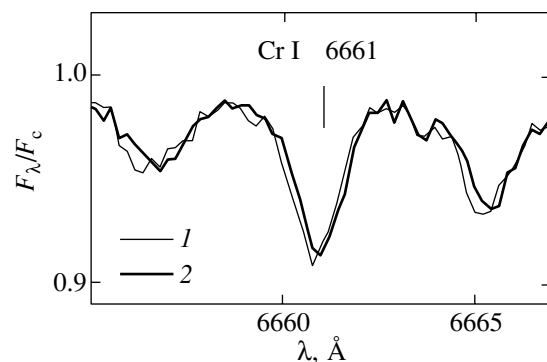

Fig. 1. A portion of the spectrum for the magnetic star HD 30466 in right-hand (*1*) and left-hand (*2*) circularly polarized light.

Table 2. Results of the observations

Species	λ, Å	g_{ef}	$\Delta\lambda_1$	$\Delta\lambda_2$	$\Delta\lambda_3$	$\Delta\lambda_4$	$\Delta\lambda_5$	$\Delta\lambda_6$
FeI	6481.87	1.14	−13.5	−27.0	+47.7	+9.4	−8.1	−
NiI	6482.80							
CaI	6493.78	0.96	−6.8	+21.5	+1.8	−5.2		13.9
FeI	6494.98							
FeI	6546.24	0.82	−5.5	+67.6	38.7	−14.5	−41.0	+3.8
FeI	6592.91	1.00	+11.8	−26.2	−0.5	−30.9	−5.4	+13.4
FeI	6633.75	1.66	+44.7	+64.9	+8.7	−14.0	+18.4	+11.6
NiI	6643.63	1.31	−97.0	−39.0	+22.9	+20.5	+16.7	+38.7
FeI	6663.23	1.53	−60.1	+31.4	+24.2	+61.2	+12.5	−43.5
FeI	6663.44							
FeI	6677.99	1.10	−	0.0	+28.7	−63.5	−47.3	−43.9
Hα^{abs}	6561.4	1.00	+13.5	−5.5	+0.2	−3.9	+0.5	+2.4
Hα^{em}	6564.3	1.00	+5.7	−9.4	−4.2	−3.2	+5.6	+3.4

et al. 1999) (see Table 2). When an absorption feature was a blend of lines of comparable intensities, we used the effective Lande factor

$$g_{ef} = \frac{\Sigma d_i g_i}{\Sigma d_i}; \qquad (3)$$

i.e., the weighted mean of the central depths d_i of the blend components taken from the VALD database. The values of g_{ef} are given in the third column of Table 2.

The measured values of $\Delta\lambda_{rl}$ for the above absorption features (in mÅ) are given in Table 2 for the six pairs of FU Ori spectra observed on January 24, 2002. In some cases, the line profiles were distorted by cosmic-ray traces and were unsuitable for measurements; such cases are marked by a dash in the table.

First, we estimated the magnetic field strength only from metal lines, i.e., without Hα. We calculated $B_{||}$ for each of these lines using formula (1) and then found the mean over all of the lines to be $+10 \pm 110$ G. To estimate the extent to which an uncertainty in the line identification affects the final result, we also calculated the mean shift $\Delta\lambda_{rl}$ from all of the metal lines in Table 2 and then calculated $B_{||}$ using formula (1) by assuming that $g = 1$. In this case, it was found to be $B_{||} = +30 \pm 130$ G. Thus, in both cases, $B_{||}$ in the metal line formation region does not exceed 350−400 G at a 3σ level.

As for the Hα line, its absorption and emission components are probably formed in different regions (Errico *et al.* 2003). Therefore, we measured $\Delta\lambda_{rl}$ for

these components separately; the parts of the profile used for our measurements are shown in Fig. 2 and the $\Delta\lambda_{rl}$ values obtained are given in the last two rows of Table 2. We obtained $B_{||} = +30 \pm 68$ G for the absorption component and $B_{||} = -10 \pm 32$ G for the emission component. Thus, at a 3σ level, $B_{||} \leq 200$ G and $B_{||} \leq 100$ G in the formation regions of the absorption and emission components, respectively.

To check our technique of allowing for systematic errors, we also observed the star 109 Tau (G8 III); since this star is a giant, it would be reasonable to assume that the strength of its magnetic field must be nearly zero. Having reduced the spectra of this star by using the same techniques as that for FU Ori, we found that the mean shift of the 18 lines used is $\simeq 0.8 \pm 0.4$ mÅ, which roughly corresponds to $B_{||} = +20 \pm 10$ G; i.e., the magnetic field of 109 Tau is actually very weak. Therefore, we believe that our method of eliminating systematic errors is highly efficient and our upper limit on $B_{||}$ for FU Ori is determined by random errors rather than by systematic ones.

DISCUSSION

As was noted above, the role of the magnetic field for FU Ori may be twofold: first, the magnetic field lines may be the guides along which the disk-wind matter moves and, second, the stellar magnetic field may be responsible for the appearance of an asymmetry in the structure of the disk and the wind.

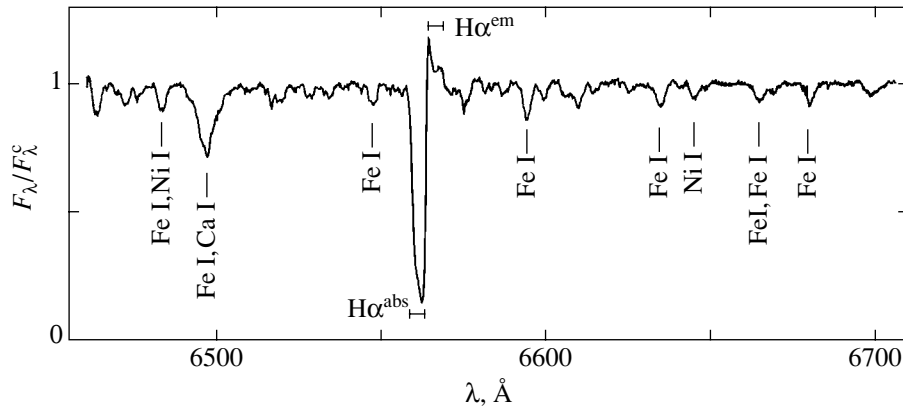Hartmann and Calvet (1995) showed that metal lines originate in the accretion disk of FU Ori and that

Fig. 2. The spectrum of FU Ori that is the sum of the right- and left-hand circularly polarized spectra taken with two exposure times.

the strongest lines also exhibit blueshifted components that originate in the wind (see also Petrov and Herbig 1992). Assume that the gas velocity and density in the wind are, respectively, 300 km s^{-1} and 3×10^{12} cm^{-3} (Calvet et al. 1993). It then follows from the relation $\rho_w V_w^2/2 \leq B^2/8\pi$ that a field strength of at least 300 G is required to control the wind motion. Our upper limits for the strength of the longitudinal field component obtained from metal lines ($B_{\parallel} \leq 350-400$ G) and from the absorption component of the Hα line ($B_{\parallel} \leq 200$ G) are consistent with this value.

According to Calvet et al. (1993), the streamlines of the disk wind from FU Ori are almost perpendicular to the plane of the disk whose rotation axis is inclined at an angle of $30° - 45°$ to the line of sight. Therefore, it follows from our constraints on $B_{\parallel}$ that the total field strength in the line formation region most likely does not exceed 300 G. Thus, if the magnetic field actually collimates the wind from FU Ori, then its strength does not significantly exceed the minimally required value.

Errico et al. (2003) estimated the magnetic field strength B in the formation region of the Hα emission component to be $\sim 10^3$ G, whereas our upper limit for $B_{\parallel}$ is only 100 G. It thus follows that either B was highly overestimated or the magnetic field lines were almost perpendicular to the line of sight at the time of our observations. Based on the available data, we cannot say which of these explanations is correct and restrict ourselves to the following remark.

In the opinion of Errico et al. (2003), the emission component of the Hα line originates in a disk region confined within a narrow range of azimuthal angles (E region), which displaces relative to the observer because of the disk rotation with a period of about 9 days. They assumed that the stellar magnetic field

lines within the E region were inclined at a small angle to the disk plane (see Fig. 5 in Errico et al. 2003). In this model, the $B_{\parallel}/B$ ratio, as well as the intensity of the Hα emission peak, must be highest when the E region crosses the plane that passes through the line of sight and the disk rotation axis. This means that a critical test for the model of Errico et al. would be an estimate of $B_{\parallel}$ near the time of maximum intensity of the Hα emission component. Meanwhile, it follows from a comparison of our Fig. 2 with Fig. 1 from Errico et al. (2003) that the intensity of the emission peak in the Hα line during our observations was more than a factor of 2 lower than its maximum value.

CONCLUSIONS

It follows from our observations that the component of the large-scale magnetic field along the line of sight $B_{\parallel}$ in the formation region of metal absorption lines (disk + wind) does not exceed $350-400$ G. We obtained a more stringent constraint from the absorption component of the Hα line formed in the disk wind: $B_{\parallel} < 200$ G. This probably implies that the strength of the magnetic field B collimating the disk wind does not exceed 300 G, in agreement with theoretical estimates based up the model of Calvet et al. (1993).

Our upper limit for the longitudinal magnetic field in the formation region of the Hα emission component ($B_{\parallel} < 100$ G) is an order of magnitude lower than the total field strength B predicted by Errico et al. (2003). It thus follows that either B was highly overestimated or the magnetic field lines were almost perpendicular to the line of sight at the time of our observations. A critical test for the model of Errico et al. would be an estimate of $B_{\parallel}$ near the time of maximum intensity of the Hα emission component.

