Astronomy Reports, Vol. 49, No. 1, 2005, pp. 1–4. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 3–7.
Original Russian Text Copyright © 2005 by Fidelis.

Observations of Very-High-Energy Gamma Rays from Mrk421 at the Crimean Astrophysical Observatory in 2002

V. V. Fidelis

Crimean Astrophysical Observatory, Nauchny, Crimea, Ukraine

Received March 18, 2004; in final form, May 27, 2004

Abstract—Observations of the Markarian galaxy Mrk421 were carried out in 2002 with the Cherenkov gamma ray telescope of the Crimean Astrophysical Observatory. Analysis of the data indicated an excess of very-high-energy photons in the direction of Mrk421. The mean flux over the observing period was 1.55 Crab ($E > 1$ TeV), with a statistical confidence level of 4.6σ. A flare with a flux of ≈2.4 Crab was detected during the observing period. © *2005 Pleiades Publishing Inc.*

1. INTRODUCTION

Electron–photon showers induced by very-high-energy (VHE) gamma rays ($E > 10^{12}$ eV) are absorbed in the atmosphere, and so do not reach the Earth. They can be detected only via Cherenkov radiation generated by electrons and positrons in the atmosphere. Most of the photons arrive from the region of maximum development of the shower or from somewhat lower in the atmosphere (heights of 5–10 km). The total amount of light from the shower is proportional to the energy of the primary gamma-ray, and the shower itself propagates in a narrow cone close to the direction of propagation of the primary gamma-ray. Second-generation gamma-ray telescopes enable imaging of a Cherenkov flash in the field of view, as well as estimation of its size, shape, and orientation [1]. The number of detected gamma-ray-like flashes can be used to estimate the fluxes of the primary gamma rays via modeling of the detected showers [2].

All the extragalactic sources detected at high gamma-ray energies by space observatories are blazars, whose radiation is dominated by broadband nonthermal emission, believed to be the synchrotron radiation emitted by electrons accelerated to high energies in relativistic jets that are oriented roughly toward the observer. The blazars that have been detected at ultra-high energies are BL Lac objects, which display weak or absent optical line emission [3]. It is believed that the dominant mechanism for the generation of VHE gamma-rays in BL Lac objects is inverse Compton scattering of photons on high-energy electrons in the relativistic jet. For BL Lac objects that possess emission lines, and therefore display a thermal radiation component, the transfer of energy to external photons coming toward the jet at large angles from the vicinity of the accretion disk of the active galactic nucleus is also possible [4]. Some combination of various emission mechanisms is likewise possible. Thus far, VHE gamma-rays have been detected for only a handful of BL Lac objects (Mrk421, Mrk501, PKS 2155–304, 1ES 2344+514, 1ES1959+650, 3C66A, BL Lac, H 1426+428, NGC253). Observations of these objects can help in refining their energy spectra and the parameters for acceleration models [5]. The rapid variability of these objects at TeV energies places constraints on the physical parameters of models, such as the source size and mechanisms for the acceleration and cooling of the charged particles.

Observations of the strong variability of Mrk421 from radio to VHE gamma-ray energies (e.g., doubling of the flux over 20 minutes) [6] can be used to infer variations in the geometry and physical parameters of the radiating region, and to construct models for the observed flares [7].

VHE gamma-rays were first detected in the direction of Mrk421 ($z = 0.031$) in March–June 1992 by the Cherenkov telescope of the Whipple Observatory. The flux level was equivalent to 0.3 Crab ($E > 0.5$ TeV) with a statistical confidence of 6σ [8] (one "Crab" corresponds to the observed flux from the Crab nebula). The mean fluxes from Mrk421 measured by the Whipple Observatory in 1995 and 1997 were 0.35 and 0.2 Crab ($E > 1$ TeV), respectively. A series of flares were detected in 1996. The most prominent of these occurred on May 7, 1996, when the flux increased by a factor of five over 2.5 h, reaching a peak value of 10 Crab, and on May 15, 1996, when the flux rose to a factor of 14 above its mean value within 30 min [9].

Mrk421 has been observed in the optical, radio, X-ray, and gamma-ray ranges. It displays variability

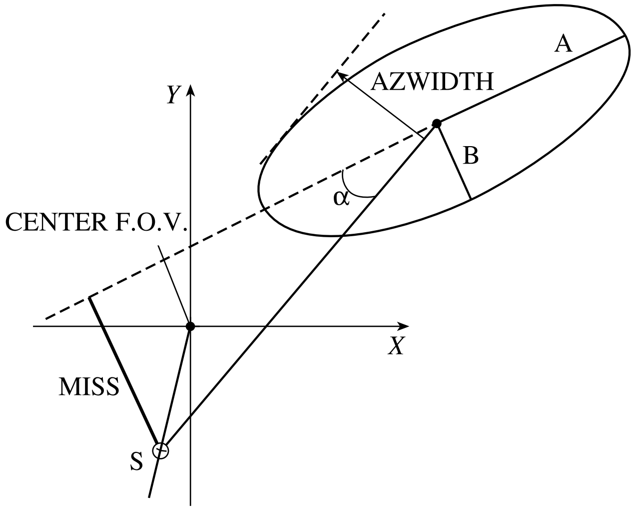


Fig. 1. Schematic of the main parameters of a flash. S is the position of the source. The point O corresponds to the center of the brightness distribution of the flash. The segment OS corresponds to the parameter DIST. "Center F.O.V." indicates the center of the imaging chamber.

on time scales from several years to months or less; its optical variability can reach 4.7^{m} [10].

A gamma-ray flare from Mrk421 was observed in December 2002 at MJD 52613–52614 on the Whipple Observatory 10-m Cherenkov telescope, accompanied by an X-ray flare (3–25 keV) detected by the RXTE satellite [11]. The data we present here confirm this variability.

2. THE CERENKOV TELESCOPE

The GT-48 Cherenkov telescope consists of two parallel sections located at an altitude of 600 m above sea level: a northern (N) and a southern (S) section, separated by 20 m in the north–south direction. On each section, four groups of parabolic mirrors focus the Cherenkov radiation to light receivers consisting of 37 photomultiplier cells forming 37 channels. The field of view of each cell is 0.4°, and the field of view of the entire receiver is 2.6°. The total area of the mirrors in both sections is 36 m^2. In each section, the signals from the photomultipliers are summed channel-by-channel. Cherenkov flashes can be identified when they arrive simultaneously at both sections. The energy threshold of the telescope is 1 TeV. A description of the setup can be found in [12].

3. ANALYSIS OF OBSERVATIONAL DATA

Mrk421 (J2002 coordinates $\alpha = 11^{\mathrm{h}}04^{\mathrm{m}}34^{\mathrm{s}}$ and $\delta = 38^{\circ}11'58''$) was observed at the Crimean Astrophysical Observatory on December 2–6, 2002.

The observations were carried out in a tracking regime. The pointing accuracy was ±0.05°. Observations of Mrk421 lasting 35^{m} (mode ON) were alternated with observations of the background at a position offset in right ascension α by 40^{m} with the same duration (mode OFF). The observations in the ON and OFF modes were conducted for the same zenith distances. The source was observed on seven moonless nights. A total of eight pairs of sessions were carried out, one of which was excluded from the subsequent analysis due to the strong irregularity in the count rate during background observations due to poor weather conditions. The source was observed for a total of 4 h 5 min.

After excluding events observed with poor pointing, applying calibration corrections, excluding flashes with a maximum amplitude in the outer ring of the receivers, and excluding events for which the charge–count transformer was saturated in at least one channel, we were left with 3667 Cherenkov flashes toward the source and 3723 toward the background.

The images of Cherenkov flashes from gamma-rays differ from those from cosmic rays in their smaller sizes and direction toward a source. Selection principles based on these differences can filter out up to 99% of flashes from charged particles while leaving an appreciable fraction of gamma-ray events.

The sizes of the flashes are characterized by the effective length A and effective width B, which are found by computing the first and second moments of the distribution remaining after the preliminary processing of the events. Flashes with low energy have large uncertainties, and a limit on their total amplitude V is imposed. The images of flashes from proton and gamma-ray showers also have different appearances. The former can have several maxima and be fragmented, while the latter are compact. The Individual Pixel Rate (IPR) characterizes the shape of flashes, and has a value of 0 for compact images and values of 1–7 for increasingly spread out images. The parameters A, B, V, and IPR do not depend on the position of the source relative to the flash image, and are accordingly coordinate-independent parameters.

The orientation of the flash relative to the direction toward the source is characterized by the coordinate-dependent parameters ALPHA, MISS, AZWIDTH, and DIST [2]. The selection parameters are depicted schematically in Fig. 1. Events whose selection parameters did not fall in the specified range were excluded from consideration.

The boundary values of the selection parameters were chosen so that the quantity $Q = (N_s - N_b)/\sqrt{N_s + N_b}$ was maximum. Here, N_s and N_b denote the number of source and background events, respectively.

Results of the Selection

Selection method	N_s	N_b	$N_s - N_b$	Q
No selection	3667	3723	−56	−0.65
Selection based on coordinate-independent parameters	325	259	66	2.73
Selection based on MISS	139	72	67	4.61

The showers used for analysis were restricted to the flare-image sizes $0.19° < A(\mathrm{N}) < 0.34°$ and $0.19° < A(\mathrm{S}) < 0.37°$, and $0.14° < B(\mathrm{N}) < 0.20°$ and $0.14° < B(\mathrm{S}) < 0.19°$. We also excluded flashes with total amplitudes $V < 150$ discrete units (i.e., 105 photoelectrons) and IPR values greater than unity for each section.

The section based on the parameter DIST was realized in the ranges $0° < \mathrm{DIST(N)} < 1.1°$ and $0° < \mathrm{DIST(S)} < 1.1°$, and the MISS selection in the range $\mathrm{MISS} < 0.225°$.

The results of successive selection are summarized in the table. The mean flux of VHE gamma-rays ($E > 1$ TeV) from the Crab Nebula in 2002 was 0.176 ± 0.033 min^{-1} [13]. The estimated flux for Mrk421 for the same season is 1.55 ± 0.63 Crab.

The distribution of the number of selected gamma-ray flashes over the field of view of the receiver is presented in Fig. 2, while Fig. 3 shows a contour "map" of this distribution. The maximum of the projected distribution ($\Delta\alpha = 0.1°$, $\Delta\delta = 0.2°$) coincides with the coordinates of Mrk421 within the uncertainties in these coordinates.

4. VARIABILITY OF THE RADIATION

The observed time scales for the rapid variability of active galactic nuclei (AGN) can be used to estimate the sizes of the radiating regions. The TeV emission from Mrk421 was variable during the December 2002 observations; the mean daily signal expressed as the VHE gamma-ray count rate is shown in Fig. 4. A flare with a duration of ~1 day was detected during our observing period. On Dec. 3 (MJD 52612), the VHE gamma-ray flux from Mrk421 rose to ~2.3 Crab. The following night, it increased to ~2.4 Crab, and fell to nearly zero a day later. The statistical confidences of these events are 3.25 and 3.75σ. Unfortunately, worsening of the weather conditions hindered continuation of this series of observations.

Let us consider a spherical model for the radiating region, $R < cT\delta/(1+z)$, where z is the cosmological redshift of the source, T is the observed variability time scale, $\delta^{-1} = \gamma(1 - \beta\cos\theta)$ is the Doppler factor of the jet, β is the speed of the jet divided by the speed of light, γ is the Lorentz factor of the jet, and θ is the angle between the jet and the line of sight to the observer [6]. We then obtain for $T \sim 1$ day, $\delta \sim 10$, and the redshift of Mrk421, $z = 0.031$, an upper limit

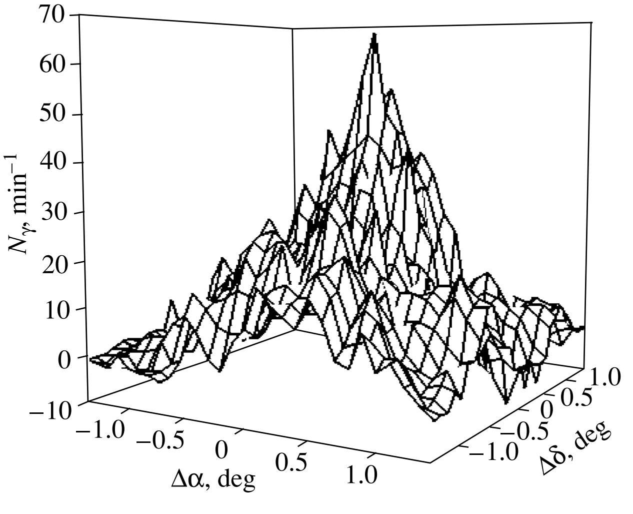


Fig. 2. Stereoimage of the distribution of gamma-ray arrival directions. $\Delta\alpha$ and $\Delta\delta$ are the offsets from the source position in right ascension and declination, respectively. N_γ is the number of events.

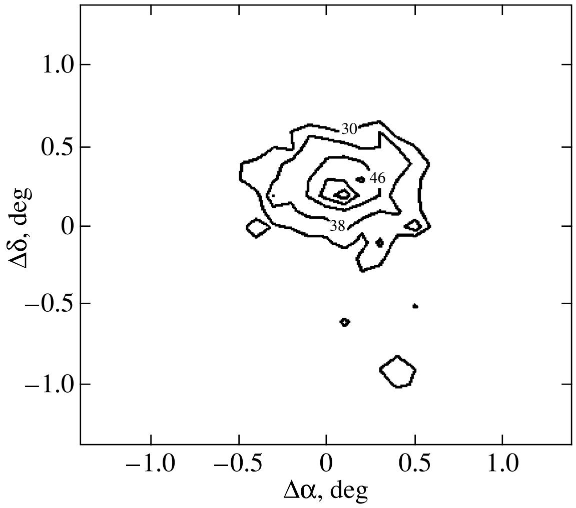


Fig. 3. Contours showing the distribution of gamma-ray arrival directions. The numbers on the curve indicate the number of events for the given contour.

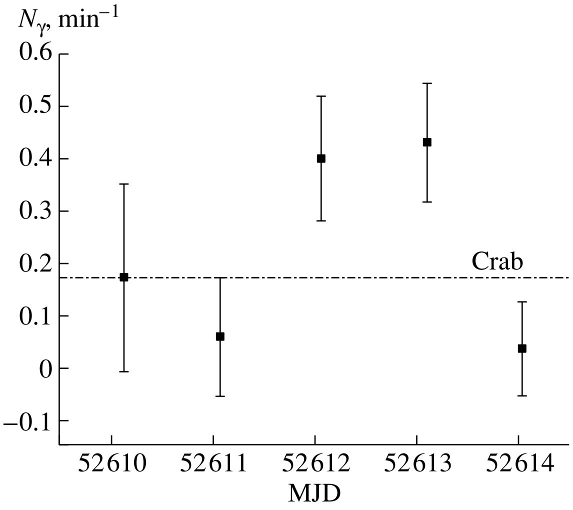


Fig. 4. Time behavior of the flux of gamma-rays (mean nightly value) with energy $E > 1$ TeV. The statistical errors are indicated. The dashed line shows the count rate from the Crab Nebula observed during the same season.

for the size of the relativistic structures responsible for the variability $R < 9.7$ light days. This simple, one-zone model assuming the presence of fairly extended radiating regions with sizes much larger than the commonly adopted values ($\sim 10^{16}$ cm) [6] does not preclude the possible presence of other more complex emission models for the variability [7]. The detected variability may also be due to an injection of high-energy charged particles into the radiating region, or the combined effect of variations in the Doppler and Lorentz factors.

5. CONCLUSION

A detected point source of VHE gamma-rays with energies exceeding 1 TeV can be identified with the galaxy Mrk421 with a statistical confidence of $\approx 4.6\sigma$. The flux of VHE gamma-rays from the galactic nucleus was variable on time scales of a day. Based on the mean integrated flux during the 2002 observational period, Mrk421 was in a more active state than the state observed by Punch *et al.* [8] in 1992 with the lower energy threshold of 0.3 Crab ($E > 0.5$ TeV).

Studies of the flare activity of blazars require the accumulation of data in various frequency bands. From this point of view, monitoring of blazars at TeV energies provides key information for our attempts to construct as full a picture of the relativistic structures in AGN as possible.

ACKNOWLEDGMENTS

In conclusion, the author thanks A.A. Stepyanyan and Yu.I. Neshpor for useful comments during discussions of this work and S.G. Kochetkova and E.M. Nekhaĭ for help in preparing the manuscript.

REFERENCES

1. T. C. Weekes and K. E. Turver, *Proc. 12-th ESLAB Symp. Frascati, Italy* (1977), p. 279.
2. A. M. Hillas, *Proc. 19th Intl. Cos. Ray Conf.* (La Jolla, USA, 1985), Vol. 3, p. 445.
3. R. M. Sambruna, L. Maraschi, and U. C. Megan, Astrophys. J. **463**, 444 (1996).
4. S. Inoue and F. Takahara, Astrophys. J. **463**, 555 (1996).
5. L. Costamante and G. Ghisellini, Astron. Astrophys. **384**, 56 (2002).
6. J. A. Gaidos, C.W. Akerlof, S. D. Biller, *et al.*, Nature **383**, 319 (1996).
7. K. Katarzynski, H. Sol, and A. Kus, Astron. Astrophys. **410**, 101 (2003).
8. M. Punch, C. W. Akerlof, M. F. Cawley, *et al.*, Nature **358**, 477 (1992).
9. J. E. McEnery, I. H. Bond, P. J. Boyle, *et al.*, Preprint No. 4545 (Harvard–Smithsonian Center for Astrophysics, 1997).
10. D. Petry, *Kruger National Park Workshop on TeV Gamma Ray Astrophysics, Towards a Major Atmospheric Cherenkov Detector V. Space Research Unit*, Ed. by O. C. de Jager (WESTPRINT–Potchefstroom, South Africa, 1997), p. 2.
11. P. F. Rebillot, S. B. Hughes, I. H. Bond, *et al.*, *Proc. 28th Intl. Cos. Ray Conf.* (Tsukuba, Japan, 2003), p. 2599.
12. B. M. Vladimirskii, Yu. L. Zyskin, A. P. Kornienko, *et al.*, Izv. Krymsk. Astrofiz. Observ. **91**, 74 (1994).
13. Yu. I. Neshpor, V. S. Eliseev, N. A. Zhogolev, *et al.*, Izv. Krymsk. Astrofiz. Observ. **99**, 43 (2003).

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 13–23. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 17–28.
Original Russian Text Copyright © 2005 by Tutukov.

Formation of Supermassive Black Holes due to the Tidal Deceleration of Stellar Black Holes, Globular Clusters, and Galaxies

A. V. Tutukov

Institute of Astronomy, Russian Academy of Sciences, Pyatnitskaya ul. 48, Moscow, 109017 Russia

Received July 10, 2004; in final form, July 15, 2004

Abstract—The formation and evolution of supermassive (10^2–$10^{10} M_\odot$) black holes (SMBHs) in the dense cores of globular clusters and galaxies is investigated. The raw material for the construction of the SMBHs is stellar black holes produced during the evolution of massive (25–$150 M_\odot$) stars. The first SMBHs, with masses of $\sim 1000 M_\odot$, arise in the centers of the densest and most massive globular clusters. Current scenarios for the formation of SMBHs in the cores of globular clusters are analyzed. The dynamical deceleration of the most massive and slowly moving stellar-mass ($< 100 M_\odot$) black holes, accompanied by the radiation of gravitational waves in late stages, is a probable scenario for the formation of SMBHs in the most massive and densest globular clusters. The dynamical friction of the most massive globular clusters close to the dense cores of their galaxies, with the formation of close binary black holes due to the radiation of gravitational waves, leads to the formation of SMBHs with masses $\gg 10^3 M_\odot$ in these regions. The stars of these galaxies form galactic bulges, providing a possible explanation for the correlation between the masses of the bulge and of the central SMBHs. The deceleration of the most massive galaxies in the central regions of the most massive and dense clusters of galaxies could lead to the appearance of the most massive (to $10^{10} M_\odot$) SMBHs in the cores of cD galaxies. A side product of this cascade scenario for the formation of massive galaxies with SMBHs in their cores is the appearance of stars with high spatial velocities (> 300 km/s). The velocities of neutron stars and stellar-mass black holes can reach $\sim 10^5$ km/s.

1. INTRODUCTION

There are two well-known manifestations of black holes: they can be observed as the components of close binaries and in the nuclei of galaxies. Stellar-mass (~ 4–$100 M_\odot$) black holes are the end products of the evolution of the most massive stars, with initial masses exceeding $\sim 25 M_\odot$ [1]. They are manifest as bright X-ray sources if they are present in close binary systems and are efficiently fed by gas from a companion. The masses of stellar black holes are close to the masses of the carbon–oxygen cores of the progenitors, since the gravitational binding energy of the products of carbon and oxygen burning is greater than the nuclear energy released in the burning [2]. The masses of observed main-sequence stars reach $\sim 150 M_\odot$ [3, 4], and estimates based on numerical modeling of the evolution of massive stars [2] show that stellar black holes with masses up to $100\ M_\odot$ can be formed. A low abundance of heavy elements, as is typical for the stars in globular clusters, weakens the stellar wind, facilitating the formation of black holes. There are well-known examples of Galactic close binaries containing black holes with masses of 4–$16 M_\odot$ [1]. Moreover, compact, ultraluminous variable X-ray sources with luminosities of up to $(2$–$3) \times 10^{40}$ erg/s have recently been found in nearby galaxies [5, 6]. Most of these sources seem to be massive close binaries in which black holes with masses of up to $\sim 100 M_\odot$ are accreting matter from close companions [7]. The donor in such a system could be a nearby Wolf–Rayet star or a main-sequence star with a mass of 4–$8 M_\odot$ that fills its Roche lobe.

Much more massive black holes are encountered in the nuclei of the massive galaxies associated with quasars and BL Lac objects. The masses of these black holes are 10^4–$10^{10} M_\odot$ [8]. Due to their accretion of the gas in their vicinity, they are the brightest sources of radiation in the Universe (up to $10^{13} L_\odot$). The characteristic apparent brightnesses of quasars increase with their redshift due to obvious selection effects [8]. Assuming that the maximum distance at which a quasar can be detected R is limited by the mass of its black hole M_{BH} ($R \sim M_{BH}^{1/2}$), the apparent mass distribution of the black holes [8] can be corrected for this observational selection effect to obtain the local mass spectrum. This yields $dN \sim M_{BH}^{-2.5} dM_{BH}$ for accreting black holes with masses of 10^5–$10^9 M_\odot$. It is important to emphasize that most supermassive black holes (SMBHs) possess relatively small masses of $\sim 10^5 M_\odot$. In addition, the catalog of Ho [8] contains two objects with smaller masses, in the nuclei of M33 ($<1500 M_\odot$) and M110 ($<9 \times 10^4 M_\odot$). Therefore, stellar-mass black holes and the SMBHs in galactic nuclei probably form a continuous common mass function, with the observed black holes being separated into two families by observational selection effects. One such family represents accreting stellar black holes in X-ray binaries, while the other represents accreting SMBHs in the nuclei of galaxies with a well-developed spherical stellar component [9].

To reconstruct the true local mass distribution for SMBHs, we must take into account another obvious selection effect having to do with the duration of the bright X-ray stage of SMBHs. The available data on the durations of the stage of active accretion for quasars [10] ($\tau \sim M_{BH}^{-0.15}$) show a weak dependence on the mass. This provides hope that the estimates of the slope of the SMBH mass function presented above for active galactic nuclei (i.e., SMBHs in the stage of active accretion) are correct enough for a first approximation. The most massive accreting black holes, observed in the nuclei of the most massive and bright cD galaxies with $L \sim 10^{12} L_\odot$, have luminosities reaching $\sim 10^{13} L_\odot$ [11], and correspond to black-hole masses of up to $\sim 3 \times 10^9 M_\odot$ and accretion rates of $\sim 6 M_\odot$/yr. It is interesting that a considerable fraction of the brightest ($L > 3 \times 10^{11} L_\odot$) galaxies possess quasars in their nuclei, whereas quasars are present in only 0.001 of galaxies with luminosities below $6 \times 10^{10} L_\odot$ [11]. The observed low occurrence of quasars in the nuclei of such galaxies seems to be related to the relative duration of the phase of active accretion by ordinary SMBHs in the galactic nuclei. Marconi *et al.* [10] also obtained a short lifetime for the active phase of a quasar (10^8 yr). It is evident that the mass of the black hole cannot increase appreciably in the case of such durations for the active accretion phase, since the mass doubling time is restricted by the Eddington accretion rate, $\sim 10^8$ yr. Nevertheless, the masses of the black holes in some quasars can reach $\sim 10^9$–$10^{10} M_\odot$, even at $z \sim 6$; i.e., at ages of $\sim 10^9$ yr [12]. Therefore, the initial masses of these black holes must be $\sim 10^6 M_\odot$, even in a regime of continuous accretion at the Eddington limit over $\sim 10^9$ yr. Therefore, at least the most massive, young, observed quasars could not have started from stellar-mass black holes ($\sim 10 M_\odot$) and gained their masses via the accretion of gas.

An additional argument against a purely accretion origin for SMBHs in quasars can be derived from their total luminosities. The observed integrated relative luminosities of quasars (~ 0.1 of the luminosity of the galaxy [13]) enables us to impose constraints on the role of accretion in increasing the black-hole mass. The mass-to-luminosity ratio in solar units for SMBHs accreting at a constant rate over the Hubble time is ~ 0.01, whereas this ratio reaches 10–30 [14] for the progenitor galaxies, whose masses are one thousand times greater than those of the central black holes [13]. Therefore, we conclude that, to obtain consistency between the three above ratios, accretion must produce no more than 3–10% of the black-hole mass. This fact forces us to turn to other possible mechanisms for the formation of SMBHs.

As was noted above, observations have revealed black holes with stellar masses (4–$16 M_\odot$) in close binary systems [1]. It has not been possible to derive an empirical mass function for stellar black holes due to the small amount of data available [1]. This function can be obtained only theoretically, by assuming that the stellar black hole contains only the products of carbon and oxygen burning in the core of a massive star, whose nuclear energy is less than the gravitational binding energy: $M_{BH}/M_\odot = 0.03(M_{MS}/M_\odot)^{1.7}$ [2]. Taking the initial stellar-mass function in the form $dN \sim M_{MS}^{-2.5}$ [15], we can obtain the initial mass function for stellar black holes: $dN/dM_{BH} \sim (M_{BH}/M_\odot)^{-1.9}$. This expression is valid for remnants of stars with initial masses of ~ 25–$150 M_\odot$. It also enables us to estimate the relative mass of stellar black holes formed at the end of the evolution of massive stars, which is about 0.04 $(M_{min}/M_\odot)^{1/2}$ of the total initial mass of stars of the same generation, where M_{min} is the minimum mass of stars satisfying the above initial mass function for them.

The minimum mass of stars in the solar neighborhood M_{min} is a few tenths of a solar mass. However, to explain the rapid enrichment of the young nuclei of galaxies and quasars in heavy elements, we must suppose that the minimum mass of the stars formed there is considerably greater than in the solar neighborhood, possibly as large as 3–$10 M_\odot$ [12, 16, 17]. A study of a young stellar cluster with a mass of $\sim 10^6 M_\odot$ in M82 showed that there are no stars with masses below $\sim 3 M_\odot$ in this cluster [18]. There are also several other compact clusters with small mass-to-luminosity ratios (~ 0.03), indicating an absence of stars with masses below a few solar masses. Studies of reionization of hydrogen in the Universe also lead to the conclusion that the minimum mass of the first stars may exceed the solar mass [19]. If $M_{min} \sim M_\odot$, the above estimate of the relative mass of stellar black holes (~ 0.04) will obviously be more than a factor of ten greater than the usual estimate of the relative mass of SMBHs in the nuclei of active galaxies (~0.001). This is important for the cascade mechanism of the formation of SMBHs in galactic nuclei, which is a sequential hierarchical merging of a small fraction of the stellar black holes, modified by the stellar structures (globular clusters and galaxies) which contain them.

Increasing the minimum stellar mass in some globular clusters to more than a solar mass results in the possible existence (in our own and other galaxies) of "ghost" stellar clusters, composed primarily of cool, degenerate, low-luminosity dwarfs. Such structures could be represented, first and foremost, by the earliest globular clusters formed in the gas of young galaxies, which contain almost no heavy elements. All stars in such clusters with masses above $\sim 0.8 M_\odot$ should have finished their evolution, resulting in black holes, neutron stars, and, predominantly, degenerate dwarfs. Old stellar black holes and neutron stars are undetectable, while the luminosities of old degenerate dwarfs will be very small ($\sim 10^{-4} L_\odot$) [20]. Searches for such clusters are difficult, even for our own Galaxy. The brightest stars in these clusters will be the secondary components of close binaries; i.e., main-sequence stars that cannot become bright giants because they are in close binaries. However, we should bear in mind that such "invisible" clusters can actively participate in the formation of SMBHs in their nuclei, while preserving a substantial fraction of their high initial masses.

2. THE FORMATION OF BLACK HOLES WITH MASSES $\sim 10^3 M_\odot$ IN GLOBULAR CLUSTERS

As was shown in the Introduction, the formation of SMBHs in the nuclei of spherical stellar clusters and galaxies can be explained if we find a method for accumulating a few percent of the mass of their stellar black holes. Such a mechanism has been known for a long time: the dynamic (tidal) deceleration of massive objects in a gravitating system of N bodies, whose characteristic time scale is [21]

$$\tau_{fr} = \frac{v^3}{4\pi G^2 m\rho \ln\lambda}, \tag{1}$$

where v and m are the characteristic velocities and masses of the gravitating bodies, G is the gravitational constant, ρ is the ambient density, and $\ln\lambda \sim 5$ is the Coulomb logarithm. In order for the deceleration of a black hole to be more efficient, its mass and ambient density must be increased, while its spatial velocity must be decreased. The late stages of merging are governed by the radiation of gravitational waves on a characteristic time scale of

$$\tau_{GWR}(\mathrm{yr}) \sim 10^8 \frac{A^4}{mM(M+m)}, \tag{2}$$

where A is the semimajor axis of the system in solar radii and m and M are the masses of the merging bodies in solar masses. The energy absorbed by a black hole is compensated by the expansion of the cluster (galaxy), followed by a partial loss of its stars. As a result of this expansion, the cluster (stellar nucleus of the galaxy) acquires a characteristic spheroidal shape. In this scenario, the reason for the observed correlation between the mass of the spheroidal stellar component and the mass of the central SMBH becomes clear. The corresponding formula can be written in a form convenient for numerical estimates:

$$\tau_{fr}(\mathrm{yr}) \sim 10^{10} \frac{V^3 R^3}{mM}, \tag{3}$$

where V is the velocity (in km/s) of a decelerated object with mass m (in $M_\odot$) in a spheroidal stellar system with mass M (in $M_\odot$) and radius R (in pc). When using this expression, we should bear in mind that the number of black holes is $N \sim V^3$ in the case of a Maxwellian velocity distribution. Therefore, for example, the deceleration of one thirtieth of all black holes will take a factor of 30 less time than the time given by (3). The reason for this is the coincidence of the velocity dependences for the number of black holes and the deceleration time [see (1) and (3)].

We are interested here only in SMBHs, globular clusters, and galaxies that decelerate within the Hubble time, $\sim 1.4 \times 10^{10}$ yr. This enables us to simplify the last formula to

$$V^3 R^3 < mM. \tag{4}$$

Let us now consider a three-stage sequence of merging black holes, ultimately resulting in the formation of a SMBH with a mass of 10^9–$10^{10} M_\odot$ in the nucleus of a cD galaxy with a mass of 10^{12}–$10^{13} M_\odot$. Stellar-mass black holes initially merge in the cores of the most massive and dense globular clusters, forming black holes with masses of $\sim 10^3 M_\odot$. Next, the globular clusters that are the most massive and closest to the centers of their parent galaxies move into the galactic cores, resulting in an accumulation of black holes with masses of 10^5–$10^7 M_\odot$. In the final stage of this process, the most massive galaxies from the central regions of the densest and most massive clusters of galaxies move into the core regions of these clusters, forming cD galaxies whose nuclei contain the most massive black holes. Due to the foamlike structure of visible matter in the Universe, with a characteristic size for voids of about 100 Mpc, galaxy clusters usually separate with velocities of a few thousand km/s, preventing the extention of this process to larger scales. The simultaneous development of all the above stages shortens the total time for the formation of SMBHs.

Many publications (e.g. [22–28]) have been devoted to theoretical studies of the formation of and observational searches for black holes with intermediate masses ($\sim 10^3 M_\odot$). It is possible that the cores of many globular clusters (first and foremost, the most massive and dense) contain such black holes. Taking into account the inevitable merging of the end products of the evolution of massive close binaries under the action of gravitational-wave radiation [29, 30] and the observed maximum stellar masses of $\sim 150 M_\odot$ [31], the upper limit for the mass of stellar black holes appears to be about $100 M_\odot$. Three basic scenarios for forming more massive black holes have been considered. The first supposes that such black holes could result from the evolution of supermassive ($\sim 1000 M_\odot$) stars, which can probably be formed in the collisional evolution of stars in globular clusters [32] or during the formation of such clusters. In the second scenario, SMBHs are formed via the accretion of gas. Finally, in the third scenario, SMBHs result from mergers of stellar black holes that have moved into the core of a globular cluster due to the action of tidal friction. These black holes form compact binary black holes in the core and subsequently merge due to the radiation of gravitational waves [33]. Let us consider the specific features of these scenarios.