ACKNOWLEDGMENTS

REFERENCES

1. H. W. Babcock, Astrophys. J., Suppl. Ser. **3**, 207 (1958).
2. N. Calvet, L. Hartmann, and S. J. Kenyon, Astrophys. J. **402**, 623 (1993).
3. G. A. Chuntonov, God. Otch. Spets. Astrofiz. Obs. **56**, 36 (1997).
4. G. D'Angelo, L. Errico, M. T. Gomez, *et al.*, Astron. Astrophys. **356**, 888 (2000).
5. L. Erriko, A. Vittone, and S. A. Lamzin, Pis'ma Astron. Zh. **29**, 125 (2003) [Astron. Lett. **29**, 105 (2003)].
6. L. Hartmann and N. Calvet, Astron. J. **109**, 1846 (1995).
7. L. Hartmann and S. J. Kenyon, Ann. Rev. Astron. Astrophys. **34**, 207 (1996).
8. G. H. Herbig, Vistas Astron. **8**, 109 (1966).
9. C. M. Johns-Krull, J. A. Valenti, A. P. Hatzes, and A. Kanaan, Astrophys. J. **510**, L41 (1999).
10. V. G. Klochkova, S. V. Ermakov, and V. E. Panchuk, Bull. Spets. Astrofiz. Obs. **41**, 58 (1996).
11. F. Kupka, N. Piskunov, T. A. Ryabchikova, *et al.*, Astron. Astrophys., Suppl. Ser. **138**, 119 (1999).
12. V. E. Panchuk, Preprint No. 154 (Special Astrophysical Observatory, Russian Academy of Sciences, Nizhnii Arkhyz, 2001).
13. P. P. Petrov and G. H. Herbig, Astrophys. J. **392**, 209 (1992).

Translated by N. Samus'

Peculiarities of the Microwave Emission from Active Regions Generating Intense Solar Flares

V. M. Bogod[*] and S. Kh. Tokhchukova

Special Astrophysical Observatory, Russian Academy of Sciences,
Nizhnii Arkhyz, Stavropolskii krai, 357147 Russia
Received August 28, 2002

Abstract—RATAN-600 multiwavelength observations of the Sun reveal sharp spectral inhomogeneities in the polarized radiation from active regions that produce intense flares. These events occur in a wide range of radio fluxes (0.05–10 s.f.u.) in a relatively narrow wavelength range (2–5 cm). They are detected on time scales from several hours to several days before and during an intense flare. We analyze the detected events and their relationship to the preliminary phase of intense flares. Significant statistical material was obtained in 2001. The new flare-plasma properties can be used to test existing solar-flare models and to develop new criteria of flaring activity. © *2003 MAIK "Nauka/Interperiodica".*

Key words: *Sun, solar radio emission, polarization, solar flares.*

INTRODUCTION

Studying the preflare plasma is important both in understanding the processes that lead to a flare and in developing criteria for predicting the flaring activity and, in particular, the proton activity. Such studies are being carried out in a wide frequency range, including ultraviolet, X-ray, optical, and radio satellite observations. Since flares most often originate in the upper chromosphere and the lower corona and subsequently extend to higher levels (into the X-ray corona) and to lower-lying levels (into the optical photosphere), it should be assumed that the main preflare-plasma properties must clearly show up precisely in the radio frequency range.

Varied and long-term studies using various radio instruments revealed a number of preflare-plasma properties. The criterion of Tanaka and Enome (1975), which is used to compare polarization measurements at two wavelengths (3 and 8 cm), is widely known. The reliability of this criterion was high (∼70%) in solar cycle 20 and decreased sharply in the next cycle.

We know the flare precursors: small increases in the total solar flux 20–40 min before a flare (Pustil'nik 1973), increases in the total radio flux fluctuation 20–40 min before a flare (Kobrin *et al.* 1978), and decreases in the radio flux before a flare (Covington 1973). Recently, Kundu *et al.* (2001) have reported radio fluctuations several minutes before a flare, which correlated with X-ray fluctuations.

High-spatial-resolution observations revealed a prolonged microwave darkening of an active region (AR) before a flare (Bogod *et al.* 1999a), the emergence of peculiar sources (Lang *et al.* 1993) and sources above the neutral magnetic field line (Kundu *et al.* 1985), and the existence of active longitudes (Maksimov *et al.* 1988).

Since the sensitivity of radio waves to variations in the chromospheric magnetic fields is high (Bogod *et al.* 1993; Nindos *et al.* 2000), various methods of their measurement were developed (Gelfreikh 1994; Ryabov *et al.* 1999) based on detailed spectral and polarization analyses in a wide wavelength range. Although there are various radio telescopes with a high spatial resolution in the world (from interferometric radioheliographs to reflective antennae), the spectrum is analyzed in detail only with two telescopes, RATAN-600 (Korol'kov and Parijskij 1979) and OVRO (Hurford *et al.* 1984), which completely cover the frequency range 1–18 GHz. The latter analyzes this frequency range sequentially, passing from one frequency to another, and has narrow pass bands, which significantly reduces the polarization measurement sensitivity. RATAN-600 has broader pass bands in the frequency channels, performs a simultaneous analysis with a frequency resolution of about 5%, and has a large effective area. However, this instrument has a one-dimensional beam, which reduces its efficiency when studying weak sources because of the influence of spurious sources (confusion effect). However, when studying bright ARs, this effect is negligible. We have been using the mode of multiple azimuthal observations (with a sampling interval
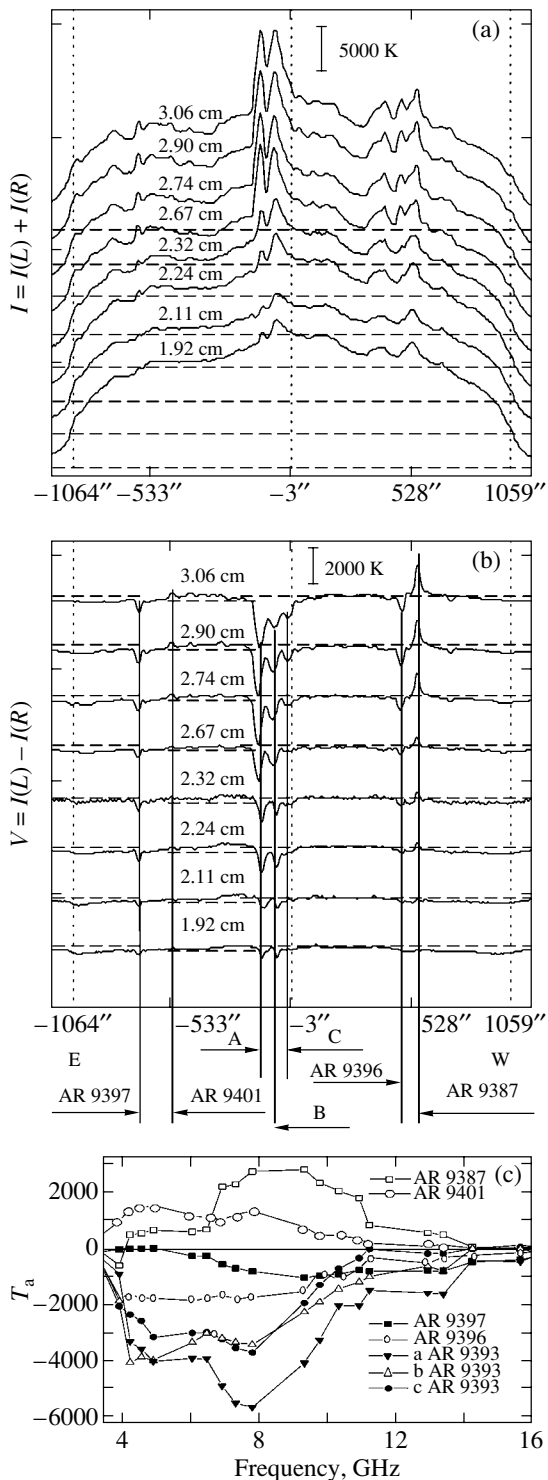
---

[*]E-mail: `vbog@gao.spb.ru`

Fig. 1. The solar-disk scans for March 27, 2001, obtained with the RATAN-600 one-dimensional beam at several close centimeter wavelengths: (a) the intensity scans vertically displaced from each other for the convenience of comparison, (b) the vertically displaced circular polarization scans (in this figure and below, the signs of the left- and right-hand circular polarizations are positive and negative, respectively, and (c) the maximum-amplitude spectra of polarized radio sources for the ARs shown in the upper Figs. 1a and 1b.

of 4−8 min/scan) on RATAN-600 since 2000. This mode significantly improved the reliability of spectral measurements of the polarized radiation from ARs and allowed a number of events directly related to flaring activity to be detected in a wide time interval.

Observations reveal ARs on the Sun that differ in stability. Stable ARs produce virtually no flares over the entire period of their presence on the disk. These regions have a stable magnetic-field structure at the photospheric level and monotonic radio intensity and polarization spectra in the microwave range with a peak near 5−6 cm (Bogod et al. 1999b; Gary et al. 1997) (see also Fig. 1). The behavior of these spectra is determined by the combined action of thermal mechanisms (cyclotron bremsstrahlung at the first three or four gyrofrequency harmonics (Zheleznyakov 1977; Akhmedov et al. 1982)). The cyclotron radiation from an AR has a circular polarization that corresponds to an excess of the extraordinary-wave radiation; the polarization sign is determined by the magnetic-field polarity of the dominant sunspot.

Of particular interest are flare-productive active regions (FPARs), which are distinguished by flux variability and high flaring activity. Our study of FPARs indicates that microwave spectral-polarization observations are highly informative; they appear to directly detect plasma at the sites of primary energy release before a flare. Several new manifestations were found in the radio polarization spectrum of FPARs on time scales from several hours to several days before and during a flare. We also show that these new effects are associated with the generation of intense flares and that they can be used to test existing solar-flare models (see Syrovatskii et al. (1983) and the reviews of Altyntsev et al. (1982) and Shibata (1998)) and develop more efficient prediction criteria than the existing ones.