There are two ways to form supermassive ($\sim 1000 M_\odot$) stars in a globular cluster. In the first, such a star is formed due to a high accretion rate μ [34], due, in turn, to the high temperature T of the protostellar gas: $\mu = 10^{-7} T^{3/2} M_\odot/\mathrm{yr}$ [35]. To accumulate a star with a mass of $\sim 10^3 M_\odot$ over the time for hydrogen burning in its core, $\sim 2 \times 10^6$ yr, the initial temperature of the gas must exceed ~ 250 K. The temperature in the center of a globular cluster during its formation can be high, but the probability of forming supermassive stars in the cores of globular clusters via this scenario remains unclear, partly due to the large uncertainty in the expected intensities of the stellar winds from such massive stars. These winds may prevent the formation of such stars, since, as follows from observations, their intensities increase sharply with the mass of the star.

The second scenario for the formation of supermassive stars in globular clusters assumes the collisional growth of stars in the dense cluster cores [32]. Let us estimate the probability of the realization of this scenario. The energy conservation for a star with a mass of M, a radius of R, and a luminosity approximately equal to the Eddington luminosity can be written

$$\frac{M}{R}\frac{dR}{dt} = \frac{dM}{dt} - 4\pi c \frac{R}{\kappa}, \tag{5}$$

where dM/dt is the rate of accretion of stellar matter by the superstar, c is the speed of light, and $\kappa = 0.34\ \mathrm{cm^2/g}$ is the opacity of the ionized gas. In order for the size of the star to increase during the collisional accumulation of mass, the right-hand side of this equation must be positive. This increase is necessary in order for the merging of globular-cluster stars moving near the superstar to accelerate.

If we define the radius of the cross section for absorption to equal the radius of the superstar, R, and the average spatial velocity of the stars in the cluster is $v_0 = (GM/R)^{1/2}$, the impact parameter r will be equal to

$$r = (2GMR)^{1/2}/v_0, \tag{6}$$

where G is the gravitational constant and M is the mass of the supermassive star. Finally, the rate at which the supermassive star accretes stellar matter will be

$$dM/dt = 2\pi\rho GMR/v_0. \tag{7}$$

In order for the star's size to increase, the right-hand side of (5) must be positive, or the density of the cluster core must satisfy the condition

$$\rho > \frac{2cv_0}{3\kappa GM}. \tag{8}$$

Taking as an estimate $v_0 = 20$ km/s, we obtain the condition that the size of the superstar increase due to the collisional accumulation of mass: $\rho > 10^{-9}\ \mathrm{g/cm^3}$. In fact, the density of stars in the core of a globular cluster is $\sim 10^{-17}\ \mathrm{g/cm^3}$, and the density in the nucleus of our Galaxy $\sim 10^{-15}\ \mathrm{g/cm^3}$, so that the collisional increase in the sizes of superstars is impossible.

As a result, the supermassive star must remain on the main sequence, slowing the increase in its mass. The efficient increase in its mass requires that the mass-doubling time be shorter than the time for hydrogen burning in its core, T_H, which is $\sim 2 \times 10^6$ yr for a star with a mass of $\sim 1000 M_\odot$ [34]. Consequently, taking (3) into account, the condition that the mass of the superstar increase will be

$$\rho > \frac{v_0}{2\pi G R T_H}. \tag{9}$$

Assuming now $v_0 = 20$ km/s and $R = 50 R_\odot$ [19], we conclude that the ambient density must exceed $\sim 3 \times 10^{-14}\ \mathrm{g/cm^3}$ in order for the collisional accumulation of mass for a superstar with a mass of $\sim 1000 M_\odot$ to occur over the hydrogen-burning time. It is obvious that this requirement likewise cannot be satisfied in the core of a globular cluster or the nucleus of a galaxy. Consequently, the masses of stars in a globular cluster or the nucleus of a typical galaxy cannot increase appreciably due to collisions. Some stars may exhibit a faster collisional increase in their masses purely statistically, but the requirement for a large number of collisions ($\sim 10^3$) for the formation of a supermassive star can be satisfied only in the case of an improbable event ($\sim 2^{-1000} \simeq 10^{-300}$), even if the average mass-accumulation rate is doubled. Therefore, the collisional scenario is not promising.

Another possible mechanism for the formation of supermassive stars was based on the dynamical collapse of stars in the core of a globular cluster. However, the binary nature of a considerable fraction of stars in the cluster [27, 28] rules out this scenario. The close binaries accumulate the energy of the dense central stellar cluster, thereby decelerating the increase in the density and preventing the collapse of the stellar core of the globular cluster.

The role of the disk accretion of gas during the formation of a SMBH can be constrained by observations of the X-ray background, whose total flux at 2–10 keV is $\sim 6 \times 10^{-8}$ erg s^{-1} cm^{-2} [35, 36]. Assuming that disk accretion by a SMBH results in the release of $\sim 10^{19}$ erg/g in X-rays [10, Fig. 3], the local density of matter in the black holes must be $\sim 6 \times 10^{-37}$ g/cm^3 within the cosmological horizon if the rate of accretion is constant over the cosmological time. On the other hand, the observed local density of matter in black holes with masses above $\sim 10^6 M_\odot$ is $\sim 1.7 \times 10^{-35}$ g/cm^3 [37]. Consequently, the observed X-ray background can be responsible for only a small fraction of the local matter density contained in SMBHs.

Next, let us estimate the average density in stellar-mass black holes. As was shown in the Introduction, about 4% of the mass transformed into stars ultimately becomes stellar-mass black holes ($5-100 M_\odot$). For an overall density of baryonic matter of 2×10^{-31} g/cm^{-3} [39] to 4×10^{-30} g/cm^{-3} [33], the local density of matter in stellar-mass black holes should be $\sim 10^{-32}-10^{-31}$ g/cm^3. This density is obviously much greater than both the density of matter in SMBHs and the density of matter accreted by black holes of all masses. Therefore, most of the "visible" matter contained in black holes must be produced by the evolution of stars with initial masses of $25-150 M_\odot$ [38].

If the SMBHs in the cores of globular clusters and the nuclei of galaxies are cosmological, the reason for the correlation between the masses of the black holes and the masses of the spherical stellar components of these clusters and galaxies remains unclear. This makes this mechanism for the formation of SMBHs unlikely.

Let us consider the last of the above scenarios for the formation of SMBHs ($\sim 1000 M_\odot$) in the cores of globular clusters, in which they are produced by the sequential merging of stellar-mass black holes. As follows from the estimates of [40, 41], stellar-mass black holes move toward the centers of globular clusters under the action of tidal deceleration. Next, they sequentially merge with each other, emitting gravitational waves in the process. The characteristic time for the tidal deceleration of a star with mass M and velocity v in a medium with average density ρ can be estimated using (3) and (4). When $v_0 = 30$ km/s, $R = 3$ pc, $m = 10 M_\odot$, and $M = 3 \times 10^5 M_\odot$, this is about equal to the Hubble time. Since this time shortens with increasing black-hole mass, precisely the most massive ($\sim 30-100 M_\odot$) stellar black holes with small spatial velocities located close to the cluster center participate most actively in increasing the mass of the central black hole. The most massive stars, which produce the most massive black holes, are usually located in the central regions of their clusters. Therefore, a SMBH with a mass of $10^3 M_\odot$ can be formed even in the core of a fairly young ($\sim 10^9$ yr) globular cluster. We emphasize that the evolution of the orbits of the merging black holes in the latest stages of the merging process, just before coalescence, has not been studied. However, it is clear that the stellar density near the SMBHs must be fairly large in order for the initial components to approach each other to the distance required for an efficient action of gravitational radiation.

Only a few percent of stellar black holes in ordinary globular clusters participate in the formation of the central SMBHs. Indeed, if the relative mass of the central black holes in the spherical stellar systems—globular clusters and bulges—is about 0.001 of their mass [37], the relative mass of the stellar black holes should be a few percent of the stellar mass of the clusters. Due to the conservation of energy, the deceleration of black holes by the stellar field of a globular cluster should be followed by an expansion of the stellar core of the cluster. Such a deceleration has, in fact, been detected [41]: the sizes of the cores of globular clusters in the Large and Small Magellanic Clouds increase with the age of the clusters, from 1 pc for clusters with ages of $\sim 10^7$ yr to 3 pc for clusters with ages of $\sim 10^{10}$ yr. This expansion may reflect an increase of the central SMBHs in these cluster cores.

The expansion of the stellar component of a globular cluster accelerates its tidal dissipation. As a result, the globular clusters with the most massive central black holes could be completely evaporated, so that the formation of such black holes could be an efficient mechanism for the disintegration of the parent globular clusters. Another efficient mechanism for the dissipation of the central, most massive, globular clusters is merging with the nucleus of the parent galaxy due to the tidal deceleration of their spatial motion. The decrease in the number of globular clusters with time must indeed be explained, since, according to the observations of [42], an increase in the age of clusters by a factor of ten is associated with a decrease in their number by almost a factor of ten.

As a result, it is quite possible that the cores of many, or even all, globular clusters contain black holes with masses of $\sim 10^3 M_\odot$ [22, 23]. Their existence can be verified either via the dynamics of the circumnuclear stars or the X-ray emission of the cluster core (if the corresponding black hole has a close companion or other source of gas from which it efficiently accretes matter). If the spectrum of the X-ray radiation is soft (0.1–0.3 keV), the X-ray luminosities of such black holes can exceed $\sim 10^{40}$ erg/s. X-ray observations have revealed six black holes with masses of $700-2700 M_\odot$ in the cores of globular clusters in nearby galaxies [28]. The accretion rate in the brightest of these reaches $10^{-5} M_\odot$/yr [43, 44]. Therefore, the existence of SMBHs in the cores of at least some globular clusters seems to be well established. The detection of accreting SMBHs in the cores of globular clusters [43] poses the problem of the sources for their accretion. The observation of pulsars in the globular cluster 47 Tucanae made it possible to estimate the density of the interstellar ionized gas, which was found to be $\sim 10^{-25}$ g/cm^3 [45]. This gas seems to be ionized by the nuclei of planetary nebulas, whose formation rate is only $\sim 10^{-6}$ yr^{-1} and whose lifetimes are $\sim 10^4$ yr. However, the long recombination time ($\sim 10^{11}$ yr) leaves this gas ionized. It is carried away by the cluster wind [46], which is probably supported by millisecond pulsars [47]. The central SMBH obviously cannot accrete an appreciable fraction of this gas, since this would lead to extremely large X-ray luminosities ($\sim 10^{40}$ erg/s) for all globular clusters and masses of the central black holes comparable to the masses of the clusters. Recall that some fraction of the globular clusters might be "ghosts" composed predominantly of a small number of secondary components of binaries, old degenerate dwarfs, neutron stars, and stellar black holes. Such clusters could also have central black holes with masses of $\sim 1000\ M_\odot$.

3. MERGERS OF GLOBULAR CLUSTERS WITH THEIR PARENT GALAXIES

Some globular clusters are clusterings of the first stars of the parent galaxies, while others result from collisions (close encounters) of galaxies [48]. The number of globular clusters in a massive galaxy can reach several thousand; for example, about 6000 clusters were found in the E galaxy NGC 4472 [49]. Most of the globular clusters are formed in the central, dense regions of galaxies. For example, about 20% of the stars in a region with a radius of 400 pc in M82 are contained in globular clusters [50]. To estimate the time for the deceleration of the cluster in the parent galaxy, we must know its density (including the "dark" component), which is reflected by the rotation curves of disk galaxies [51, 52]. The masses of the central regions of galaxies similar to our own (whose mass is typical for the galaxies containing most of the visible matter in the Universe [39, 53]) can be estimated from the rotational velocity of its disk component, $V_{rot} = 250$ km/s:

$$M = \frac{V_{rot}^2 R}{G} = 10^7 R M_\odot, \tag{10}$$

where R is the distance to the center in pc. Substituting this expression into (3), we obtain for the deceleration time for a cluster with mass m_6 (in units of $10^6 M_\odot$)

$$\tau_{fr}(\mathrm{yr}) = 5 \times 10^4 R^2/m_6. \tag{11}$$

Finally, we conclude that globular clusters with masses above $\sim 10^6 M_\odot$, radii $R < 500$ pc, and circular orbits will merge into the nucleus of the parent galaxy over the Hubble time. On the other hand, the orbits of most globular clusters are probably elliptical, thereby increasing their tidal-capture zones, since the deceleration time depends strongly on the velocity [see (1)] and clusters with elliptical orbits spend most of their time at their apocenters, where they have fairly small velocities. Precisely the most massive clusters from the central regions of galaxies will initially move into the galactic nuclei, thereby affecting the observed mass function. The energy released by the deceleration of the globular clusters is absorbed by stars in the galactic nucleus, resulting in the formation of a galactic bulge; i.e., a spheroidal stellar cluster, whose mass is correlated with the mass of the central SMBH [54]. Some fraction of the stars acquire velocities exceeding the escape velocity, and so will be lost by the galaxy and form a stellar corona.

Since the observed number of globular clusters in massive galaxies can reach several thousand [49] (and these are only the clusters that have avoided merging into the nucleus of the parent galaxy), the total mass of the black hole that forms in the nucleus of a galaxy similar to our own can reach $\sim 10^6 M_\odot$. Indeed, precisely black holes with such masses are usually detected in the nuclei of nearby ($< 6 \times 10^8$ pc) massive galaxies [55]. A similar SMBH is located in the nucleus of our Galaxy; we may now be observing the merger of the next globular cluster with this central SMBH [56]. If a sufficient amount of gas is available near the SMBH, it can accrete at a high rate, resulting in a high brightness of the gas accretion disk around the black hole. This is the nature of quasars. If there is no gas, the SMBH in a galactic nucleus can be detected only via its gravitational influence on the motions of nearby stars.

4. MERGERS OF GALAXIES IN THE CORES OF DENSE GALACTIC CLUSTERS

Observations have shown that a substantial fraction of galaxies are members of galactic clusters with average masses of $\sim 10^{15} M_\odot$ and characteristic radii of about 3 Mpc [57, 58]. The typical velocities of the galaxies in clusters is $\sim$500 km/s [39]. According to (3), the characteristic time for the tidal deceleration of galaxies with masses of $\sim 10^{11} M_\odot$ in their clusters is $\sim 10^{10}$ yr. However, for the most massive galaxies, this time is less than the Hubble time. As a result, supermassive ($\sim 10^{13} M_\odot$) cD galaxies can be formed in the centers of such clusters via the merging of about 1% of the slowest, most massive central galaxies [59]. Therefore, cD galaxies with central SMBHs will appear in the cores of dense, massive galactic clusters, even at $z \sim 5$ (which corresponds to an age of 10^9 yr). Galaxies observed in the process of merging are well known [60]. Some of the mass of a cD galaxy may belong to the merging galaxies, forming a dense cluster core, and some of the stars of a cD galaxy could be formed during star formation in so-called cooling flows of the cluster gas, which flows into the cores under the action of gravity [61].

Observations indicate that massive cD galaxies are located in the cores of galactic clusters at $z = 1.24$ (which corresponds to an age of about 4×10^9 yr) [62]. We also know of nearby galaxies observed in the process of merging that possess double [63, 64], and even triple, nuclei [65]. The sequential merging of galaxies and the SMBHs in their nuclei results in the fairly rapid formation of cD galaxies with SMBHs (10^9–$10^{10} M_\odot$) in their nuclei in the young Universe with an age of only $\sim 10^9$ yr [66]. The formation of SMBHs is already appreciably accelerated at $z \sim 6.4$ by hidden mass, which should represent a considerable fraction of the gravitating matter [67]. We may be observing the merger of two SMBHs in a binary quasar with a distance between the components of only 20 kpc [68]. Some low-mass galaxies should accumulate the kinetic energy of the system during the merging of massive galaxies, so that they are ejected from the parent clusters [69]. Other galaxies will be disrupted by tidal forces, so that their stars enrich the intergalactic space of the clusters.

The merging of the central binary SMBHs occurs under the influence of gravitational-wave radiation. According to theoretical estimates, most of the flux of gravitational waves comes from close events with $z < 1$ [70]. LISA is expected to detect one event per year from merging black holes with masses of $\sim 10^7 M_\odot$ and one event per 10 years from merging black holes with masses of $\sim 10^8 M_\odot$ [70].

There is an important effect that must be taken into account when studying the merging of the binary black holes due to gravitational-wave radiation. Simple estimates show that the characteristic time for the merging of binary black holes when the components fill their Roche lobes is appreciably shorter than the orbital period. Consequently, a substantial fraction of the binding energy of the resulting black hole ($\sim 0.1 M_{BH} c^2$) is emitted as a directional pulse. Due to the conservation of momentum, the merger product can acquire a large velocity if the masses of the components are not equal [71]. As a result, some black-hole merger products can leave their parent systems, and, therefore, will not participate in the cascade mass increase. The unidirectional character of the gravitational-wave pulses leads to problems in their detection due to the presence of the cosmological horizon. A similar phenomenon occurs in the detection of gamma-ray bursts [72].

5. THE FORMATION OF ULTRAFAST STARS VIA MERGERS OF SUPERMASSIVE BLACK HOLES

Along with the generation of powerful (and, possibly, sharply directed) bursts of gravitational-wave radiation, the cascade scenario for the formation of SMBHs involves another interesting side effect: the acceleration of some fraction of the stars. This is due to the dynamic disintegration of unstable triple stars or the acceleration of stars by the nonstationary gravitational potential of a close binary SMBH. This last process results in a gravitationally bound binary system composed of two SMBHs and a solitary star moving with a velocity approximately equal to the orbital velocity of the binary components. Since the orbital velocities of SMBHs just before merging are of the order of the speed of light, the velocities of the ejected stars are limited only by the escape speed at their surface. Therefore, this mechanism can accelerate main-sequence stars to $\sim$600 km/s, degenerate dwarfs to $\sim$6000 km/s, and neutron stars and stellar black holes to $\sim 10^5$ km/s. The maximum velocities produced by this mechanism were estimated by Barkov *et al.* [72].

We can now estimate the minimum orbital velocities of the components of close binary black holes with masses m and M that merge within the Hubble time using (2): $a/R_\odot \sim 3(m/M_\odot)^{1/4}(M/M_\odot)^{1/2}$, where we assume that $m \ll M$. The orbital velocity of the component with the smaller mass will then be

$$V(\text{km/s}) \sim 250(M/M_\odot)^{1/4}(m/M_\odot)^{-1/8}. \quad (12)$$

When $M \sim 10^9 M_\odot$ and $m \sim 10^3 M_\odot$, we obtain $V \sim$ 20 000 km/s. This clearly shows the possibility of accelerating some fraction of the neutron stars and black holes to velocities above $V_{\max} \sim 1000$ km/s during mergers of binary SMBHs, which cannot be produced by the explosions of supernovas in close binary systems [2].

It is more difficult to estimate the spatial-velocity distribution of the stars ejected via the gravitational acceleration of a binary SMBH, since it depends substantially on the distribution and diffusion of stars in the galactic nucleus in velocity and in space. The merging of the components of a binary SMBH into the galactic nucleus is separated into two stages from the viewpoint of the main deceleration mechanism. The components are first decelerated by the tidal friction of the stellar environment and the ejection of stars with velocities comparable to the orbital speeds of the components of the binary SMBH [see (12)]. Much greater velocities can be gained by the stars ejected by the binary SMBH in the second stage, when the merger of the components is governed by gravitational-wave radiation. The number of stars accelerated at this stage N should clearly be proportional to the product of the lifetime of the binary SMBH in the stage with the given orbital speeds V and the cross-sectional area for the interaction: $N \sim V^{-10}$. It is obvious that the number of ejected stars depends strongly on their spatial velocity and, as a result, the number of relativistic stars—black holes and neutron stars—ejected with speeds much greater than that given by (12) may be quite small. To obtain a quantitative estimate of the velocity spectrum of ultrafast stars formed during the merging of SMBHs in galactic nuclei, we must solve the corresponding problem numerically. However, it is already clear that it should be possible to accelerate some fraction of the relativistic stars from the nuclei of merging galaxies containing SMBHs to relativistic speeds $\sim 10^5$ km/s [72].

Observational evidence for the important role of galaxy mergers in the formation of the most massive galaxies is the increase in the maximum luminosity of galaxies from $\sim 10^{10} L_\odot$ to $\sim 10^{11} L_\odot$ as their age increases from 1 to 7 billion years ($1 < z < 4$) [73]. Most collisions occurs at $z \sim 1$, but collisions are fairly frequent even at the present time [74]. Galaxies with characteristic surface densities of $\sim 100-1000 M_\odot/\mathrm{pc}^2$ in clusters with masses of $\sim 10^{15} M_\odot$ and radii of about 1 Mpc should inevitably experience collisions, as follows from the total volume covered by such galaxies during the Hubble time. Simple estimates indicate that this area is much greater than the cluster size, resulting in unavoidable collisional evolution of the galaxies in the clusters, and, consequently, mergers of some of the galaxies that lead to the growth of the masses of the black holes in their nuclei. Merging of galaxies leads to the formation of spheroidal stellar halos surrounding them. As was noted above, one side product of the merging of SMBHs in galactic nuclei should be the production of a certain number of ultrafast neutron stars, stellar black holes, and degenerate dwarfs in the Universe. Due to their large velocities (10^3-10^5 km/s), these can fill both galaxy clusters and the space between them almost uniformly.

6. CONCLUSION

The aim of this paper was to analyze the cascade (hierarchical) mechanism for the formation and growth of supermassive black holes in the cores of globular clusters, galaxies, and supermassive cD galaxies in the cores of rich and dense galactic clusters. There are two stages of merging. In the first stage, the main decelerating mechanism is tidal friction by the background stars (in globular clusters) or galaxies (in galactic clusters). In the second stage, the main mechanism is the radiation of gravitational waves. The formation of SMBHs is facilitated by the fact that only a few percent of the total number of stellar black holes participate in this process: the slowest black holes and SMBHs, whose mergers require short times (probably less than $\sim 10^9$ yr). This enables us to explain the existence of merging galaxies and the formation of SMBHs in their nuclei already at $z \sim 1-2$ [75]. It is important to emphasize that the cascade mechanism for the formation of SMBHs in galactic nuclei can explain the observed relation between the black-hole mass and the stellar-velocity disperson in galactic nuclei ($M_{BH} \sim v^4$), as well as between the black-hole mass and the masses of galactic bulges ($M_{BH} \sim M_b$) [76].

Mergers of galaxies result in changes in their morphology (growth in the mass of the spherical component) and increase the accretion activity in their nuclei. One particular observational example of merging is M87, the central galaxy of the Virgo cluster, whose complex nuclear-gas morphology probably resulted from a recent merger of two galaxies [77]. The nearest merging SMBHs may be located in the center of our own Galaxy, but are hidden in the optical by absorption ($\sim 38.5^m$) [78]. A young, massive stellar cluster, marked by a dense (<0.02 pc) group of OB stars, is probably in the stage of merging into the nuclear stellar population of the Galaxy. The young cluster is assumed to be bound by a SMBH ($\sim 1300 M_\odot$). Another consequence of galaxy mergers is the appearance of a considerable number of intergalactic stars (2 to 20%) [79] that are ejected from the parent galaxies when the merger products relax. Observations show that these stars are among the oldest, and are concentrated toward the central cD galaxies, where the frequency of collisions between the galaxies is high due to the enhanced density.

A SMBH in a galactic nucleus can be revealed as an X-ray source if there is a source of gas that can be accreted at a rate sufficient to produce X-rays that are detectable at cosmological distances. The gas-accretion rate of active nuclei is usually an appreciable fraction of the accretion rate corresponding to their Eddington luminosity. The accreted gas can be supplied either by stars close to the SMBH or by the gaseous disk component of the galaxy or its nucleus. In this last case, the accretion rate increases if the galaxy collides with another or has a bar. Central massive cD galaxies have cooling flows in which intergalactic gas flows into their nuclei at a rate $100-1000 M_\odot/\mathrm{yr}$. The accretion of gas by a SMBH results in the formation of a quasar. It is more difficult to produce a high accretion rate in the cores of globular clusters, which have very little gas, making the reliable identification of SMBHs in globular clusters more difficult [28]. This obvious observational selection effect can probably explain the lack of detection of black holes with masses of $\sim 10^3 M_\odot$, which appears as an absence or very low occurrence of such black holes. On the other hand, it is possible that bright, soft X-ray sources (0.1–0.4 keV) [80] represent precisely this kind of SMBH. The reconstruction of the corresponding mass spectrum is an important task for the future.

It is interesting that the compact (<100 pc) gas disks around SMBHs in the nuclei of Seyfert and other active galaxies not only produce the activity of their nuclei, but also often show signs of active star formation [81]. For example, about 10% of the massive stars of our Galaxy are currently being formed within only about 50 pc of the Galactic center [82]. The formation of stars in a galactic nucleus results in the appearance of dense stellar clusters [83], in whose cores SMBHs can be formed and later merge with the central black hole. Many elliptical galaxies exhibit a substantial blue excess, indicating the presence of active star formation in their nuclei [84]. This, first and foremost, provides evidence for the presence of gas. In addition, star formation in a gaseous disk near the nucleus and the explosions of short-lived massive stars as supernovas turbulize the disk, which amplifies the accretion of gas onto the central SMBH. As a result, the mass of the black hole increases with time and, more importantly, the features of its activity become observable. Bursts of star formation in galactic nuclei may be recurrent, thereby resulting in periodicity of the accretion activity of SMBHs [85].

The cascade mechanism for the formation of SMBHs in the nuclei of globular clusters and galaxies may also have some morphological consequences for the distribution of the old stars in them. The deceleration of stellar black holes in globular clusters results in their expansion and acquisition of a spherical shape with time [41, 86]. Some fraction of the stars should be lost in this process [87]. The deceleration of globular clusters and galaxies in the more massive systems produces spheroidal stellar components in these systems, again with some stars being lost. As a result, the galactic bulges may represent the products of the cascade mechanism for the formation of SMBHs. This may provide an explanation for the observed correlation between the masses of the SMBHs and the spheroidal stellar components of galaxies.

An important effect of the cascade mechanism for the formation of SMBHs is the generation of powerful gravitational-wave radiation. The characteristic period of these waves is determined by the mass of the merging SMBHs, and is equal to $\sim 10^{-4} M_{BH}/M_\odot$ s. If we take the typical mass of a black hole to be $M_{BH} \sim 10^5 M_\odot$, the period of the gravitational waves will be about 10 s, and the energy released by the merger of two such SMBHs will be $E_{GWR} \sim 0.1 M_{BH} c^2 \sim 10^{56}$ erg. It is quite possible that the first detectable gravitational waves will be short ($\sim 10^{-4} M_{BH}/M_\odot$ s) bursts resulting from mergers of SMBHs in the nuclei of sufficiently close merging galaxies. To elucidate the fate of the products of SMBH mergers, it is important to find the amplitude of the kick experienced by the merger product due to the sharp directivity of the gravitational-wave pulse [71].

Another consequence of the cascade mechanism for the formation of SMBHs that should be mentioned here is the inevitable appearance of ultrafast relativistic stars (degenerate dwarfs, neutron stars, and stellar-mass black holes). The velocities of the last two types of object can reach $\sim 10^5$ km/s. The more precise formulation of the requirements for detecting these ultrafast stars requires numerical modeling and estimation of their velocity spectrum. Any binary systems containing such objects will be disrupted when they are accelerated due to their weak gravitational binding. This complicates searches for such stars, since a number of nonstationary features typical of close binaries will be excluded. At present, only long-lived single millisecond radio pulsars are promising candidates for detecting ultrafast relativistic stellar objects.

ACKNOWLEDGMENTS

We are grateful to Prof. G.S. Bisnovatyĭ-Kogan for fruitful discussions of the proposed scenario. This work was supported by the Program of Support for Leading Scientific Schools of Russia (project no. 00-15-96533), the Russian Foundation for Basic Research (project nos. 02-02-17524 and 03-02-162254), NNIO-02-02-04008, and the Federal Science and Technology Program in Astronomy.

REFERENCES

1. A. Tutukov and A. Cherepashchuk, Astron. Zh. **80**, 419 (2003) [Astron. Rep. **47**, 386 (2003)].
2. A. Tutukov and L. Yungel'son, Nauchn. Inform. Astrosoveta Akad. Nauk USSR **27**, 3 (1973).
3. R. Humphreys, Astrophys. J., Suppl. Ser. **29**, 389 (1975).
4. O. Figer, R. Rich, and S. Kim, astro-ph/0309757.
5. E. Colbert and R. Mushotzky, Astrophys. J. **519**, 89 (1999).
6. G. Fabbiano *et al.*, Astrophys. J. Lett. **584**, L5 (2003).
7. A. Tutukov and A. Fedorova, Astron. Zh. **81**, 589 (2004) [Astron. Rep. **48**, 534 (2004)].
8. I. Ho, Astrophys. J. **564**, 120 (2002).
9. I. Gioia, J. Henry, C. Mullis, *et al.*, astro-ph/0309788.
10. A. Marconi, G. Risaliti, and R. Gilli, astro-ph/0311619.
11. J. Hamilton, S. Casertano, and D. Turnishek, Astrophys. J. **576**, 61 (2002).
12. C. Warner, F. Hamman, and M. Dietrich, astro-ph/0307247.
13. M. Elvis, G. Risaliti, and G. Zamorani, Astrophys. J. **565**, L75 (2002).
14. K. R. Lang, *Astrophysical Data: Planets and Stars* (Springer, New York, 1991).
15. A. G. Masevich and A. V. Tutukov, *Stellar Physics and Evolution: Investigation of the Cosmic Space* (VINITI, Moscow, 1981), Vol. 17 [in Russian].
16. A. Finogenov, A. Burkert, and H. Bohringer, Astrophys. J. **594**, 136 (2003).
17. D. Wiebe, A. Moiseev, A. Tutukov, *et al.*, Astron. Zh. **81**, 403 (2004) [Astron. Rep. **48**, 367 (2004)].
18. L. Smith and J. Gallagher, Mon. Not. R. Astron. Soc. **326**, 1027 (2001).
19. J. Wyithe, B. Stuart, and A. Loeb, Astrophys. J. **588**, L69 (2003).
20. I. Iben and A. Tutukov, Astrophys. J. **282**, 615 (1984).
21. J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, 1987), p. 425.
22. K. Gebhardt, R. Rich, and L. Ho, Astrophys. J. **578**, L41 (2002).
23. J. Gerssen, R. Marel, K. Gebhardt, *et al.*, Astron. J. **125**, 376 (2002).
24. R. Humphreys, G. Fabbiano, M. Elvis, *et al.*, Mon. Not. R. Astron. Soc. **344**, 134 (2003).
25. K. Menou, Z. Haiman, and V. Narayanan, Astrophys. J. **558**, 535 (2001).
26. P. Zwart, F. Simon, and L. Stephen, Astrophys. J. **576**, 899 (2002).
27. J. Fregeau, M. Gurkan, K. Joshi, *et al.*, Astrophys. J. **593**, 772 (2003).
28. G. S. Bisnovatyĭ-Kogan, Pis'ma Astron. Zh. **4**, 130 (1978) [Sov. Astron. Lett. **4**, 69 (1978)].
29. A. V. Tutukov and L. R. Yungelson, Mon. Not. R. Astron. Soc. **260**, 675 (1993).
30. A. V. Tutukov and L. R. Yungelson, Astron. Zh. **79**, 738 (2002) [Astron. Rep. **46**, 667 (2002)].
31. H. Lamers, N. Panagia, S. Scuder, *et al.*, Astrophys. J. **566**, 818 (2002).
32. S. Zwart, H. Baumgardt, P. Hut, *et al.*, astro-ph/0402622.
33. M. Miller, M. Coleman, D. Hamilton, *et al.*, Mon. Not. R. Astron. Soc. **330**, 232 (2002).
34. A. G. Masevich and A. V. Tutukov, *Stellar Evolution. Theory and Observations* (Nauka, Moscow, 1988) [in Russian].
35. C. Vignali, astro-ph/0403100.
36. A. Finogenov, U. Briel, J. Henry, *et al.*, astro-ph/0403216.
37. J. Wyithe and A. Loeb, Astrophys. J. **595**, 614 (2003).
38. A. Tutukov and A. Cherepashchuk, Astron. Zh. **80**, 419 (2003) [Astron. Rep. **47**, 386 (2003)].
39. C. W. Allen, *Astrophysical Quantities* (Athlone Press, London, 1973; Mir, Moscow, 1977).
40. Yo. Taniguchi, Ya. Shiova, T. Tsuru, *et al.*, Publ. Astron. Soc. Jpn. **52**, 533 (2000).
41. D. Merritt, S. Piatek, S. Zwart, *et al.*, astro-ph/0403331.
42. B. C. Whitemore, astro-ph/0403709.
43. Q. Wang, T. Yao, W. Fukui, *et al.*, astro-ph/0403413.
44. M. Gilfanov, astro-ph/0403552.
45. P. Freier, M. Kramer, A. Lyne, *et al.*, Astrophys. J. Lett. **557**, L105 (2001).
46. G. Knapp, J. Gunn, P. Bowers, *et al.*, Astrophys. J. **462**, 231 (1996).
47. D. Spergel, Nature **352**, 221 (1991).
48. C. Conselice, M. Bershady, and M. Dickinson, Astron. J. **126**, 1183 (2003).
49. K. Rhode and S. Zepf, Astron. J. **126**, 2307 (2003).
50. S. Lipsey and P. Plavchan, astro-ph/0311297.
51. S. Tremaine, J. Ostriker, and L. Spitzer, Astrophys. J. **196**, 407 (1975).
52. V. Surdin and N. Arkhipova, Pis'ma Astron. Zh. **24**, 407 (1998) [Astron. Lett. **24**, 343 (1998)].
53. S. Anderson, J. Cuillandre, and R. Pello, astro-ph/0009431.
54. R. McLure, J. Dunlope, and C. Conselice, astro-ph/0312350.
55. I. Gioia, J. Henry, C. Mullis, *et al.*, astro-ph/0309788.
56. J. Maillard, T. Paumard, S. Stolovy, *et al.*, astro-ph/0404450.
57. M. Girardi, S. Borgani, and G. Giuricin, Astrophys. J. **506**, 45 (1998).
58. A. Sanchez, N. Padilla, and D. Lambas, Mon. Not. R. Astron. Soc. **337**, 161 (2002).
59. L. Koopmans and T. Tren, Astrophys. J. Lett. **568**, L5 (2002).
60. R. Ibata, G. Gilmore, and M. Irwin, Nature **370**, 194 (1994).
61. A. Jordan, P. Cote, M. West, *et al.*, Astron. J. **127**, 24 (2004).