OBSERVATIONS

For our analysis, we used the regular spectral-polarization observations of the Sun carried out in 2001 with the RATAN-600 radio telescope. This telescope has parameters that are currently most suitable for studying flare plasma at early stages of FPAR activity. The most important quality of this instrument was the combination of an instantaneous frequency spectrum from 16 to 1 GHz taken with a panoramic spectrum analyzer (PSA) (Bogod et al. 1999b) (in this work, we used the frequency range 4−16 GHz) with a frequency resolution of about 5−7%, a high sensitivity to variations in circular polarization (about 0.05%), and a high flux sensitivity reaching 0.001 s.f.u. (1 s.f.u. = 10^{-22} W m^{-2} Hz^{-1}). Such a high sensitivity was achieved due to a broad band of

received frequencies (200–600 MHz), a simultaneous analysis of the spectrum, and a large collecting surface area ($\sim$400–600 m^2). The RATAN-600 radio telescope has a moderate spatial resolution in the horizontal plane ($<15''$ at a wavelength of 2 cm) and a low resolution in the vertical plane ($\sim15'$ at 2 cm). The beam size is proportional to the wavelength. Our observations were performed regularly both in single-scanning mode (three to five observations per day) and in multiple azimuthal scanning mode (23–25 scans per day at 8-min intervals for 4 h). Such frequently repeated scans with a narrow beam for 4 h allow the radiation from individual components of the AR structure (local sources above sunspots, interspot sources, flocculi, and others) to be clearly separated.

For our analysis, we used the observational data obtained in the regular RATAN-600 observations from January 4 through October 28, 2001, except for small technical interruptions from April 16 until April 24 and from August 15 until August 25. Nevertheless, almost all of the FPARs present on the disk for ten months were recorded in their spectral observations in intensity (the Stokes parameter $I = I(L) + I(R)$) and polarization ($V = I(L) - I(R)$). In addition, we carried out multiazimuthal observations, which allowed the dynamic properties of FPARs to be studied at microwaves.

Figure 1a shows one-dimensional RATAN-600 scans of the Sun. For the convenience of comparison, the scans at different wavelengths were vertically displaced from one another. The observations were reduced with the Workscan processing program (Garaimov 1997), which is freely available at http://www.sao.ru/~sun. Observations of the Toyokawa flux service at frequencies of 1.0, 2.0, 3.75, and 9.4 GHz, which are published on the Internet at ftp://solar.nro.nao.ac.jp/ pub/norp/data/daily/, were used to calibrate and reference the scans at different wavelengths. The flux at intermediate PSA frequencies is determined by a cubic interpolation and is used to calibrate the Stokes parameter $I = I(L) + I(R)$. The Stokes parameter $V = I(L) - I(R)$ is calibrated proportionally to the original signal.

Short-Wavelength Polarization Inversion in FPARs

Figures 1a and 1b show one of the multiwavelength observations performed on March 27 at 8^h50^m UT. On this day, all of the ARs detected on the solar disk were stable and their intensity (Fig. 1a) and circular polarization (Fig. 1b) spectra monotonically rose with the wavelength. This behavior is typical of the radiation from stable cyclotron radio sources. The polarization spectra for some of the ARs present
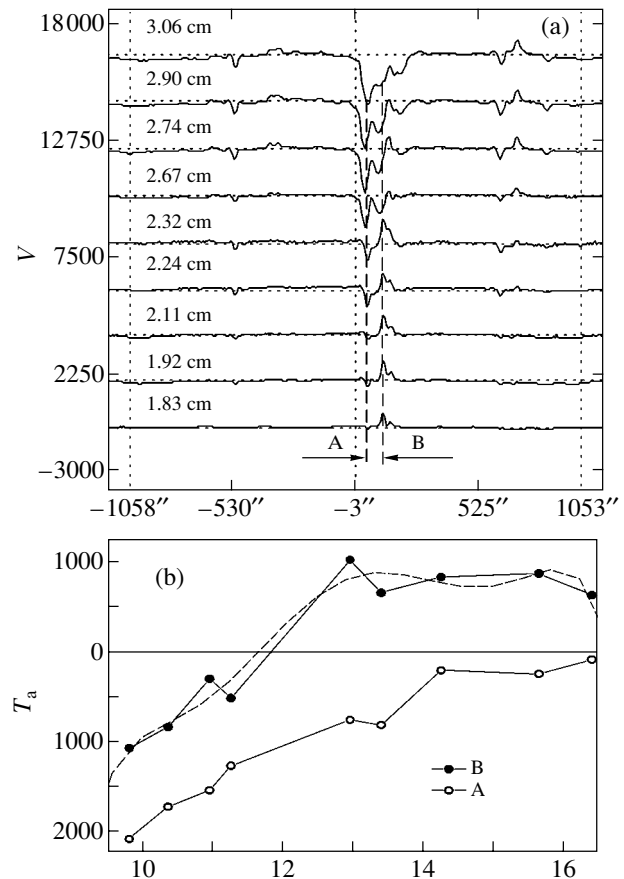
Fig. 2. (a) Circular polarization scans for the Sun at short centimeter wavelengths on March 28, 2001, 8^h46^m UT. Source B of left-hand circular polarization at wavelengths from 1.83 to 2.32 cm exhibits a rise in the western part of AR 9393, which caused a short-wavelength polarization inversion. (b) The spectra of the eastern (A) and western (B) sources in AR 9393. Source B exhibits a sign inversion at a frequency close to 12 GHz. The dotted line represents a power-law polynomial fit. Source A shows a monotonic rise in right-hand circular polarization with wavelength.

on the disk on this day are shown in Fig. 1c. They all exhibit a rise with the wavelength at frequencies above 15 GHz and a broad peak in the frequency range 6–8 GHz. The AR NOAA 9393 (see Figs. 1a and 1b) was near the disk center. It had a complex structure that consisted of three sources of right-hand circular polarization (A, B, and C) and was flare-active, although the radiation from this region was stable during March 27, 2001. Subsequently, this AR was activated sharply, which gave rise to spectral inhomogeneities in the polarized radiation.

Preflare processes began in AR 9393 on the next day. They were accompanied by a circular polarization inversion at short wavelengths (1.83–2.32 cm) in its western part (see Fig. 2a). This figure shows, on a large scale, the multiwavelength one-dimensional
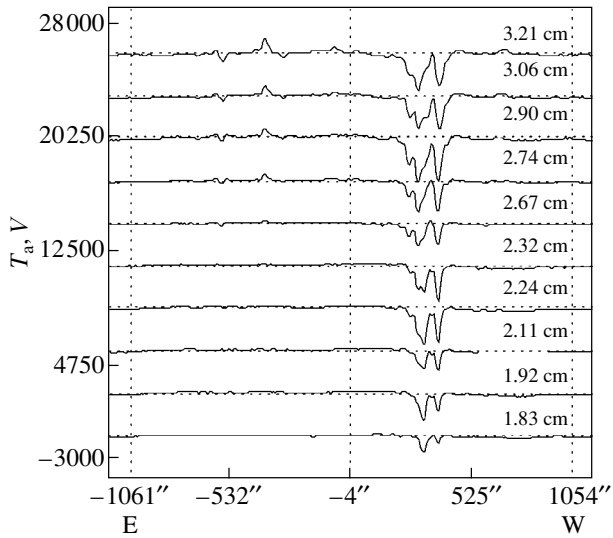
Fig. 3. Polarization scans (in antenna temperatures) for March 30, 2001, 10^h14^m UT. AR 9393 lies in the western part of the solar disk and again has a stable spectrum.

scan for the Sun in polarized radiation obtained on March 27, 2001, at 8^h46^m UT.

On this day, AR 9393 was located near the disk center and significant changes occurred in its structure. The two main sources, B in the western part of the AR and A in its eastern part, are marked in Fig. 2a. A polarization inversion was detected in the radiation from source B at wavelengths from 1.83 to 2.32 cm, whereas the polarization pattern and sign of source A did not change (see also Figs. 1b and 2b). This short-wavelength circular polarization inversion (which we called an A1 effect) persisted until the onset of the impulsive phase of an intense X1.7 flare at 10^h15^m on March 29, 2001.