62. M. Beasley, D. Forby, J. Brodie, *et al.*, astro-ph/0312354.
63. J. Gersen, astro-ph/0310029.
64. F. Liu, X. Wu, and S. Cao, astro-ph/0310045.
65. M. Seigar, Mon. Not. R. Astron. Soc. **344**, 110 (2003).
66. A. Barth, P. Martini, C. Nelson, *et al.*, astro-ph/0308005.
67. F. Munyaneza and P. Biermann, astro-ph/0403511.
68. H. Zhou, T. Wang, X. Zhang, *et al.*, Astrophys. J. **604**, L33 (2004).
69. B. Bingelli, E. Barrazo, and H. Jerjen, Astron. Astrophys. **359**, 447 (2000).
70. M. Enoki, K. Inoue, M. Nagashima, *et al.*, astro-ph/0404389.
71. P. Madau and E. Quataert, Astrophys. J. **606**, L17 (2004).
72. M. Barkov, V. Belinski, and G. S. Bisnovatyi-Kogan, Mon. Not. R. Astron. Soc. **334**, 338 (2002).
73. A. Fontana, I. Donnazumma, E. Vanzella, *et al.*, Astrophys. J. **594**, L9 (2003).
74. Ch. Conselice, M. Bershady, and M. Dickinson, Astron. J. **126**, 1183 (2003).
75. H. Chen, astro-ph/0312173.
76. M. Haehnelt and G. Kauffmann, Mon. Not. R. Astron. Soc. **318**, 35 (2000).
77. H. Feng, S. Zhang, Y. Lou, *et al.*, astro-ph/0404377.
78. J. Maillard, T. Paumard, S. Stolovy, *et al.*, astro-ph/0404450.
79. B. Willman, F. Governato, J. Wadsley, *et al.*, astro-ph/0405094.
80. R. DiStefano, F. Primini, A. Kong, *et al.*, astro-ph/0405238.
81. M. Imanishi, astro-ph/0309083.
82. D. Figer, astro-ph/0403088.
83. T. Boker, M. Sazzi, A. McLaughlin, *et al.*, astro-ph/0309761.
84. J. Kotilainen and R. Falomo, astro-ph/0404372.
85. E. Kruegel and A. Tutukov, Astron. Astrophys. **275**, 416 (1993).
86. G. S. Bisnovatyĭ-Kogan, M. Merafina, R. Ruffini, and E. Vesperini, Astrophys. J. **500**, 217 (1998).
87. H. Baumgardt, J. Makino, and T. Ebisuzaki, astro-ph/0406231.

Translated by Yu. Dumin

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 24–35. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 29–40.
Original Russian Text Copyright © 2005 by Barkov, Bisnovatyĭ-Kogan.

Interaction of a Cosmological Gamma-ray Burst with a Dense Molecular Cloud and the Formation of Jets

M. V. Barkov and G. S. Bisnovatyĭ-Kogan

Space Research Institute, Moscow, Russia

Received April 5, 2004; in final form, July 15, 2004

Abstract—The interaction of a powerful cosmological gamma-ray burst (GRB) with a dense molecular cloud is modeled. Two-dimensional gas–dynamical flows were computed for various configurations of the cloud. In the spherically symmetrical case, the gas velocity does not exceed $\sim 2 \times 10^3 \sqrt{E/1.6 \times 10^{53}}$ km/s. If the GRB precursor has an anisotropic wind, a conical cavity can form in the nearby region of the molecular cloud. The propagation of the gamma-ray pulse in this cavity leads to the formation of a rapidly moving hot clump of matter, with the gas velocity reaching 1.8×10^4 km/s for gamma-ray energy of $E = 1.6 \times 10^{53}$ erg. In all the computations, the velocity of the moving material is much lower than the velocity of light, the volume of gas affected by the motion is small, and the influence of the gas motions on the light curve of the optical afterglow is insignificant. *© 2005 Pleiades Publishing Inc.*

1. INTRODUCTION

The nature of gamma-ray bursts (GRBs) remains unclear, in spite of the fact that more than 30 years have passed since their discovery [1]. The most complete statistics on GRBs have been obtained with the BATSE experiment [2–4]. The detected GRBs are distributed uniformly on the sky, and their $\log N - \log S$ curve shows appreciable deviations from the law $N \sim S^{3/2}$, which corresponds to a uniform spatial distribution for the burst sources [2, 5]. The BATSE data have also yielded the distribution of durations for the bursts, which displays a clear two-peaked structure, with the dip between the peaks being near 1.5 s. Further, we will consider only long bursts (with durations exceeding one second). Most GRBs have durations less than 100 s, although several bursts with durations of about 1000 s have been detected.

The BeppoSAX experiment has enabled us to appreciably improve our understanding of GRBs [6]. Although its limiting sensitivity was substantially worse that of BATSE, it made it possible to obtain X-ray observations of GRB sources within a short time after their detection, enabling more accurate localization of the GRB coordinates. The first successfully localized GRB was GRB 970228 [7, 8], whose coordinates were sufficiently accurate to enable the identification of the GRB with a flaring optical source (transient) [9]. Nebulous optical emission was also observed around the optical transient [9, 10], which proved to be a distant galaxy with a cosmological redshift, $z = 0.695$.

Observations of the optical afterglows of GRBs and measurements of their redshifts demonstrated that GRBs are cosmological; i.e., they are located in distant ($\sim 10^3$ Mpc) galaxies. In this case, the release of a huge amount of gamma-ray energy ($\sim 10^{52-54.5}$ erg) in a short time (tens of seconds) is required to explain the observed gamma-ray fluxes $\sim 10^{-4}$ erg s^{-1} cm^{-2}. It is obvious that such a powerful release of energy should strongly infuence an appreciable volume of the host galaxy. Our main aim in the current paper is to demonstrate that this is the case.

Studies of optical afterglows at various wavelengths led to the conclusion that GRBs occur in dense interstellar clouds harboring star-forming regions [11, 12]. Without specifying a particular mechanism for the GRBs, we will assume only the presence of a powerful flux of gamma-rays and consider the interaction of this radiation with the dense interstellar medium of the host galaxy in the form of a molecular cloud on large spatial and temporal scales. This situation was first considered in [13–15], where solutions for a one-dimensional molecular cloud in a spherically symmetrical approximation were obtained.

Here, we investigate the response of an interstellar medium with normal chemical composition to the passage through this medium of a short pulse of gamma-rays in the axially symmetrical case. We consider various density distributions in the medium, GRB energies, and degrees of collimation for the GRB, including the case when the GRB is located at the edge of a molecular cloud. The hydrodynamical flow of the interstellar matter was computed based on a two-dimensional Piecewise-Parabolic-Method

(PPM) code. We consider the case when a directed ejection (jet) with a speed of the order of several thousand km/s forms as a result of heating of the cloud by the GRB radiation.

2. MAIN EQUATIONS AND CONDITIONS

The system of hydrodynamical equations describing the motion of matter under the action of a powerful pulse of radiation in the axially symmetrical case was solved in spherical coordinates:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0, \tag{1}$$

$$\frac{\partial(\rho v_r)}{\partial t} + \frac{\partial(\rho v_r^2 + P)}{\partial r} + \frac{1}{r}\frac{\partial(\rho v_r v_\theta)}{\partial \theta} + \frac{2\rho v_r^2 - \rho v_\theta^2 + \rho v_r v_\theta \cot\theta}{r} = \rho F_\gamma, \tag{2}$$

$$\frac{\partial(\rho v_\theta)}{\partial t} + \frac{\partial(\rho v_r v_\theta)}{\partial r} + \frac{1}{r}\frac{\partial(\rho v_\theta^2 + P)}{\partial \theta} + \frac{3\rho v_r v_\theta + \rho v_\theta^2 \cot\theta}{r} = 0, \tag{3}$$

$$\frac{\partial}{\partial t}\left(\frac{\rho v^2}{2} + \rho\varepsilon\right) + \nabla\left\{\rho \mathbf{v}\left(\frac{v^2}{2} + \varepsilon + \frac{P}{\rho}\right)\right\} = \rho H_\gamma - \rho C_\gamma. \tag{4}$$

Here, ρ is density, P is pressure, ε is internal energy, and v_r and v_θ are the radial and transverse velocities of the matter, respectively. The matter was assumed to be fully ionized. This assumption is justified by the fact that the velocity in the unheated gas is low (a few km/s), as well as the pressure. After the passage of the gamma-ray pulse, the matter is heated to high temperatures, and it can be considered to be fully ionized.

Since the time scales of interest to us ($\sim 10^5$–10^8 s) are much longer than the burst durations (<100 s), we can take the GRB to be an instantaneous pulse with energy Γ:

$$L = \Gamma\delta\left(t - \frac{r}{c}\right). \tag{5}$$

For gamma-rays with $h\nu \gg B_e^{a,i}$, where $B_e^{a,i}$ is the binding energy of an electron in an atom or ion, the cross section for the interaction between gamma-ray photons and electrons will be nearly the same for bound and free electrons. The rate of gamma-ray heating per unit mass of gas H_γ was taken from [16–18]. The GRB spectrum was taken to have the form $\frac{dL}{dE} = \frac{L}{E_{max}} e^{-E/E_{max}}$. Introducing the notation $W(E, E_{max}) = \frac{dL}{LdE}$, we have

$$H_\gamma = \frac{L}{4\pi r^2} \times \frac{\mu_e \sigma_T}{m_u} \frac{E_{max} f_h(E_{max}) - 4kT f_c(E_{max})}{m_e c^2}, \tag{6}$$

where E is the energy of the photons, m_e is the mass of an electron, and σ_T is the Thomson cross section for the interaction between an electron and a photon. We used $E_{max} = 2$ MeV and $E_{max} = 0.6$ MeV. The functions

$$f_h(E_{max}) = \frac{1}{E_{max}} \int_0^\infty W(E, E_{max}) s(E) E dE \tag{7}$$

and

$$f_c(E_{max}) = \int_0^\infty W(E, E_{max}) q(E) dE \tag{8}$$

are obtained by integrating over the spectrum. The functions $s(E)$ and $q(E)$ presented in [16] take into account the dependence of the cross section for the relativistic scattering of a photon on an electron on the photon energy (the Klein–Nishina cross section σ_{KN}). The normalization was chosen so that $f_h = f_c = 1$ when $E_{max} \ll m_e c^2$. For an energy of 600 keV, $f_h = 0.19$ and $f_c = 0.33$, while $f_h = 0.065$ and $f_c = 0.16$ for an energy of 2 MeV.

The radiative force of the gamma-ray pulse acting on a gram of matter, F_γ, is determined by the light pressure due to the gamma-rays interacting with matter with normal chemical composition via electron scattering with the cross section σ_{KN}:

$$F_\gamma = \frac{1}{c}\frac{L}{4\pi r^2}\frac{\mu_e \sigma_T}{m_u} f_f(E_{max}), \tag{9}$$

where L is the "luminosity" of the GRB source, μ_e is the number of electrons per nucleon (for a medium with normal chemical composition, $\mu_e = 0.866$), and c is the speed of light. The function $f_f(E_{max})$ takes into account variations in the electron–gamma-ray cross section associated with relativistic corrections [19]:

$$f_f(E_{max}) = \frac{1}{\sigma_T} \int_0^\infty W(E, E_{max}) \sigma_{KN}(E) dE. \tag{10}$$

For energies of 600 keV and 2 Mev, $f_f = 0.5$ and 0.32, respectively.

According to [20, 21], the cooling rate per unit mass of optically thin plasma, C_γ, is given by the relation

$$C_\gamma = \frac{\Lambda(T)n^2}{\rho}. \tag{11}$$

To increase the computational speed of the program, the cooling rate from [20] was approximated using the formulas

$$\Lambda(T) = \begin{cases} 0, \ T < 10^4 K; \\ 10^{-48.8}T^{6.4}, \ 10^4 < T < 10^{4.25}; \\ 10^{-16.5}T^{-1.2}, \ 10^{4.25} < T < 10^{4.5}; \\ 10^{-27.48}T^{1.24}, \ 10^{4.5} < T < 10^5; \\ 10^{-21.03}T^{-0.05}, \ 10^5 < T < 10^{5.4}; \\ 10^{-13.6698}T^{-1.413}, \ 10^{5.4} < T < 10^{5.86}; \\ 10^{-22.8378}T^{0.1515}, \ 10^{5.86} < T < 10^{6.19}; \\ 10^{-13.1969}T^{-1.406}, \ 10^{6.19} < T < 10^{6.83}; \\ 10^{-22.2877}T^{-0.075}, \ 10^{6.83} < T < 10^{7.5}; \\ 10^{-26.6}T^{0.5}, \ 10^{7.5} < T. \end{cases} \tag{12}$$

The computations assigned solar chemical composition to the interstellar medium [22] and assumed that all heavy elements were in gaseous form.

It was shown in [13] that the phase velocity of the cooling-wave front is of the order of (or higher than) the speed of light. We found that a zone with a temperature inversion develops in the heated region, and this region broadening due to cooling waves. In the spherically symmetrical case, two waves travel through the gas: one toward the edge of the cloud, following the heating wave, and the second toward the cloud center. The cooling wave moving toward the periphery of the cloud always travels with a phase speed that exceeds the speed of light, while the second front travels with a gradually decreasing phase velocity. The pattern of the cooling-wave motions is more complex in the non-spherically symmetrical case. However, the speed of the cooling wave exceeds the sound speed at the cooling-wave front for 10–20 years after the time of the GRB. Due to the superluminal speed of the outwardly propagating cooling front, it remains thin, and is not smeared by thermal conduction or other transport processes. Let us estimate the infuence of thermal conduction on the inwardly propagating cooling wave. In the case of constant heat conductivity χ, the heat-conduction equation has the form [23]

$$\frac{\partial T}{\partial t} = \chi \triangle T. \tag{13}$$

The heat conductivity of gases is approximately equal to the diffusion coefficient:

$$\chi = l_a \bar{v}/3, \tag{14}$$

where l_a is the mean free path and $\bar{v}$ is the thermal velocity. The rate of propagation of a thermal front via thermal conduction is [23]

$$\frac{dx}{dt} \sim \frac{d}{dt}\sqrt{\chi t} \sim \sqrt{\frac{\chi}{t}} \sim \frac{\chi}{x} \sim \frac{l_a}{x}\bar{v}. \tag{15}$$

Under the conditions in the interstellar medium, the mean free path of an electron is determined by the Larmor radius. We have for a field of $H \sim 10^{-6}$ G

$$l_a \sim r_L = \frac{\bar{v}}{\omega_B} = \frac{m_e c\bar{v}}{eB} \approx 2.9 \times 10^7 \text{ cm}. \tag{16}$$

We adopt for x the estimated thickness of the cooling front (12) obtained without including thermal conduction:

$$x = \frac{E}{dE/dt} v_{cw} \sim \frac{3/2kTc}{\Lambda(T)n} \approx 6 \times 10^{16} \frac{T_6}{n_5}, \tag{17}$$

where v_{cw} is the velocity of the cooling wave. Taking a temperature of $T = 3 \times 10^5$ K and a number density of 10^5 cm^{-3} yields $x \approx 2 \times 10^{16}$ cm. Even if we take the speed of propagation of the cooling wave to be equal to the sound speed at the front, v_s, we have

$$v_{cw} \approx v_s \approx \bar{v} \gg \frac{l_a}{x}\bar{v} \approx 1.5 \times 10^{-9}\bar{v}, \tag{18}$$

i.e., electron thermal conduction can be disregarded under the conditions relevant to our problem.