We believe that the circular polarization inversion points to a rise in the new magnetic flux whose polarity is opposite to the polarity of the overlying magnetic field with a strength of about 1400–2000 G under the assumption of cyclotron radiation at the third gyrofrequency harmonic in accordance with the formula $B[G_S] = 3570/\lambda$ [cm]. In this case, the polarization sign is determined by the radiation of an extraordinary wave and the magnetic-field polarity. In general, the short-wavelength polarization inversion first appears at the shortest wavelength of 1.74 cm and then gradually displaces with time toward longer wavelengths, 3–4 cm. This type of polarization inversion should be distinguished from the well-known passage of the cyclotron radiation from a spotted radio source through a region of quasi-transverse magnetic fields (Cohen 1960; Zheleznyakov 1964; Ryabov *et al.* 1999). The inversion in quasi-transverse magnetic fields is known from observations to commonly begin at long
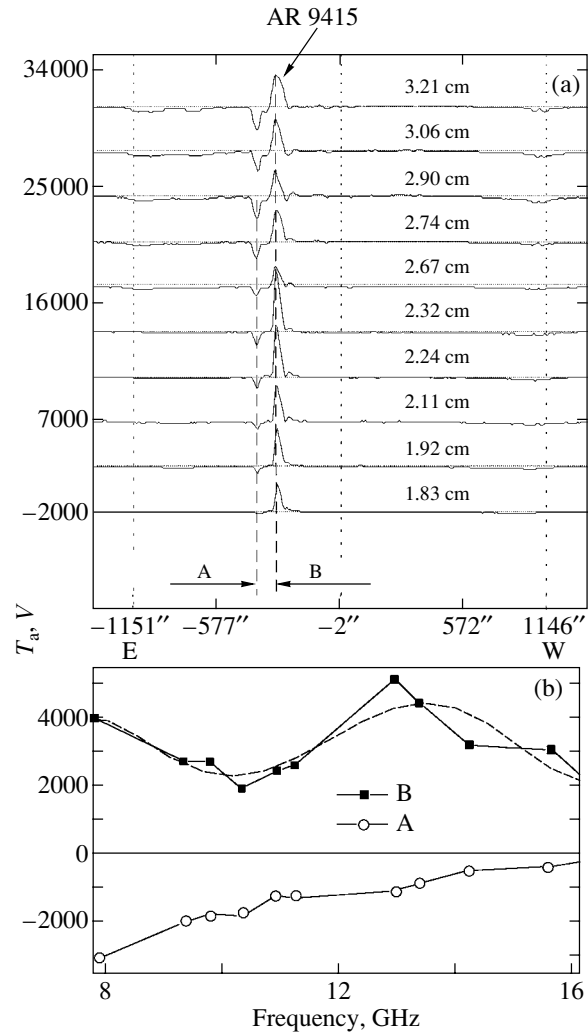

Fig. 4. (a) Polarization scans for the Sun (in antenna temperatures) on April 8, 2001, 9^h15^m UT. AR 9415 lies in the eastern part of the solar disk and has a bipolar polarization pattern in the from of sources A and B. Source B exhibits a rise in intensity at short wavelengths, while source A shows no such dependence. (b) The spectra of the western (B) and eastern (A) sources in AR 9415. Source A shows a monotonic rise in left-hand circular polarization; source B of left-hand polarization exhibits a rise of the spectrum at short wavelengths. The dotted line represents a power-law polynomial fit.

wavelengths (in the range 8–9 cm) in stable ARs near the limb and displace with time toward shorter wavelengths. On the next day, in the observations of March 30, 2001, a stable structure (see Fig. 3) similar to the structure (see Fig. 1) observed on March 27, 2001, was restored in AR 9393.

Short-Wavelength Polarization Brightening in FPARs

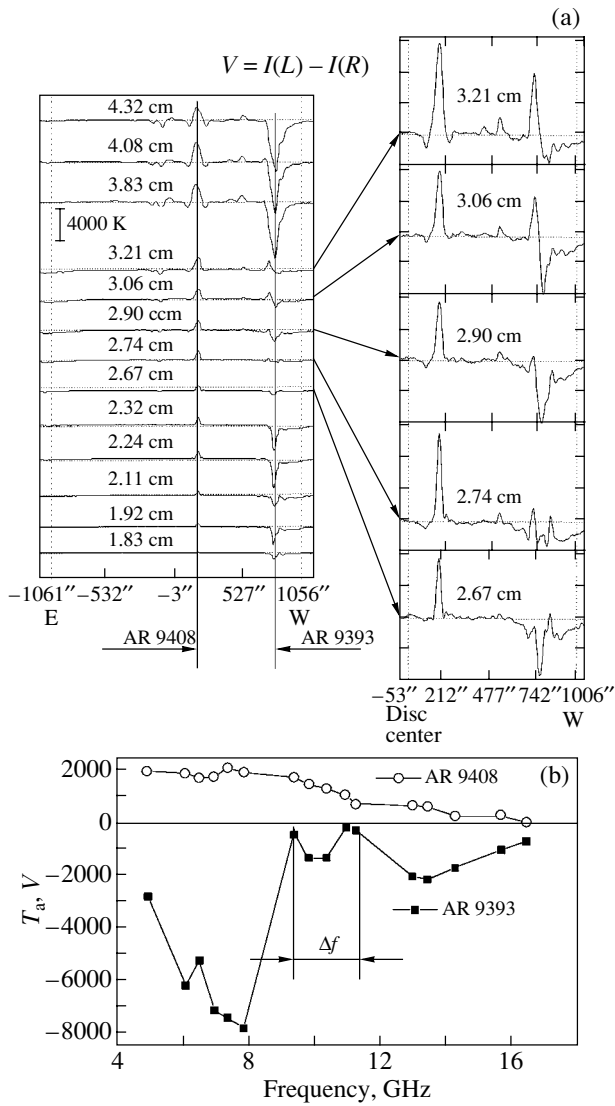Another commonly encountered effect is a brightening at short centimeter wavelengths. In Fig. 4a,

Fig. 5. (a) Polarization scans for the Sun on April 1, 2001, 9^h17^m UT. AR 9393 lies in the western part of the solar disk and has a simple polarization pattern at short (2.32 cm or shorter) and long (3.83 cm or longer) wavelengths. The scans in the wavelength range from 2.67 to 3.21 cm are shown on the right on a large scale. This range exhibits a complex pattern of mode coupling with the appearance of narrow pointlike sources, with sharp spectral slopes, polarization inversions, and a low degree of polarization. (b) The circular polarization spectra for the stable AR 9408 and FPAR 9393. Here, Δf is the width of the frequency band in which abrupt variations occur at a low level of the polarized signal.

Fig. 6. Multiwavelength scans in the polarization channels at 9^h08^m UT. The arrows on the right indicate the locations of the frequency polarization inversions that occurred in AR 9393 before an M8.4 flare.

magnetic field. The spectrum exhibits a rise toward shorter wavelengths (see source B in Fig. 4b).

The A1 and A2 effects persist from several hours to several days. The maximum frequency band in which the two effects occur is broad and reaches one octave (from 1.7 to 3.5 cm). The polarized fluxes typically lie within the range 0.05−0.5 s.f.u.; the A effects are commonly observed in FPARs (in 60% of the cases).

A Frequency-Limited Mode Coupling Region in the Radio Polarization Spectrum

The next manifestation of FPARs, called a B effect, is usually recorded in well-developed ARs with prolonged flaring activity. This effect often results from the development of the A1 effect with time. In this case, the polarization inversion displaces toward

this effect is illustrated for AR 9415: an increase in the circularly polarized flux is observed at short wavelengths for the western source of this group. We call this case an A2 effect. We believe that this effect is related to the rise of a new magnetic flux whose polarity matches the polarity of the overlying
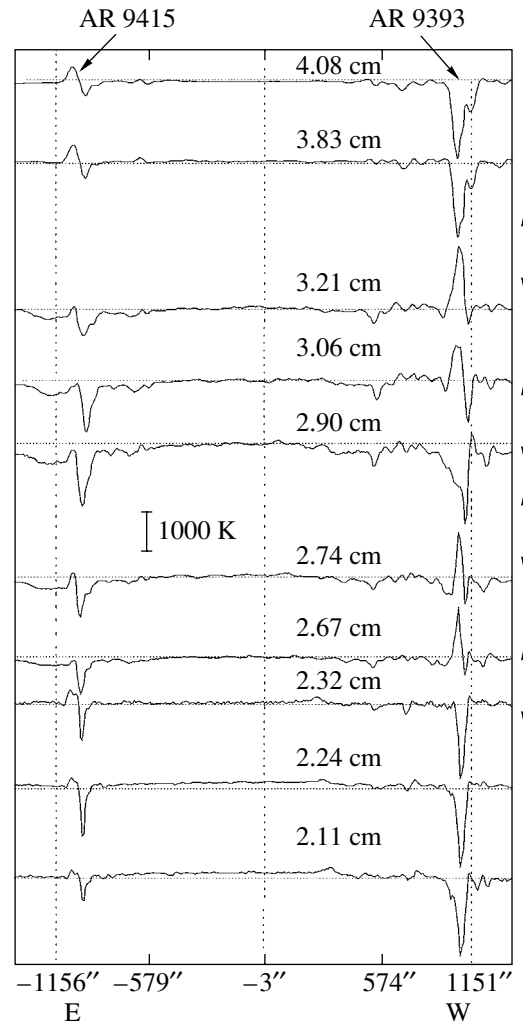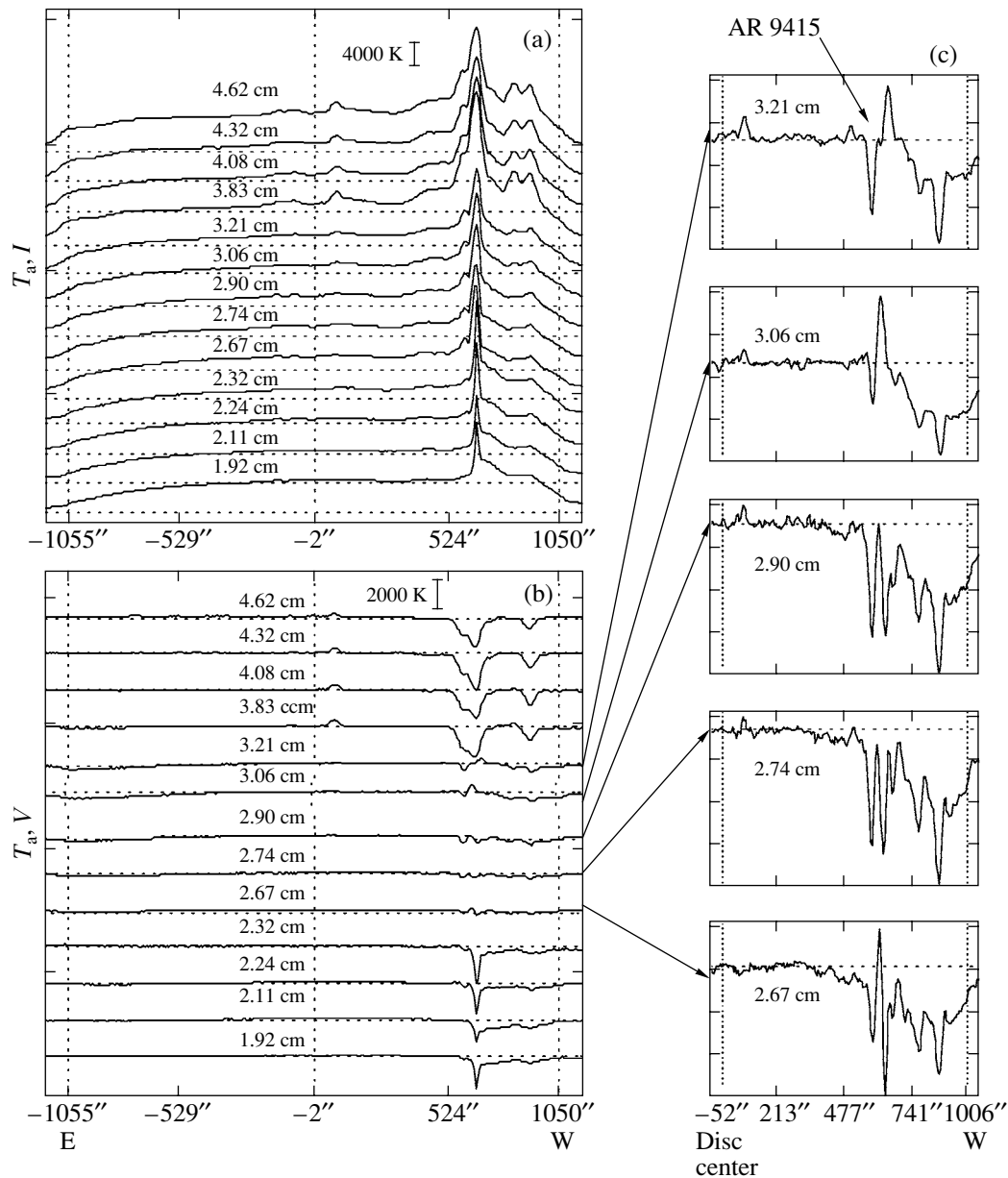
Fig. 7. The observations of April 14, 2001. The FPAR 9415 lies in the western part of the solar disk. (a) Intensity scans of the Sun; (b) circular polarization scans; and (c) scans of the western part of the solar disk on a large scale in the wavelength range 2.67−3.21 cm. The polarization pattern of AR 9415 in this range differs sharply from its pattern at short and long wavelengths, forming the so-called frequency mode coupling region.

longer wavelengths, while at short wavelengths, polarized radiation with the original polarization sign appears. The region of the B effect is characterized by a decrease of the polarized flux in a narrow frequency band (20−30%) in the wavelength range 2−5 cm. Accordingly, the degree of polarization decreases from 10−15% outside the band to fractions of a percent inside the band. In addition, narrow pointlike sources with sharp spectral slopes and multiple polarization frequency inversions appear in this region. Figure 5 illustrates the formation of the B effect at wavelengths

2.67−3.21 cm for AR 9393. We see the appearance of new (pointlike) sources with multiple polarization inversions within a limited band at neighboring wavelengths. This band often displaces toward longer wavelengths immediately before the flare itself. In this case, an intense X20 flare occurred on April 2 at 3^h25^m UT.

We believe that the region of the B effect reflects the processes that take place in the region of coupling between the ordinary and extraordinary modes,
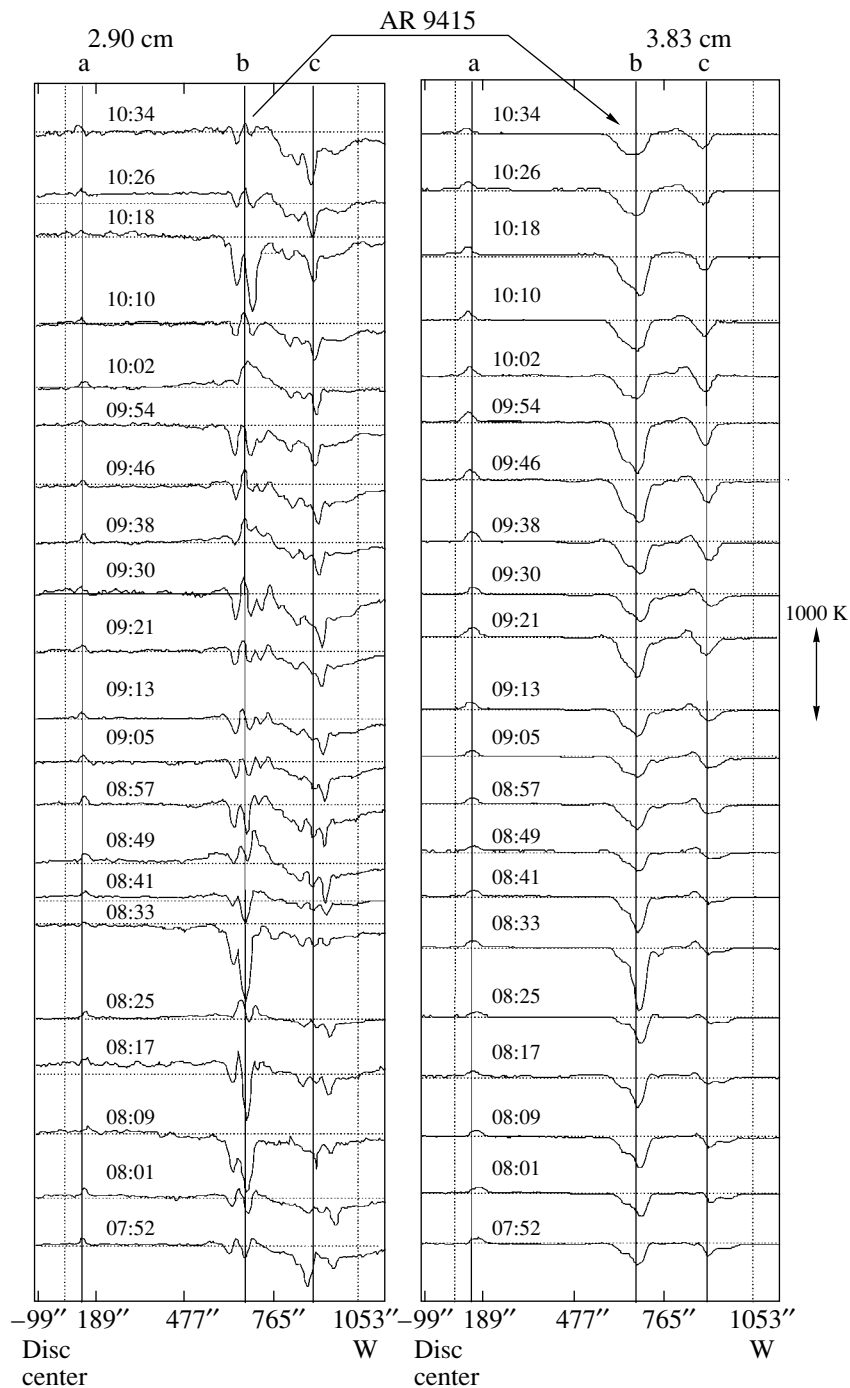
Fig. 8. An example of the dynamic polarization behavior for the FPAR 9415 at two wavelengths inside and outside the mode coupling region. The observations were carried out from 7^h52^m until 10^h34^m UT with an 8-min interval between the scans.

because polarization inversions in frequency and time are commonly observed here (see also Fig. 6).

Multiple Spectral and Temporal Circular Polarization Reversals in FPARs

Our study of several FPARs indicates that a polarization inversion is often present in their microwave radiation and is probably a characteristic property of ARs that should be investigated further. A polarization inversion commonly appears in the polarization spectrum before an intense flare and can also occur during the flare itself. Figure 6 shows the observations of AR 9393 carried out on April 5, 2001, at 9^h08^m UT immediately before the onset of the main phase of an

M8.4 limb flare at 9^h22^m UT. We see four polarization inversions in frequency marked by the arrows on the right.

Our multiazimuthal observations also revealed fast temporal variations of the polarization pattern in the wavelength range from 2 to 8 cm with an 8-min interval between the scans for 4 h. Figure 7 shows the observations of the FPAR 9415 on April 14, 2001, at 9^h54^m UT in the intensity and polarization channels at 13 wavelengths. We see that the degree of polarization decreased sharply in the wavelength range 2.67−3.21 cm. Therefore, this range is shown in Fig. 7c on a large scale. Interestingly, pointlike polarized radio sources with sharp variations over the spectrum up to the sign reversal appear in this range.

To analyze the temporal variations in the parameters of FPAR 9415, we selected two wavelengths shown in Fig. 8. One ($\lambda_1 = 2.90$ cm) was taken from the mode coupling region and the other ($\lambda_2 = 3.83$ cm) was taken from the long-wavelength range outside it. The figure shows 22 scans for the two wavelengths with an 8-min interval. For comparison with other ARs, we drew the vertical lines. The structure of AR 9415 at $\lambda_1 = 2.90$ cm consists of various narrow polarized radio sources with frequent polarization inversions in time. On the other hand, only smooth variations in the amplitude of the polarized signal occur at $\lambda_2 = 3.83$ cm.

THE RESULTS OF A STATISTICAL ANALYSIS

We performed a statistical analysis of the correlation of FPARs with detected features in the microwave radiation with the productivity of intense flares. For our analysis, we chose the period from January through October 2001. In this year, three observations were carried out daily with RATAN-600 at various azimuths. The flaring activity in the period being analyzed is presented in Table 1. The table lists all of the FPARs that produced the most intense flares. Columns 3 and 4 characterize the X-ray activity; columns 5 and 6 characterize the optical activity.