A program based on a fourth-order difference scheme for a Godunov PPM aproximation was written to solve the system of equations (1)–(4) [24, 25]. This program is able to solve this system of gas–dynamical equations in a two-dimensional approximation. The use of a complex numerical algorithm is due to the need to carry out thorough computations of the shock front that arises in the inner region of the cloud.

3. PHYSICAL FORMULATION OF THE PROBLEM AND INITIAL CONDITIONS

In order to more fully understand the gas–dynamical flows arising in a molecular cloud heated by a nearby GRB, we investigated models with a whole series of initial conditions, from a simple, spherically symmetrical, uniform molecular cloud and an isotropic gamma-ray pulse to the case of a conical cavity and an anisotropic (well collimated) gamma-ray pulse.

We investigated models corresponding to the following sets of initial conditions.

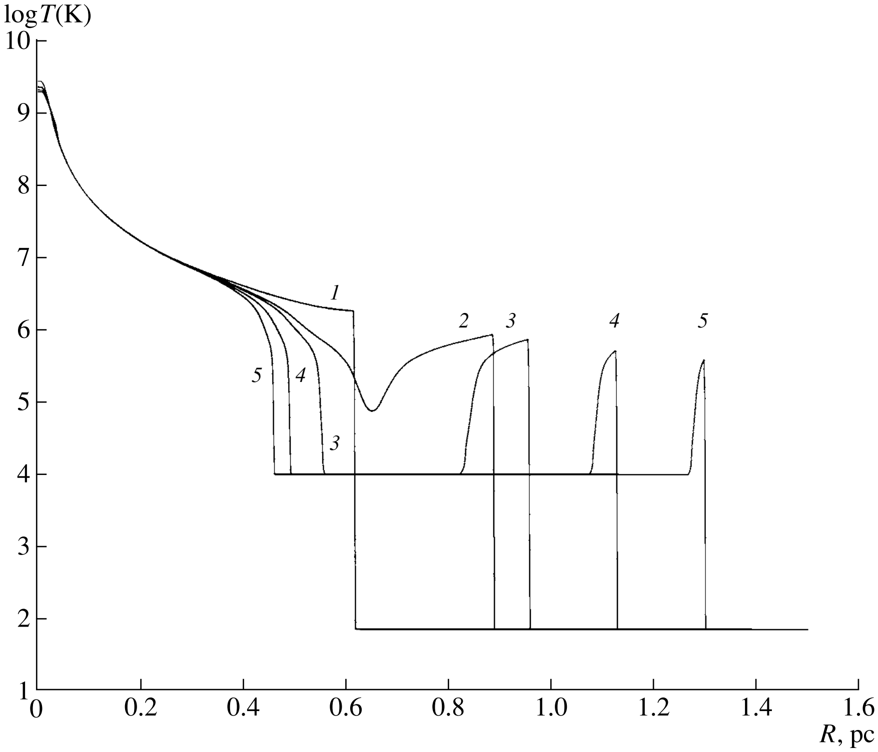


Fig. 1. Dependence of the temperature in the cloud on the distance to the GRB source at various times for model 1a; the numbered curves correspond to times after the GRB of (*1*) 2.02, (*2*) 2.91, (*3*) 3.14, (*4*) 3.70, and (*5*) 4.26 yrs.

(1) An isotropic GRB occurs at the center of a molecular cloud with radius $R = 1.5$ pc and constant density $n_H = 10^5$ cm^{-3}, and has an energy of

(a) 1.6×10^{53} erg for $E_{max} = 2$ MeV or 1.7×10^{53} erg for $E_{max} = 0.6$ MeV;

(b) 1.6×10^{54} erg for $E_{max} = 2$ MeV or 1.7×10^{54} erg for $E_{max} = 0.6$ MeV.

(2) An isotropic GRB occurs in a molecular cloud whose density falls with radius according to the law $n_H = 10^5 e^{-r^2/r_0^2}$ cm^{-3}, where $r_0 = 0.75$ pc and the radius of the computational domain is $R = 1.5$ pc. The GRB has an energy of 1.6×10^{53} erg for $E_{max} = 2$ MeV or 1.7×10^{53} erg for $E_{max} = 0.6$ MeV, and is located at the center of the molecular cloud.

(3) An isotropic GRB occurs in a molecular cloud consisting of two regions. In a large volume of the cloud, the density is constant and equal to $n_H = 10^5$ cm^{-3} (this corresponds to the condition $r \cos\theta < x_0 = 0.15$ pc). The second region is separated from the first by a plane located a distance x_0 in parsecs from the GRB (this corresponds to the condition $r \cos\theta > x_0 = 0.15$ pc), and has a density of $n_H = 10^4$ cm^{-3}. The radius of the computational domain is $R = 1.5$ pc. The GRB has an energy of 1.6×10^{53} erg for $E_{max} = 2$ MeV or 1.7×10^{53} erg for $E_{max} = 0.6$ MeV.

(4) An isotropic GRB occurs in a molecular cloud whose density falls off with radius according to the law $n_H = 10^5 e^{-r^2/r_0^2}$ cm^{-3}, where $r_0 = 0.75$ pc. The GRB has an energy of 1.6×10^{53} erg for $E_{max} = 2$ MeV or 1.7×10^{53} erg for $E_{max} = 0.6$ MeV, and is located a distance $x_0 = 1.0$ pc from the center of the cloud. The radius of the computational domain is $R = 2.5$ pc. This case is similar to model 2, but with the GRB at some distance from the center of the cloud.

(5) A GRB has an energy of 1.6×10^{53} erg for $E_{max} = 2$ MeV or 1.7×10^{53} erg for $E_{max} = 0.6$ MeV and occurs in a molecular cloud with the density distribution $n = 10^5 e^{-2-2\cos(10\theta)}$ cm^{-3}. The computational domain is restricted to $0 \le \theta \le \pi/10$ and $R = 1.5$ pc. The condition $v_\theta = 0$ is imposed at the boundary $\theta = \pi/10$. This model corresponds to a somewhat artificial conical density distribution, which could imitate a GRB occurring between two nearby molecular clouds or a GRB precursor with an anisotropic wind that "blew out" a cavity in the molecular cloud.

(6) An anisotropic GRB occurs in a molecular cloud with radius $R = 1.5$ pc and constant density $n_H = 10^5$ cm^{-3}, with the energy distributions

(a) $E = 1.6 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV

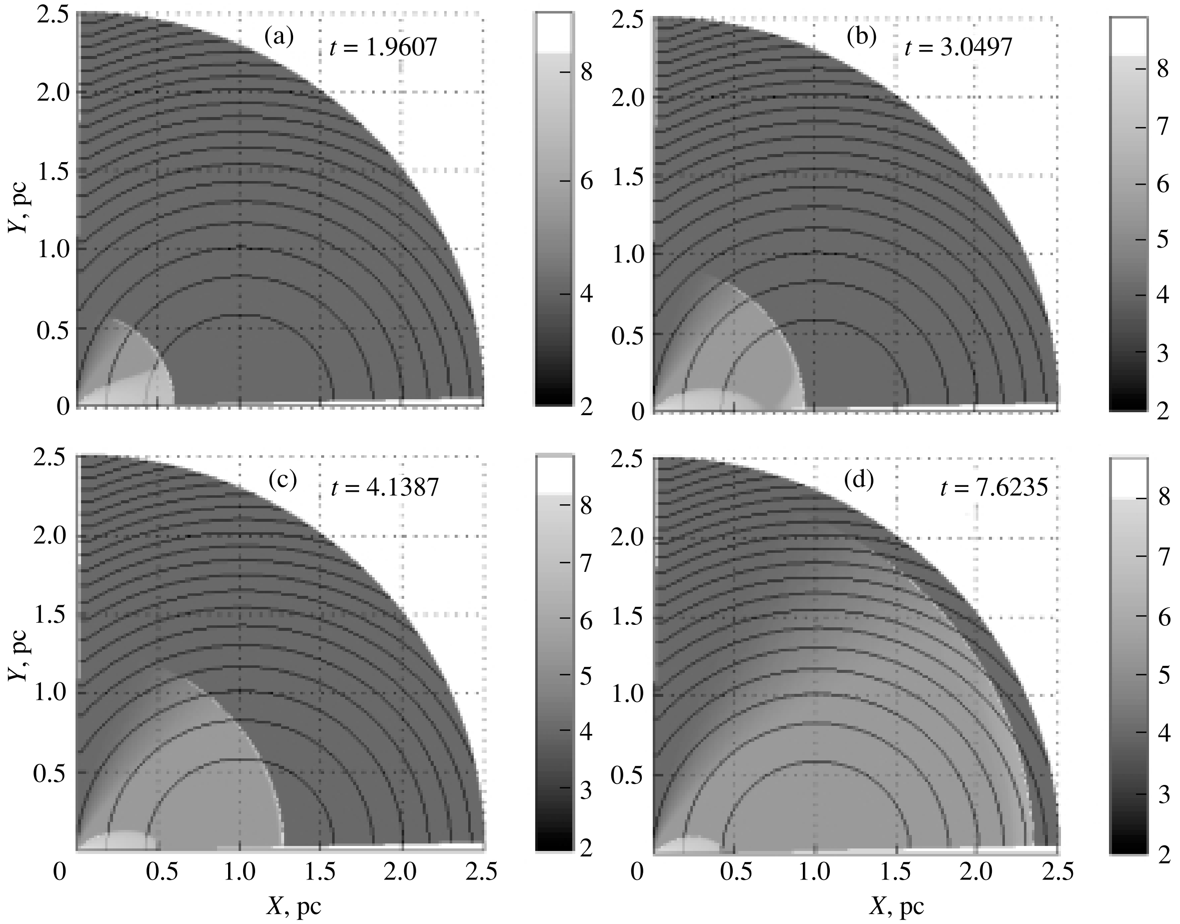


Fig. 2. Distribution of temperature (grey scale) and contours of constant density for the cloud at various times in model 8a. The distance from the GRB source is indicated in parsecs and time is measured in years since the GRB. The right-hand bars give the log T(K) temperature scale.

or $E = 1.7 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ Mev, with $\theta_0 = 0.4$ rad;

(b) $E = 1.6 \times 10^{54} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV or $E = 1.7 \times 10^{54} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ Mev, with $\theta_0 = 0.4$ rad.

Due to the energy distribution in the pulse, the computational domain was restricted to $0 \leq \theta \leq \pi/2$ and $R = 1.5$ pc.

(7) An anisotropic GRB occurs in a molecular cloud whose density falls off with radius as $n_H = 10^5 e^{-r^2/r_0^2}$ cm^{-3}, where $r_0 = 0.75$ pc. The GRB is located at the center of the cloud, and has the energy distributions $E = 1.6 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV or $E = 1.7 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ MeV, with $\theta_0 = 0.4$ rad. Due to the energy distribution in the pulse, the computational domain was restricted to $0 \leq \theta \leq \pi/2$ and $R = 1.5$ pc.

(8) An anisotropic GRB occurs a distance $x_0 = 1.0$ pc from the center of a molecular cloud whose density falls off with radius as $n_H = 10^5 e^{-r^2/r_0^2}$ cm^{-3}, where $r_0 = 0.75$ pc. The GRB has the energy distributions

(a) $E = 1.6 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV or $E = 1.7 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ MeV, with $\theta_0 = 0.4$ rad;

(b) $E = 1.6 \times 10^{54} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV or $E = 1.7 \times 10^{54} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ MeV, with $\theta_0 = 0.4$ rad. Due to the energy distribution in the pulse, the computational domain was restricted to $0 \leq \theta \leq \pi/2$ and $R = 2.5$ pc.

(9) An anisotropic GRB occurs in a molecular cloud with the density distribution $n = 10^5 e^{-2-2\cos(10\theta)}$ cm^{-3}. The computational domain was restricted to $0 \leq \theta \leq \pi/10$ and $R = 1.5$ pc. The condition $v_\theta = 0$ was imposed at the boundary $\theta = \pi/10$. The GRB has the energy distributions $E = 1.6 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 2$ MeV or $E = 1.7 \times 10^{53} e^{-(\theta/\theta_0)^2}$ erg for $E_{max} = 0.6$ MeV, with

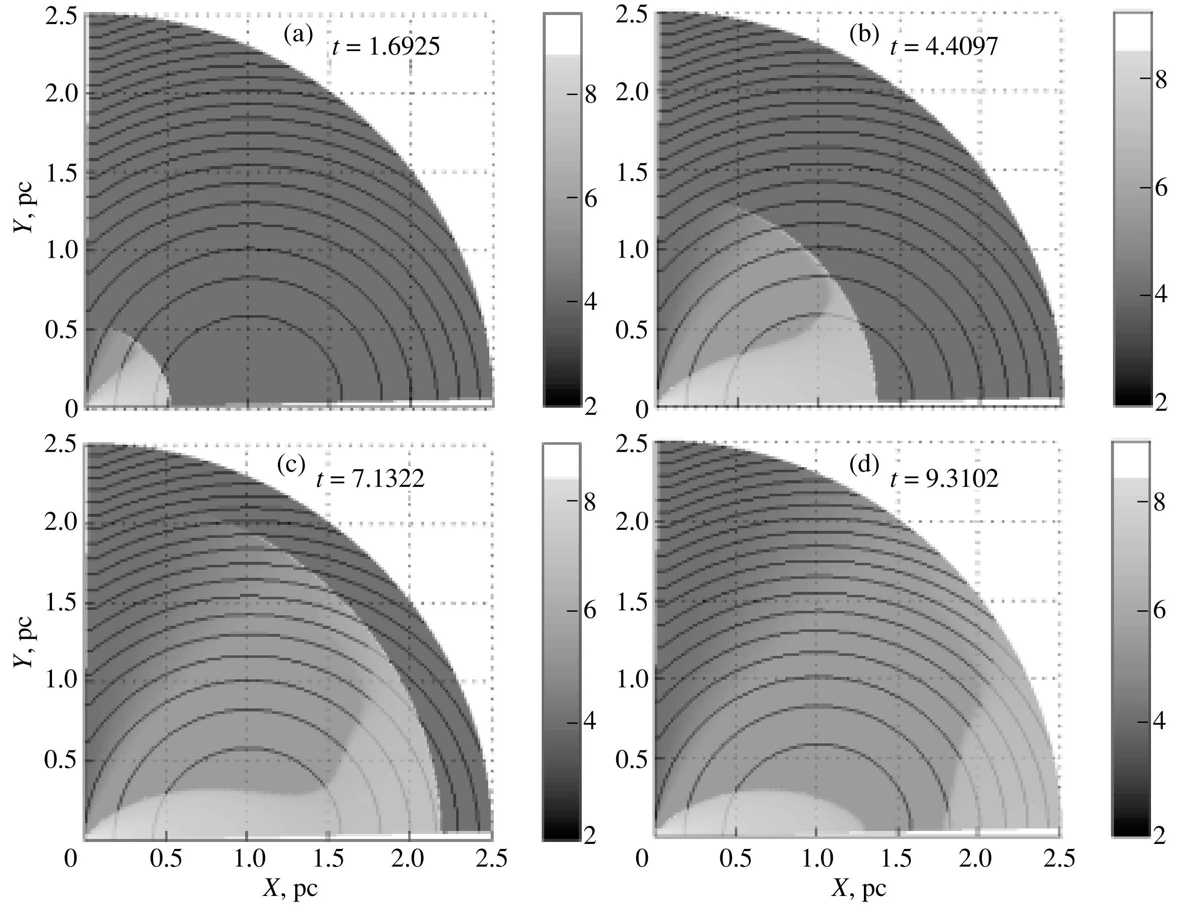


Fig. 3. Same as Fig. 2 for model 8b.

$\theta_0 = 0.4$ rad. This model corresponds to a conical density distribution analogous to model 5.

4. COMPUTATIONAL RESULTS. THERMAL PATTERN

Let us consider the detailed evolution of the thermal pattern of the molecular cloud after the GRB for the two extreme cases corresponding to models 1a and 8a.

A. An Isotropic GRB in a Uniform Molecular Cloud

Consider the time dependence of the temperature distribution. This model is topologically fully analogous to the case considered in [13]. At the initial time, the pulse from the GRB heats the central part of the molecular cloud to temperatures of the order of 10^9 K; heating to higher temperatures is hindered by the joint action of the Klein–Nishina correction and the decrease in the efficiency of Compton heating at temperatures $kT \sim E_{max}/4$. On the spatial scales we have studied (from 10^{16} to 10^{18} cm), the momentum transferred to the matter by radiation is small, and the matter acquires most of its velocity via the pressure gradient. Since matter is heated to high temperatures in the central region ($T > 10^7$–10^8 K), it rapidly becomes fully ionized. Thus, the main radiative losses are due to free–free transitions in the plasma. The thermal-loss rate due to free–free radiation per cubic centimeter of optically thin plasma is proportional to $T^{1/2}$; at such temperatures, the gas can be considered to be ideal, so that its heat capacity does not depend on the temperature. We find that, the more strongly the plasma of the interstellar medium is heated, the longer it takes to cool. Therefore, we can disregard gas cooling near the GRB at early times after the GRB, and assume that the temperature profile will repeat the profile of the GRB energy absorbed by the matter; i.e., the temperature of the matter is proportional to R^{-2} (Fig. 1, curve *1*).

When the GRB pulse has propagated to a distance where it heats matter to temperatures of the order of a million degrees or lower, the volume cooling rate of the matter grows sharply due to the increased efficiency of atomic line radiation [see (12)], and the volume cooling time of the gas becomes shorter than

Results of gas–dynamical flow computations

Model	$v_{r,max}$, 10^3 km/s	$v_{\theta,max}$, 10^3 km/s	$v_{abs,max}$, 10^3 km/s
1a	2.2	0	2.2
1b	5.2	0	5.2
2	2.2	0	2.2
3	1.9	0.00074	1.9
4	5.7	0.39	5.7
5	18	4.4	18
6a	1.4	2.8	2.9
6b	2.6	4.0	4.3
7	1.4	2.8	2.9
8a	4.7	2.2	4.7
9	14	3.4	14

the time for the passage of the GRB pulse to a given radius: $t_{\rm cool} < R/c$ (Fig. 1, curves *2–5*). We find the following pattern. At the initial time, the temperature profile formed by the GRB pulse is described by the dependence $T \sim R^{-2}$. After reaching some radius (which depends on the GRB energy, the hardness of the spectrum, and the density in the molecular cloud), a turnover appears in the temperature profile (the temperature drops to $\sim 10^4$ K, which is in the range of temperatures at which the matter will become virtually completely neutral and the cooling rate will sharply decrease), and one heating wave (due to the GRB) and two cooling waves (one in the direction of increasing and one in the direction of decreasing distance from the GRB) travel radially in the cloud. As a result, the gamma-ray pulse becomes so weak at large distances from the GRB source that it is not able to heat the matter to temperatures of 10^4 K or higher, the interstellar matter remains neutral, and the outwardly propagating cooling wave catches up to the heating front. The energetic gamma-rays will ionize the matter, but our approach is no longer valid in this case—we are not considering such advanced stages in the evolution in this study.

B. An Anisotropic GRB at some Distance from the Center of a Nonuniform Molecular Cloud

The most probable situation is that an anisotropic GRB occurs somewhere other than the exact center of the molecular cloud. We accordingly consider our model 8a to be the one most often realized in nature.

Immediately after the burst, the temperature distribution in the cloud near the main direction of the GRB is determined entirely by the angular energy distribution of the GRB and the distance from the source. Very soon after the GRB, the temperatures at the edges of the gamma-ray pulse (far from the burst axis) do not reach high values (more than a few million degrees), and the matter cools rapidly to temperatures of the order of 10^4 K (Fig. 2a). As the gamma-ray front travels in increasingly dense regions in the cloud, the temperature to which it heats the interstellar matter decreases, while the cooling rate rises, decreasing the thickness of the region of hot gas (Fig. 2b). At this point, the region of heated gas separates into two parts. One follows the gamma-ray pulse, decreases its geometrical thickness, and has the form of a thin meniscus. The second region is oblong and gradually cools, decreasing in size (although the matter in this region expands, the expansion rate is lower than the velocity of the cooling wave in the computations; Fig. 2c). After the gamma-ray pulse passes through the center of the molecular cloud, it proceeds into a region of decreasing density. The falling density leads to a decrease in the cooling rate, which results in an increase in the thickness of the region of heated gas behind the gamma-ray pulse (Fig. 2d).

Let us consider model 8b, for which the GRB energy is an order of magnitude higher. This increase in the GRB energy raises the temperature to which the molecular cloud is heated at a given radius (Fig. 3a). Near the GRB axis, matter at a distance of 2 pc is heated to temperatures exceeding 2×10^6 K. This region encompasses the central regions of the molecular cloud (Figs. 3b, 3c). At such temperatures, the efficiency of cooling is modest, even in the central region of the molecular cloud where the density is highest. The separation of the zone of hot gas into two regions occurs much later (Fig. 3d).

5. GAS–DYNAMICAL EFFECTS DUE TO A COSMOLOGICAL GRB

An important goal of our work was to analyze the motions of gas in the molecular cloud resulting from its heating by the GRB, and the possible formation of strong shocks.

It was shown in [13] that, in the case of a uniform molecular cloud and an isotropic GRB, the central shock wave that develops is nonrelativistic. These computations indicated that the maximum velocity acquired by matter in a uniform molecular cloud with an isotropic GRB with energy 1.6×10^{53} erg for $E_{max} \gtrsim 0.6$ MeV was 2.2×10^3 km/s (table, model 1a); if the GRB energy is increased to 1.6×10^{54} erg, the speed of the flow in the molecular cloud increases to 5.2×10^3 km/s (table, model 1b).

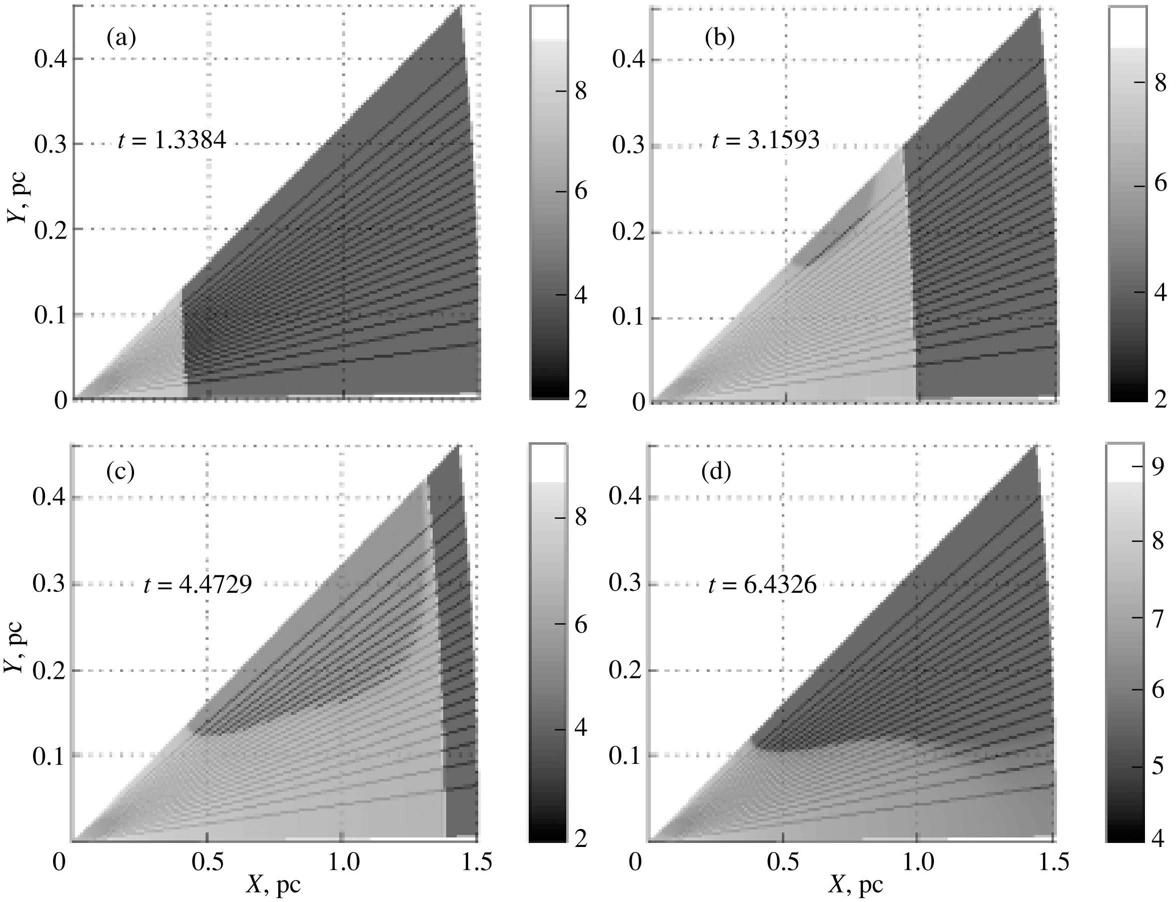


Fig. 4. Flow velocity (arrows) and temperature (grey scale) of the matter at various times for model 5. The distance from the GRB source is given in parsecs, and the indicated times are in years since the GRB. The right-hand bars give the log T(K) temperature scale.

Further growth in the GRB energy does not lead to the formation of a relativistic shock, since the relatively soft observed spectrum of the GRB and the influence of the Klein–Nishina correction limit the efficiency of heating. The mean energy of the photons participating in the heating is $\langle E \rangle \sim (300-600)$ keV over a wide range of spectral hardnesses. Compton heating cannot heat matter to temperatures higher than $T_c = \langle h\nu \rangle / 4k \approx 2 \times 10^9$ K, and the sound speed for this temperature and a normal chemical composition is $c_s = 5.3 \times 10^3$ km/s. As is known from studies of the decay of an arbitrary explosion, the speed of a shock cannot exceed $2c_s/(\gamma - 1)$ [26], which corresponds to $3c_s$ or 1.5×10^4 km/s for $\gamma = 5/3$.

We computed the evolution of the shock formed at the center of the molecular cloud in model 1a. The temperature jump, $T_{\text{after}}/T_{\text{before}}$, remained constant and equal to three while the shock traveled through the heated part of the cloud. The shock front reaches the cooling wave one hundred years after the GRB. Because the shock is not very powerful (the shocked material is not very strongly heated) and the density is not very high, the shock rapidly cools during its interaction with the cooling wave. A weak excess appears in the optical light curve for several tens of days at the level of the numerical noise associated with the discreteness of the computational grid; i.e., the presence of the shock has virtually no effect on the light curve of the optical afterglow.

In model 2 (a moderately inhomogeneous cloud) over the computed time (20 years), the shock propagates similarly to the case of a uniform cloud, and there is not sufficient time for the shock to be strengthened by the density decrease. Therefore, from the viewpoint of the gas flows, this case is virtually identical to model 1a.

The situation with model 3 is very similar to that for model 2. The expected increase in the speed of the shock was not observed, since it would not be physical for the corresponding “step” in the cloud-density profile to be so close to the GRB (hundredths or thousandths of a parsec), and such a situation is very unlikely. The decrease in the size of the computational

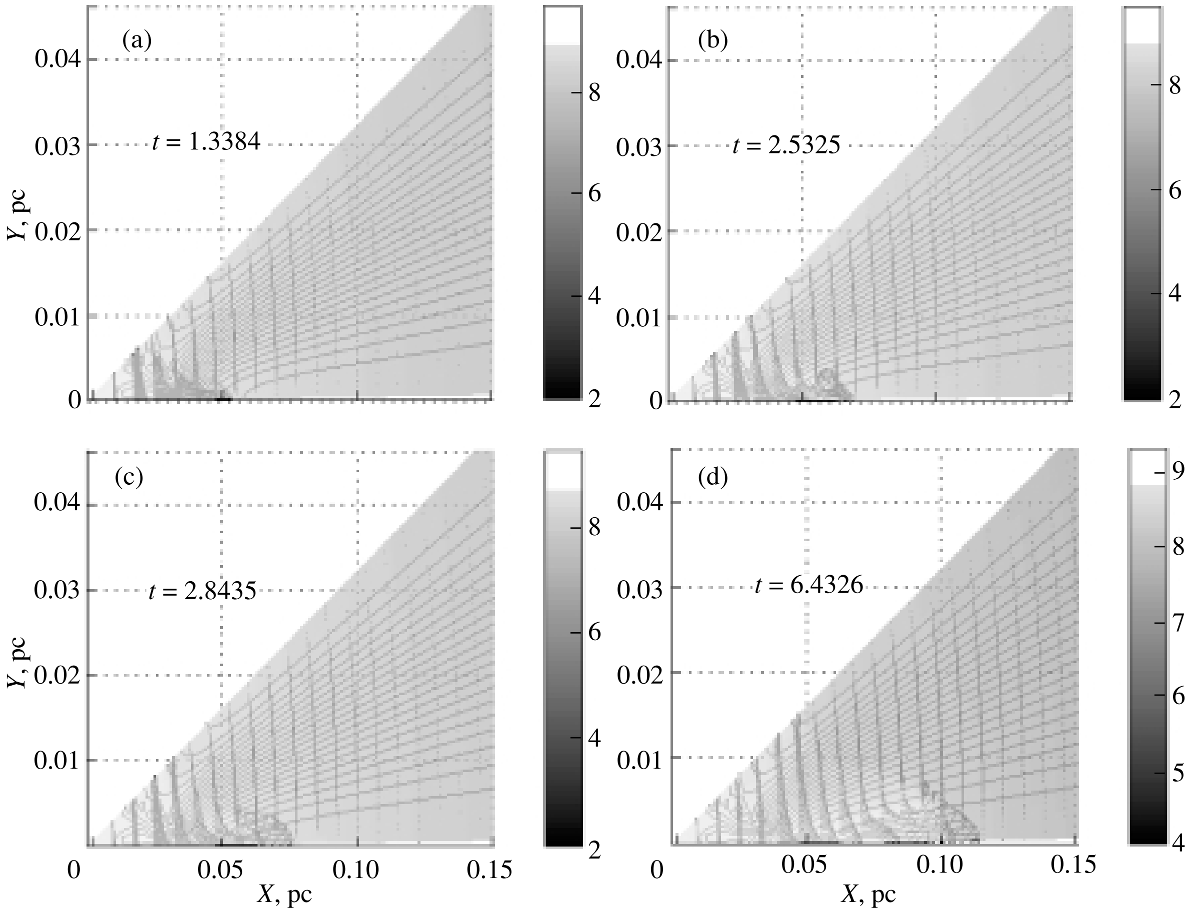


Fig. 5. Same as Fig. 4 on an expanded scale for the central region.

region means that only a small part of the molecular-cloud volume was considered, and this region would not significantly affect the light curve.

In model 4, the shock velocity is a factor of 2.6 higher than the velocity for model 1a. The reason for this is that the GRB source is located in a region with a large density gradient right from the onset of computations, so that the shock is already accelerating as it propagates from the center of the GRB. Note that the region of the molecular cloud with the highest density is not involved in the motion of the shock.

In model 5, the GRB occurs in a cone with low density near its axis and higher density at its edges. Since the molecular cloud is transparent to gamma-rays, the energy absorbed per gram of matter depends only on the distance to the GRB source and the temperature to which the matter is heated, and is independent of the density of the matter. In the case of an isotropic GRB, the temperature of the matter does not depend on the angle θ, giving rise to a pressure gradient that leads to motion of matter toward the axis of the cone. Accordingly, flows arriving from various directions meet at the cone axis and collide. As a result, we find that the phase velocity of the mixing at the point where the flows collide can be estimated as $v_\theta/\sin\theta_0$, where v_θ is the gas flow velocity and θ_0 is the opening angle of the cone (which we took to be $\pi/10$ in the computations). A cumulative effect develops, and matter that ends up near the flow-collision point is accelerated to high velocities by these collisions.

Figure 4 shows the time evolution of model 5, and Fig. 5 an enlargement of the central part of this figure. The arrows show the flow rate of the matter, and the temperature is shown in grey scale. A hot, rapidly moving clump (bullet) with a velocity of about 1.7×10^4 km/s forms near the GRB source, at the point where the gas flows collide. The situation described above can come about in various ways: the GRB precursor can have an anisotropic stellar wind that blows out a cone (hole) in the molecular cloud, or the matter of the inhomogeneous wind itself can form a funnel-like structure. Note that, if the GRB occurs between two nearby molecular clouds, the resulting ejection will have the form of an expanding ring.