The intense FPARs listed in Table 1 exhibited the features in the polarization spectrum described above. We experimentally established that all of the ARs that produced at least three intense flares of importance M or higher could be placed into this group. Only two regions, AR 9488 and AR 9636, which produced only two intense flares each, constitute an exception. Another exception is AR 9658, which produced 16 X-ray flares, four of which are of importance M, but we found no features in the spectrum of its circular polarization.

The results of our statistical analysis are presented in Table 2. We see from this table that almost all of the

activity from January 4, 2001, through October 28, 2001, was attributable to the activity of the FPARs that accounted for only 5.9% of the total number of ARs regions on the Sun. The FPARs produced 83.3% of the intense flares, which usually generate high-velocity protons. Only one region of the 24 FPARs exhibited no features in the polarization spectrum. Such a high correlation coefficient may indicate a high prognostic value of the detected effects.

DISCUSSION

This paper emphasizes the importance of radio observations in studying the nature of the powerful flaring activity, because they are highly sensitive to polarization measurement, which is the main subject of discussion.

ARs should be distinguished by the stability of their radiation: stable, quasi-stable, and FPAR. All three types of ARs differ by the height distribution of plasma parameters.

Stable ARs produce no flares at all and they account for most of the ARs. These ARs have stable magnetospheres and the magnetic-field structure can be calculated in the potential approximation from the photospheric level into the corona. There is no substantial energy release in these ARs and the following condition is dominant: $E \leq H^2/8\pi$, where $E = n_e \kappa T_e$ is the kinetic electron energy and H is the magnetic field.

Quasi-stable ARs produce several low- and medium-intensity flares (of X-ray importance A, B, or C) and, occasionally, one or two M flares. In general, they retain stable magnetospheres, but conditions with short-duration energy release in the form of intermediate-level flares can be created in them.

FPARs usually generate at least three intense flares of importance M or higher and many A, B, and C flares. FPARs are fundamentally unstable ARs. Energy is continuously released and accumulated in these regions and the condition $E \geq H^2/8\pi$ is satisfied here. The magnetospheres of these ARs are unstable, have no ordered structure, and cannot be calculated as an extrapolation of the potential magnetic field from the photosphere into the corona. This view explains the short (20−30 min) preflare phase in quasi-stable ARs (Kundu *et al.* 2001) and the long (from several hours to several days) preliminary phase in FPARs. The abrupt polarization inversions (see Figs. 7 and 8) over a long period suggest a continuously lasting flare process with a weak energy release. This process accumulates energy in the AR magnetosphere and, subsequently, the mechanism of energy release is triggered.

Table 1. A list of FPARs from January through October 2001

NOAA AR number	Number of M and X X-ray flares	Total number of X-ray flares	Number of optical S-level flares	Number of optical flares with level >S
9313	4M	10	15	4
9368	3M	10	11	4
9373	4M	19	36	1
9393	24M, 3X	55	12	3
9401	5M	12	29	5
9415	6M, 5X	26	34	9
9433	15M	60	83	4
9455	5M	25	28	5
9488	2M	13	18	5
9502	3M	6	7	3
9511	4M, 1X	18	24	5
9557	3M	26	36	2
9591	9M, 1X	59	80	8
9600	3M	11	8	4
9601	10M	33	56	9
9608	12M	38	37	10
9616	5M	14	23	6
9628	7M	31	81	5
9632	2M, 1X	12	34	1
9636	2M	11	36	2
9658	4M	16	30	3
9661	1M, 2X	19	37	5
9672	4M, 2X	22	28	5
9682	9M	40	61	1

Table 2. The results of our statistical analysis for the period from January through October 2001

Number of all solar ARs from January 4, 2001, through October 28, 2001	406	100%
Number of FPARs that produced three or more M and X X-ray flares	24	5.9%
Number of all intense M and X X-ray flares	192	100%
Number of intense flares from FPARs	160	83.3%
Number of intense flares from FPARs with effects in microwave polarization spectra	156	81.2%

Our observational data provide a new insight into the phenomenon of intense solar flares as processes with a long preliminary phase. At this phase, weak energy release and accumulation take place in the AR magnetosphere. This process continues until the instability level is reached. The subsequent scenario for the main flare phase was well developed (Petchek 1964; Heyvaerts *et al.* 1977; Shibata 1998; Zaitsev and Stepanov 1983). According to the RATAN-600 data in the wavelength range 2–
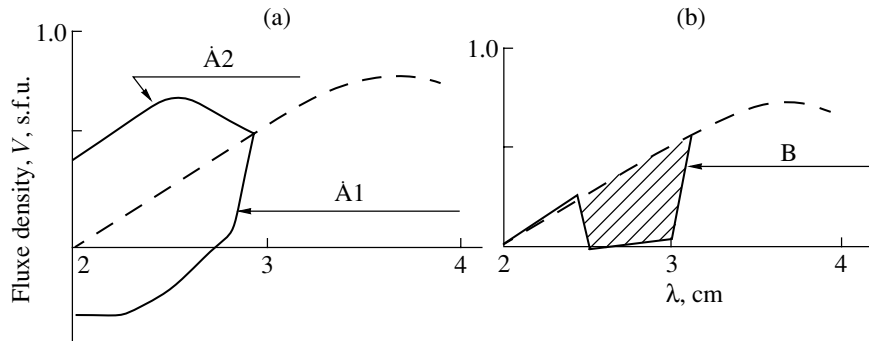
Fig. 9. (a) Comparison of the spectrum for a stable AR with a dominant cyclotron radio source and the spectrum for a FPAR with the A1 and A2 effects. (b) The same in comparison with the B effect.

5 cm, the preliminary phase can have several varieties in the polarization spectra of FPARs.

We treat two of them as the appearance of a new magnetic flux in the structure of the old overlying AR magnetic field. The polarity of the new magnetic flux can be identical to the dominant magnetic field polarity and, in this case, the polarization spectra exhibit a brightening toward shorter wavelengths (Fig. 4b). When a new magnetic flux with the polarity opposite to that of the old magnetic field rises, a short-wavelength polarization inversion is observed (Fig. 2b). These two effects persist in FPARs for several hours and days and, thus, differ from the so-called preflare phase. The latter generally lasts several minutes or, rarely, several tens of minutes (Kundu *et al.* 2001). The magnetic loops of the new magnetic flux appear to interact with the loops of the old magnetic field through reconnection according to the scenario of Hanaoka (1996). However, here, this process acts as a preheating. As a result, energy is accumulated in the AR magnetosphere to a level at which the magnetic structure becomes unstable, which triggers intense energy release and coronal mass ejection (see the review of Shibata 1998).

Another manifestation of the long preliminary phase found in radio observations is the formation of a frequency-limited region of coupling between the ordinary and extraordinary modes. Abrupt variations in the polarization pattern in the frequency band 20–25%, in which the degree of polarization decreases to fractions of a percent and in which the formation of pointlike polarized radio sources with sign reversals is observed (see Figs. 5a, 5b, 7, 9a), justify this name. Here, we probably also deal with weak energy release in higher layers of the lower corona. The multiple polarization inversions, in both frequency and time, indicate that the preliminary phase can arise in the form of prolonged low-energy miniflares. Such prolonged miniflares also result in the accumulation of energy in the magnetosphere, the attainment of

the instability threshold, and the turn-on of the main phase of an intense flare.

Based on the data obtained during 2001, we constructed generic spectra for the A and B effects (see Fig. 9).

What are the mechanisms of the detected spectral-polarization properties of the microwave radio emission (B effect) at the preliminary phase of intense flares? The discussion of this question is beyond the scope of this paper. Here, we only list the basic properties of this effect:

(1) The frequency inversions occur in the wavelength range 2–5 cm, suggesting a connection to the magnetobremsstrahlung mechanism at the first three or four gyrofrequency harmonics.

(2) The polarized flux in the mode coupling region decreases sharply to 0.05–10 s.f.u and the degree of circular polarization falls to 0.3–1%. Explanations of the presence of current sheets in the FPAR magnetosphere should probably be invoked.

(3) The polarization inversions at the preliminary flare phase occur at intervals of 8 min or, possibly, shorter.

(4) Multiple polarization inversions in frequency occur, suggesting a strong vertical magnetic field nonuniformity at the preliminary phase and during flares.

Magnetic reconnection is the most plausible interpretation. On the other hand, we cannot rule out the possibility that the frequency region results from the turbulence of a multiarch magnetic structure or many current sheets on the pathway of radio waves to the Earth (Zheleznyakov *et al.* 1996). The model of Zlotnik (2001), which explains the polarization inversion as resulting from the appearance of an ordinary mode in the cold plasma of a prominence located above a hot radio source, should also be mentioned. However, this model disregards the turbulence of cold and hot plasmas.

CONCLUSIONS

(1) Based on RATAN-600 spectral-polarization observations, we obtained a new observational result concerning the properties of the magnetospheres of FPARs.

(2) We obtained data on the existence of a long preliminary phase during the generation of intense flares, which manifests itself in detailed polarization spectra over a wide microwave range.

(3) We introduced the concept of FPAR with prolonged energy release at the preliminary phase, energy accumulation, and explosive energy release during the main phase.

(4) We provided statistical data that shows a special role of FPARs in general solar activity.

(5) We showed the necessity of using large instruments in investigating FPARs when both studying the physical processes and developing prediction criteria for intense flares.

Because of their high sensitivity to the processes in the lower corona, where primary energy release takes place, spectral-polarization radio observations can form the basis for developing techniques for predicting intense flare events, which are often accompanied by the release of a large number of high-energy protons.

Thus, spectral radio observations are important in understanding the physical processes in FPARs and for applied problems, for example, those associated with the Cosmic Weather Program.

REFERENCES

1. Sh. B. Akhmedov, G. B. Gelfreikh, V. M. Bogod, and A. N. Korzhavin, Sol. Phys. **79**, 41 (1982).
2. A. T. Altyntsev, V. G. Banin, G. V. Kuklin, and V. M. Tomozov, *Solar Flares* (Nauka, Moscow, 1982).
3. V. M. Bogod, A. Grebinskij, V. Garaimov, *et al.*, Bull. Spec. Astrophys. Obs. **47**, 5 (1999a).
4. V. M. Bogod, V. I. Garaimov, N. P. Komar, and A. N. Korzhavin, ESA Spec. Publ. **448**, 1253 (1999b).
5. V. M. Bogod, S. M. Vatrushin, V. E. Abramov-Maximov, *et al.*, Astron. Soc. Pac. Conf. Ser. **46**, 306 (1993).
6. M. H. Cohen, Astrophys. J. **131**, 664 (1960).
7. A. E. Covington, Sol. Phys. **33**, 439 (1973).
8. V. I. Garaimov, Preprint Spec. Astrophys. Obs. Russ. Acad. Sci. **27**, 1 (1997).
9. D. E. Gary, M. D. Hartl, and T. Shimuzu, Astrophys. J. **477**, 958 (1997).
10. G. B. Gelfreikh, in *Solar Coronal Structures. Proceedings of the 144th Colloquium of the IAU, Tatranska Lomnica, Slovakia, 1993,* Ed. by V. Rusin, P. Heinzel, and J.-C. Vial (VEDA Publishing House of the Slovak Academy of Sciences, Tatranska Lomnica, 1994), p. 21.
11. Y. Hanaoka, Sol. Phys. **165**, 275 (1996).
12. J. Heyvaerts, J. M. Priest, and D. M. Rust, Astrophys. J. **216**, 123 (1977).
13. G. J. Hurford, R. B. Read, and H. Zirin, Sol. Phys. **94**, 413 (1984).
14. M. M. Kobrin, A. I. Korshunov, S. I. Arbuzov, *et al.*, Sol. Phys. **56**, 359 (1978).
15. D. V. Korol'kov and Yu. N. Parijskij, Sky Telesc. **57**, 324 (1979).
16. M. R. Kundu, V. Gaizauskas, B. E. Woodgate, *et al.*, Astrophys. J., Suppl. Ser. **57**, 621 (1985).
17. M. R. Kundu, V. V. Grechnev, V. I. Garaimov, and S. M. White, Astrophys. J. **563**, 389 (2001).
18. M. R. Kundu, S. M. White, K. Shibasaki, and J. P. Raulin, Astrophys. J., Suppl. Ser. **133**, 467 (2001).
19. K. R. Lang, R. F. Willson, J. N. Kile, *et al.*, Astrophys. J. **419**, 398 (1993).
20. V. P. Maksimov, V. P. Nefed'ev, and G. Ya. Smol'kov, Issled. Geomagn. Aeronom. Fiz. Soln. **82**, 155 (1988).
21. A. Nindos, M. R. Kundu, S. M. White, *et al.*, Astrophys. J., Suppl. Ser. **130**, 485 (2000).
22. H. E. Petchek, in *Proc. AAS-NASA Symp. on Physics of Solar Flares, NASA SP-50*, Ed. by W. N. Hess (Washington, D. C., 1964), p. 425.
23. L. A. Pustil'nik, Astron. Zh. **50**, 1211 (1973) [Sov. Astron. **17**, 763 (1973)].
24. B. I. Ryabov, N. A. Pilyeva, C. E. Alissandrakis, *et al.*, Sol. Phys. **185**, 157 (1999).
25. K. Shibata, in *Proc. of Nobeyama Symposium, Kiyosato, Japan, 1998*, Ed. by T. S. Bastian, N. Gopalswamy, and K. Shibasaki, NRO Report **479**, 381 (1998).
26. S. I. Syrovatskii, Izv. Akad. Nauk SSSR, Ser. Fiz. **48**, 37 (1979).
27. H. Tanaka and S. Enome, Sol. Phys. **40**, 123 (1975).
28. V. V. Zaitsev and A. V. Stepanov, Sol. Phys. **88**, 297 (1983).
29. V. V. Zheleznyakov, *Electromagnetic Waves in Cosmic Plasma* (Nauka, Moscow, 1977).
30. V. V. Zheleznyakov, *Radio Emission from the Sun and Planets* (Nauka, Moscow, 1964).
31. V. V. Zheleznyakov, V. V. Kocharovsky, and Vl. V. Kocharovsky, Astron. Astrophys. **308**, 685 (1996).
32. E. Ya. Zlotnik, Radiophys. Quantum Electron. **44**, 53 (2001).

Translated by V. Astakhov

Hill Stability in the Many-Body Problem

L. G. Luk'yanov, L. P. Nasonova[*], and G. I. Shirmin

Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119992 Russia
Received September 23, 2002

Abstract—We established a criterion for the Hill stability of motions in the problem of many spherical bodies with a spherical density distribution. The region of Hill stability was determined. The sizes of this region are comparable to the total volume of all of the bodies in the system, which sharply increases the probability of mutual collisions. This result may be considered as a confirmation that a supermassive core can be formed at the center of a globular star cluster. The motions in the n-body problem are shown to be unstable according to Hill. © *2003 MAIK "Nauka/Interperiodica"*.

Key words: *celestial mechanics, many-body problem, zero kinetic energy surface, Hill stability, collisions of bodies.*

Luk'yanov and Shirmin (2001, 2002) explored the possibility of motions that are stable according to Hill in the general three-body problem. They showed that all motions are unstable according to Hill in the problem of three point bodies (Luk'yanov and Shirmin 2001) and that there is Hill stability in the problem of three spherical bodies with a spherical mass distribution when certain conditions are satisfied (Luk'yanov and Shirmin 2002).

Here, we explore the possibility of motions that are stable according to Hill in the problem of many rigid bodies in the shape of spheres with a spherical mass distribution. We study several possibilities where the initial n-body system ceases to exist either during its breakup, when one or more bodies recedes to an infinitely great distance from the remaining bodies, or during a collision of at least one pair of bodies. We also consider the case where the bodies are mass points.

Consider a system of n absolutely rigid bodies M_1, M_2, …, M_n that are spheres of radii a_1, a_2, …, a_n and masses m_1, m_2, …, m_n, respectively. All of the bodies have a spherical density distribution; i.e., their mass densities depend only on the distance from the center of the corresponding body. Such bodies are attracted to one another as mass points (Duboshin 1961).

The force function of the mutual attraction between such bodies is

$$U = \sum_{j>i} U_{ij}, \qquad (1)$$

where

$$U_{ij} = f\frac{m_i m_j}{r_{ij}}, \qquad (2)$$

f is the universal gravitational constant, and r_{ij} is the distance between the centers of bodies M_i and M_j.

The force function of the attraction between zero-size bodies (mass points) is unbounded, while the force function for finite-sized bodies is always bounded. Since the force function is bounded, we obtain the following results on the Hill stability of the motions of spherical bodies.

The force function for finite-sized bodies is bounded for two reasons. First, the centers of mass of the bodies can never approach each other to a zero distance and, second, the force function is always finite inside a three-dimensional body (Duboshin 1961). The latter is not realized in the statement under consideration, because all bodies are assumed to be absolutely rigid, but it is of interest, for example, during a collision of galaxies, where the galaxy bodies penetrate each other.

The smallest distance to which two spherical bodies can approach each other is determined by the sum of their radii; i.e., the following inequality is always satisfied:

$$r_{ij} \geq a_i + a_j. \qquad (3)$$

The smallest distance is reached when the bodies are in contact (collide). The initial system ceases to exist during the collision of at least one pair of bodies.

The boundedness of the force function is specified by the inequality

$$U \leq V, \qquad (4)$$

[*]E-mail: `nason@sai.msu.ru`

where the constant quantity V corresponds to some compact (optimal) contact of all n bodies when each body is in contact with at least two other bodies. Obviously, no other compact arrangements of n bodies can yield U values larger than V.

Such a compact contact of all bodies can be treated as an analogue of an n-fold collision in the problem of n mass points. We call it an n-fold collision of spherical bodies.

During the n-fold collision, there can be pairs of bodies the distances between the centers of which exceed the sum of their radii. However, this imposes no fundamental constraints on the results described below.

Denote the force function of the mutual attraction between $(n-1)$ bodies that remain after the removal of body M_k from the initial n-body system (e.g., when body M_k goes to infinity) by $U^{(k)}$. This force function is also bounded:

$$U^{(k)} \leq V^{(k)}, \qquad (5)$$

where $V^{(k)}$ is the maximum value of $U^{(k)}$ that is reached during the $(n-1)$-fold collision of bodies $M_1, \ldots, M_{k-1}, M_{k+1}, \ldots,$ and M_n.

The bodies are numbered in such a way that the following inequalities are valid:

$$V^{(1)} \leq V^{(2)} \leq \cdots \leq V^{(n)} \leq V. \qquad (6)$$

Conditions (6) are definitely satisfied if the following inequalities are valid:

$$m_1/a_1 \geq m_2/a_2 \geq \cdots \geq m_n/a_n \qquad (7)$$

and

$$m_1 \geq m_2 \geq \cdots \geq m_n. \qquad (8)$$

The differential equations of motion for the many-body problem admit the energy integral that can be represented in the barycentric coordinate system as

$$T - U = h, \qquad (9)$$

where T is the system kinetic energy and h is the barycentric energy constant.

The condition that the kinetic energy is nonnegative for $h < 0$ specifies the regions where the motions of bodies are possible in the form of the inequalities

$$U \geq C, \qquad (10)$$

where we introduced a new arbitrary constant $C = -h$ that can take on real values in the range $0 < C \leq V$.

The boundaries of these regions are the surfaces of zero kinetic energy

$$U = C. \qquad (11)$$

The surfaces of zero kinetic energy are hypersurfaces in the $(3n)$-dimensional space of Cartesian coordinates of the bodies. Passing to relative coordinates allows the space dimension to be reduced to $3(n-1)$. If we use the Cartesian coordinate system of mutual distances $Or_{12}, \ldots, r_{n-1n}$, as was done by Luk'yanov and Shirmin (2001, 2002) for the three-body problem, then the dimension of the above space is $C_n^2 = n(n-1)/2$. Introducing such a coordinate system decreases and increases the space dimension for $n < 6$ and $n > 6$, respectively, compared to the space of relative coordinates.

In the space of relative distances, all of the motions of the bodies are described by continuous curves located within the region of (physically) admissible values.

This region of admissible values is characterized by the following inequalities:

(i) C_n^2 conditions for pair collisions of bodies M_i and M_j

$$r_{ij} \geq a_i + a_j; \qquad (12)$$

(ii) $3C_n^3$ inequalities of triangles $M_i M_j M_k$

$$r_{ij} + r_{jk} \geq r_{ik}; \qquad (13)$$

(iii) $4C_n^4$ inequalities of triangular pyramids $M_i M_j M_k M_p$ (polyhedrons with four vertices)

$$s_{ijp} + s_{jkp} + s_{kip} \geq s_{ijk}; \qquad (14)$$

where s_{ijk} is the area of triangle $M_i M_j M_k$,

(iv) $5C_n^5$ inequalities for polyhedrons $M_i M_j M_k M_p M_q$ with five vertices

$$v_{ijkq} + v_{jkpq} + v_{kipq} + v_{ijpq} \geq v_{ijkp}; \qquad (15)$$

where v_{ijkp} is the volume of pyramid $M_i M_j M_k M_p$, and

$$\ldots\ldots\ldots$$
$$\ldots\ldots\ldots$$

(v) n inequalities for polyhedron $M_1 M_2 \ldots M_n$ with n vertices

$$v^{(1)} + v^{(2)} + \cdots + v^{(n-1)} \geq v^{(n)}, \qquad (16)$$

where $v^{(i)}$ is the volume of polyhedron $M_1, \ldots, M_{i-1}, M_{i+1}, \ldots, M_n$ with $(n-1)$ vertices.

The subscripts $i, j, k, p,$ and q in inequalities (12)–(16) and in the other formulas are always different and take on all possible integer values from 1 to n.

The regions where motion is possible (10) are confined between the corresponding surfaces of zero kinetic energy (11) and the boundaries of the region of admissible values (12)–(16).

The regions where motion is possible for the problem of four or more bodies are difficult to represent

graphically because of the large dimension of the space of mutual distances.

However, some qualitative properties of the motions of n bodies can be inferred directly from inequalities (10). Of greatest interest is the study of the Hill stability. We use the following definition of the Hill stability.

We call the motions in the n-body system where the mutual distances act as the Cartesian coordinates stable according to Hill (in the $Or_{12} \ldots r_{n-1n}$ coordinate system) if there is such a constant C^* that for any initial mutual distances $r_{ij}(t_0) = r_{ij}^{(0)}$ that satisfy the inequality

$$U(r_{ij}^{(0)}) = C^{(0)} > C^*,$$

the regions where motions are possible $U(r_{ij}) \geq C^{(0)}$ have finite sizes.

If there is no such C^*, then we say that the motions in the n-body problem in the coordinate system under consideration are unstable according to Hill.

It is clear from the definition that when the inequality $C^{(0)} > C^*$ is satisfied, all of the mutual distances $r_{ij}(t)$ are bounded for any time $t\epsilon(-\infty, \infty)$.

The main difference between our definition of stability and its classical definition used by Hill in the restricted problem of three point bodies is that one passively gravitating mass is located within the region where motion is possible in the restricted three-body problem, while in our problem, all n bodies are located within this region.

In addition, all of the regions where motion is possible in the many-body problem are simply connected. Therefore, the definition of stability does not include the additional requirement that the initial conditions belong to the finite region where motion is possible. This requirement is necessary for the restricted three-body problem, because the regions where the motion of a low-mass body is possible can be multiply connected due to the rotation of the coordinate system and due to the presence of main finite-mass bodies.

The following Hill stability criterion can be easily derived for a many-body system:

$$V^{(n)} < C \leq V. \tag{17}$$

Indeed, if $C > V^{(n)}$, then it follows from (4), (6), and (10) that

$$V^{(n)} < C \leq U < V.$$

It is then clear from the definition of $V^{(n)}$ that all of the bodies remain at finite distances from each other as the time varies between the infinite limits, $-\infty < t < +\infty$. None of the bodies can recede to an infinitely large distance from the remaining bodies,

because this is possible only for $C \leq V^{(n)}$. Therefore, the body trajectories for $C > V^{(n)}$ in the space of mutual distances are always bounded by the closed surface of zero kinetic energy $U = V^{(n)} + \varepsilon$, where $\varepsilon > 0$. This proves the Hill stability of the motions in the n-body system.

If the constant C does not satisfy condition (17), then at least one of the bodies can recede to an infinitely large distance. Indeed, if

$$V^{(n-1)} < C \leq V^{(n)},$$

then body M_n can go to infinity. This is confirmed by the inequalities

$$V^{(n-1)} < C \leq U^{(n)} \leq V^{(n)}$$

and by the definitions of $V^{(n-1)}$ and $V^{(n)}$. Body M_{n-1} (and the other bodies) cannot go to infinity, because this would lead to the contradictory inequalities

$$V^{(n-1)} < C \leq U^{(n-1)} \leq V^{(n-1)}.$$

If the constant C takes on the values that satisfy the inequalities

$$V^{(n-2)} < C \leq V^{(n-1)},$$

then it can be easily shown by similar reasoning that either body M_n or M_{n-1} leave the system.

Finally, for

$$0 < C \leq V^{(1)},$$

any body can recede to an infinitely large distance. Moreover, two or more bodies can simultaneously leave the system in this case. However, to establish this possibility (i.e., to determine the corresponding C), we must investigate the systems of $n-1$, $n-2$, etc., bodies by a similar method.

The satisfaction of the Hill stability criterion does not rule out collisions between bodies. It can be asserted that when condition (17) is satisfied, the n-body system either exists eternally (for $t \to \pm\infty$) or two or more bodies collide at some time t and the system ceases to exist in its original state.

When condition (17) is satisfied, the constant C is so large that most (possibly all) of the mutual distances between the bodies are, according to inequality (10), comparable to the sizes of the bodies themselves. We assume, for simplicity, that all mutual distances are equal ($r_{ij} = r$); in defining $V^{(n)}$, we arbitrarily assume that in an $(n-1)$-fold collision, each pair of bodies is in contact. Under these assumptions, from the inequalities

$$U \geq C > V^{(n)}$$

we obtain the following estimate for the size of the region where criterion (17) is satisfied:

$$r < \frac{n}{n-2}(a_1 + a_2). \qquad (18)$$

Hence, we see that to satisfy the Hill stability conditions it is required that at least some of the bodies be so close to each other that the probability of their collisions becomes high. Strictly speaking, our assumption that the mutual distances are equal ($r_{ij} = r$) can be valid only at $n \leq 4$. It yields (18, which underestimates the size of the region; actually, the bodies can occupy a slightly larger space. However, even if we increase r severalfold, the probability of body collisions remains high.

Let us now consider the case where all or some of the bodies are mass points. The radii of these bodies become zero, while their masses remain the same.

First, let only one body M_1 be a mass point. Physically, this implies that the size of this body is negligible compared to the sizes of the remaining bodies, while the masses of all of the bodies are of the same order of magnitude. All of the previous results remain valid for such a system. We should only set $a_1 = 0$ in all of the formulas. In this case, the Hill stability regions are completely preserved.

If two bodies, M_1 and M_2, are mass points and the rest of the bodies are finite, then V and all quantities $V^{(k)}$ except $V^{(1)}$ and $V^{(2)}$ become infinite. No Hill stability region exists. Strictly speaking, this region exists but its size is so small (comparable to the negligibly small sizes of bodies M_1 and M_2) that none of the finite bodies can be accommodated in it. At any C, any body except point bodies M_1 and M_2 can always go to infinity. Bodies M_2 and M_1 can also leave

the system for $V^{(1)} < C \leq V^{(2)}$ and $0 < C \leq V^{(1)}$, respectively.

If the number of mass points in an n-body system is equal to or larger than three, then the pattern of motions is qualitatively the same as that for a system of n mass points. More specifically, no Hill stability region exists and any body can leave the system at any C.

The final result of our analysis of Hill stability in the n-body problem can be formulated as follows.

In the problem of n spherical bodies with a spherical density distribution, the motions of all bodies are stable according to Hill if condition (17) is satisfied. However, the sizes of the stability regions are small and comparable to the sizes of the bodies themselves. Therefore, the satisfaction of criterion (17) is related to the high probability of collisions between the bodies. This result can be considered as an argument for the possible formation of a supermassive core at the center of a globular star cluster.

In the problem of n mass points, all motions are unstable according to Hill, because under any initial conditions, any body can leave the system and recede to an infinite distance from the remaining bodies.

REFERENCES

1. G. N. Duboshin, *The Theory of Attraction* (Fizmatgiz, Moscow, 1961).
2. L. G. Luk'yanov and G. I. Shirmin, Pis'ma Astron. Zh. **27**, 948 (2001) [Astron. Lett. **27**, 816 (2001)].
3. L. G. Luk'yanov and G. I. Shirmin, Pis'ma Astron. Zh. **28**, 477 (2002) [Astron. Lett. **28**, 419 (2002)].

Translated by A. Dambis