Comparing models 1a and 5 in the table, we find

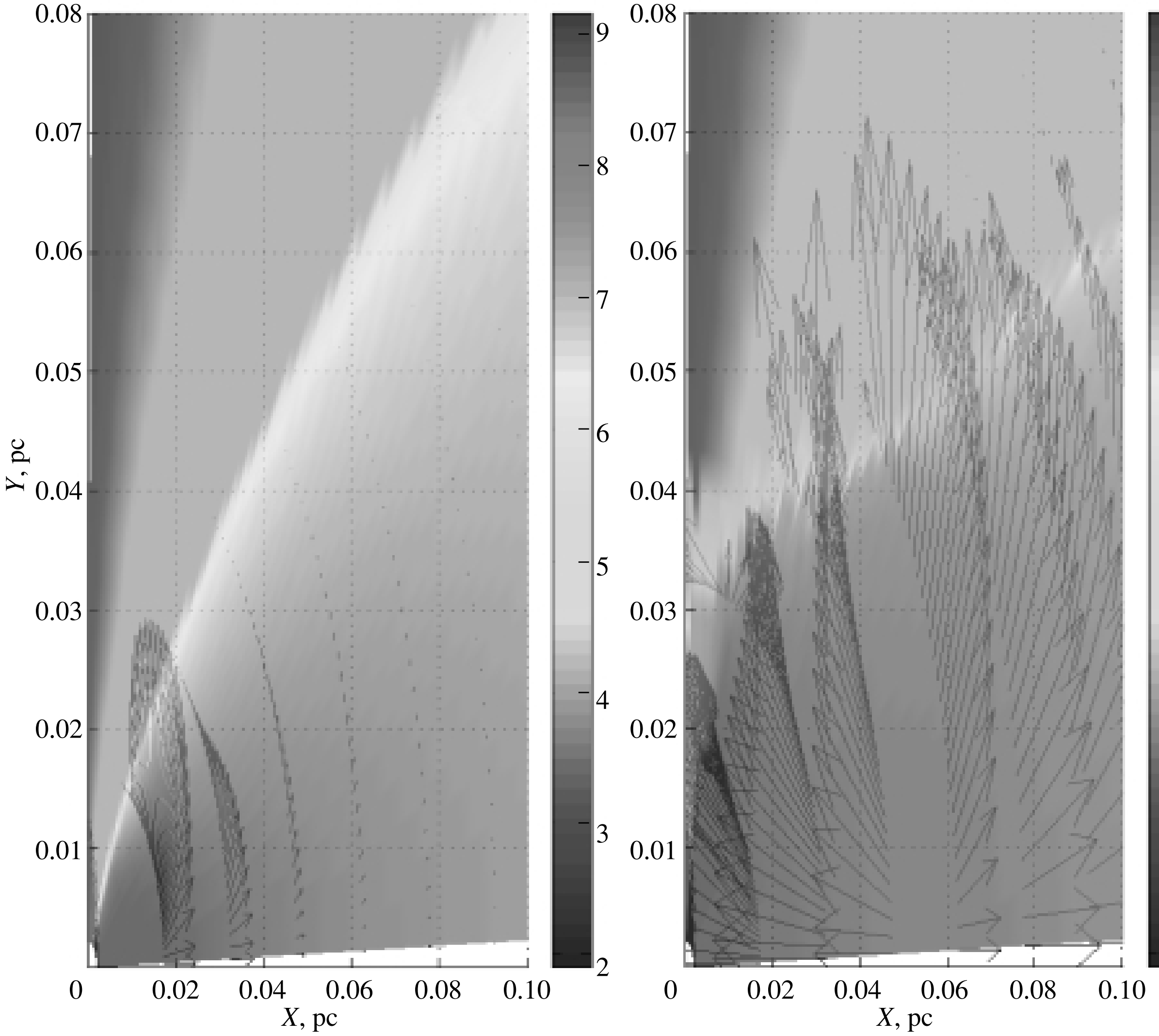


Fig. 6. Same as Fig. 5 for model 6a at two times.

that the maximum velocity of the gas flow is a factor of 8.2 higher in model 5. This indicates that the velocity of the ejected matter can grow by nearly an order of magnitude purely due to geometrical effects.

Models 6a and 6b are fully analogous to models 1a and 1b, but with an anisotropic GRB occurring in a uniform molecular cloud. The anisotropic energy distribution of the burst leads to a large temperature gradient near the GRB source, with this gradient being larger the more narrowly collimated the burst. In the case of uniform density, the temperature gradient coincides with the pressure gradient, leading to the formation of a ringlike shock. The shock moves predominantly perpendicular to, rather than along, the symmetry axis (see the table). In contrast to models 1a and 1b, due to the form of the burst energy distribution, the shock near the GRB source rapidly approaches the cooling wave. Since the material behind the front is hot ($T = 10^8$–10^9 K), the cooling wave does not appreciably influence the shock over the first 20 years (Fig. 6).

The gas flow in model 7 fully coincides with that in model 6a (table). The GRB source is located at the center of the cloud. The density varies only very weakly within 0.1 pc of the cloud center, where the main gas motions occur—the relative density variations are no larger than 0.02. This leads to full agreement with the behavior of the gas in models 7 and 6a.

In model 8a, the structure of the gas motions displays various tendencies in different regions of the molecular cloud. In the immediate vicinity of the GRB source (less than 0.05 pc from it), the gas motion immediately after the passage of the gamma-ray pulse is directed along the temperature gradient (perpendicular to the direction of the GRB pulse). After the neighboring regions are heated by the shock (as in model 6a), the gas flow changes to become parallel to the density gradient (toward the GRB source). Far (0.2 pc) from the GRB source, the velocity of the gas is predominantly perpendicular to the direction toward the GRB, along the temperature gradient.

Model 9 resembles model 5 in terms of the structure and velocities of the gas motions. This is due to the fact that the effective opening angle of the GRB cone (0.4 rad) is larger than the opening angle of the density cone ($\pi/10$ rad). The gas-pressure distribution after the passage of the gamma-ray pulse is close to that for model 5. The structure of the motions for more narrowly collimated GRBs may be close to that for model 6a.

Our computations of the gas flows in a molecular cloud heated by a GRB show that the velocity of the gas flows can vary over nearly an order of magnitude, depending on the density distribution in the molecular cloud. However, it was not possible to obtain a relativistic shock in any of the cases we considered, due to the nature of the heating mechanism and the large distance from the GRB source. The shock itself has virtually no effect on the optical light curve, since the spatial region heated by the shock is small (usually smaller than 0.1 pc), and the temperature at the shock front varies by a factor of three, which is much smaller than the overall temperature variations in the region in which the shock develops (10^6–10^9 K).

6. CONCLUSION

The interaction of a cosmological GRB with dense material in the host galaxy leads to a number of interesting effects. The gamma-ray burst acts on the interstellar medium via the pressure of its radiation and heating. The radiation pressure from the GRB can be important [27], but the computational results presented here and in [13, 14] show that it does not exert an appreciable influence on the interstellar medium at distances of 0.1–1 pc. The main source of gas-dynamical flows in the interstellar medium is the gas-pressure gradient brought about by the temperature or density gradients. The velocity of the front of a nonrelativistic shock cannot exceed the sound speed behind the front. Via the Compton effect, the GRB radiation can heat the material to temperatures of half an MeV, while the mean rest energy of a proton is on the order of a GeV, so that it is not possible to obtain a relativistic flow (without invoking the action of a relativistic shock formed in the GRB itself). As a result, the gamma-ray pulse immediately separates from the shock that is formed, and their evolution can be considered to be independent: the shock travels along the temperature profile formed by the gamma-ray pulse.

We obtained four types of shocks in the computations. In the spherically symmetrical models 1a, 1b, and 2 (as well as model 3), a spherically symmetrical shock forms due to the temperature gradient and propagates from the center with a speed of $\sim 2 \times 10^3 \sqrt{E/1.6 \times 10^{53}}$ km/s. The origin of the shock that forms in model 4 and travels outward from the center of the molecular cloud is the strong density gradient. If the GRB is well collimated (models 6a, 6b, 7, 8a, and 8b), a ringlike shock forms due to the strong temperature gradient, and propagates perpendicular to the GRB collimation axis with a speed of $\sim 2.8 \times 10^3$ km/s. If the GRB precursor had an anisotropic wind, a conical cavity can form in the molecular cloud (models 5 and 9). The propagation of the gamma-ray pulse in this cavity leads to the formation of a rapidly moving hot clump of material (the shock velocity can reach 1.8×10^4 km/s). A similar situation is possible when a GRB occurs between two neighboring molecular clouds; in this case, a rapidly expanding ring arises in a plane parallel to the cloud boundaries.

ACKNOWLEDGMENTS

The authors thank N.G. Bochkarev for useful advice and discussions. This work was partially supported by the Russian Foundation for Basic Research (project no. 02-02-06598, 03-02-06786, 02-02-16900) and the INTAS foundation (grants 99-120 and INTAS 00-491).

REFERENCES

1. R. W. Klebesadel, I. B. Strong, and R. A. Olson, Astrophys. J. Lett. **182**, L85 (1973).
2. M. S. Brigs, Astrophys. Space Sci. **231**, 3 (1995).
3. G. J. Fishman and C. A. Meegan, Annu. Rev. Astron. Astrophys. **33**, 415 (1995).
4. C. A. Meegan, G. J. Fishman, R. B. Wilson, *et al.*, Nature **355**, 143 (1992).
5. C. Kouveliotou, Astrophys. J., Suppl. Ser. **92**, 637 (1994).
6. G. Boella *et al.*, Astron. Astrophys. **122**, 299 (1997).
7. F. Frontera, E. Costa, D. dal Fiume, *et al.*, Astron. Astrophys. **122**, 357 (1997).
8. R. Jager *et al.*, Astron. Astrophys., Suppl. Ser. **125**, 557 (1997).
9. J. van Paradijs *et al.*, Nature **386**, 686 (1997).
10. M. R. Metzger, S. R. Kulkarni, *et al.*, IAU Circ. No. 6588 (1997).
11. S. R. Kulkarni *et al.*, Nature **393**, 35 (1998).
12. V. V. Sokolov, T. A. Fatkhullin, A. J. Castro-Tirado, *et al.*, Astron. Astrophys. **372**, 438 (2001).
13. G. S. Bisnovatyĭ-Kogan and A. N. Timokhin, Astron. Zh. **74**, 483 (1997) [Astron. Rep. **41**, 423 (1997)].
14. A. N. Timokhin and G. S. Bisnovatyĭ-Kogan, Astrophys. Space Sci. **231**, 323 (1995).

15. R. Perna, J. Raymond, and A. Loeb, Astrophys. J. **533** (2), 658 (2000).
16. A. M. Beloborodov and A. F. Illarionov, Astrophys. J. **450**, 64 (1995).
17. G. S. Bisnovatyĭ-Kogan and S. I. Blinnikov, Mon. Not. R. Astron. Soc. **191**, 711 (1980).
18. L. L. Cowie, J. P. Ostriker, and A. A. Stark, Astrophys. J. **226**, 1041 (1978).
19. V. B. Berestetskii, E. M. Lifshits, and A. P. Pitaevskii, *Relativistic Quantum Theory* (Pergamon Press, Oxford, 1971; Nauka, Moscow, 1968, 1971).
20. A. B. Kirienko, Pis'ma Astron. Zh. **19**, 27 (1993) [Astron. Lett. **19**, 11 (1993)].
21. J. Raymond, D. Cox, and B. Smith, Astrophys. J. **204**, 290 (1976).
22. E. Anders and N. Grevesse, Geochim. Cosmochim. Acta **53**, 197 (1989).
23. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1963; Academic Press, New York, 1966, 1967), Vols. 1 and 2.
24. A. A. Boyarchuk, D. V. Bisikalo, O. A. Kuznetsov, and V. M. Chechetkin, *Mass Transfer in Close Binary Stars* (Taylor & Francis, London, 2002).
25. P. Woodward and P. Colella, J. Comput. Phys. **54**, 115 (1984).
26. B. L. Rozhdenstvenskii and N. N. Yanenko, *Systems of Quasi-Linear Equations and Their Application to Gas Dynamics* (Nauka, Moscow, 1968) [in Russian].
27. A. M. Beloborodov, Astrophys. J. **565**, 808 (2002).

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 36–43. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 41–49.
Original Russian Text Copyright © 2005 by Glushkov, Esipov, Kondrat'eva.

Visual Extinction Towards the Omega Nebula (M17) and the Ionization of Nitrogen in the Nebula

Yu. I. Glushkov[1], V. F. Esipov[1], and L. N. Kondrat'eva[2]

[1]*Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia*
[2]*Fesenkov Astrophysical Institute, Academy of Sciences of Kazakhstan,*
Kamenskoe Plato, Almaty, 480068 Kazakhstan

Received March 28, 2004; in final form, July 15, 2004

Abstract—We estimate the absolute intensity of the Hα emission line in M17 based on spectroscopic observations with the 70 cm telescope of the Fesenkov Astrophysical Institute and the 1.25 m telescope of the Sternberg Astronomical Institute's Crimean Laboratory. The visual extinction, A_v, is derived for about 250 regions in the nebula via a comparison of the optical (Hα) and radio (21 cm continuum) emissions. The A_v values for the eastern, optically bright, part of the nebula are $2–6^m$. For the western part, which is covered by a dark cloud, the extinction is $A_v = 7–14.5^m$. We derived $I(\text{H}\alpha)/I([\text{NII}]\,\lambda6584)$ ratios and estimated the degree of ionization of nitrogen in the nebula; we find that most nitrogen is in the [NIII] state.

1. INTRODUCTION

The Omega Nebula (M17) is one of the Galaxy's brightest star-forming regions in any spectral range, and is optically the second brightest after the Orion Nebula (NGC 1976), which is at least a factor of three closer to the Earth. In contrast to NGC 1976, whose optical and radio regions almost coincide, these regions are significantly different for the Omega Nebula. This makes it important to study the extinction toward M17, and such a study is the main subject of the current paper. This work is a continuation of our earlier paper [1], where we briefly described the nebula and the problems concerning its nature, and presented a fairly detailed discussion of data on its electron density and ionization.

In 1961, Pariiskii [2] and Hobbs [3] were the first to note the difference between the nebula's emission in the optical and radio. They found that the center of the radio emission was to the west of the nebula's optically bright region, and essentially coincided with a "dark bay." In the same year, M17 was observed at the Astrophysical Institute of the Academy of Sciences of Kazakhstan, and the nebula's isophotes in Hα were derived in units of erg cm^{-2} s^{-1} sr^{-1} [4]. These results (with an angular resolution of 35″) were compared to the radio isophotes derived by Hobbs [3] (with an angular resolution of 6′), providing the first photovisual-extinction data for 57 positions in the nebula. The extinction varied from $1.5–3^m$ in the nebula's optically bright region to $5–10^m$ in its western part. A similar study was carried out in 1968 in Japan by Ishida and Kawajiri [5], also using the radio observations of Hobbs [3]; their extinction values varied from 1^m to 7^m. In the same year, Dickel [6] published a study in which she determined the extinctions for four positions in the nebula (with an angular resolution of 2.5′).

Several papers have been dedicated to studies of extinction in M17 since that time, but always with angular resolutions no better than 2′.

In the current paper, we present A_v values for ∼250 positions in the nebula, based on observations with an angular resolution of ∼10–12″ in both the optical and radio. Our optical observations were in the red, so we were able to derive $I(\text{H}\alpha)/I([\text{NII}]\,\lambda6584)$ ratios and thereby estimate the degree of ionization of nitrogen.

2. OBSERVATIONS AND RESULTS

We observed M17 in the region of the Hα emission line in 1995–1997 using the 70 cm telescope of the Fesenkov Astrophysical Institute (with a diffraction spectrograph, UM-92 image tube, and A-600 photographic film) and the 1.25 m telescope of the Sternberg Astronomical Institute's Crimean Laboratory (with the A spectrograph and ST-6 and ST-6I CCD arrays). The spectral range observed was λλ6000–7000 Å. The main goal of these observations was to determine the absolute intensities of the Hα emission line ($I(\text{H}\alpha)$) in erg cm^{-2} s^{-1} sr^{-1}. We used a fixed section of the nebula Simeiz 57 [7] as a standard. The mean uncertainty in $I(\mathrm{H}\alpha)$ was $\sim$15%. A comparison of the resultant $I(\mathrm{H}\alpha)$ values with the data of [5, 6] yields deviations that are no greater than 20%. Note, however, that Ishida and Kawajiri [5] and Dickel [6] used photographs of M17 taken at $\lambda\lambda$6300–6700 Å, and estimated the contributions of the [NIII] and HeI lines and the very strong continuum of M17 only roughly. In our reduction, we scanned the Hα line along the spectrograph slit and corrected for the continuum on both sides of the line, thereby obtaining a "pure" $I(\mathrm{H}\alpha)$ value.

We obtained about 100 spectrograms for 25 sections of the nebula with the 70 cm telescope in this way, with a mean dispersion of $\sim$70 Å/mm. Figure 1 shows the positions of the projection of the spectrograph slit relative to the nebula's stellar field. The scale in arcseconds for the horizontal and vertical axes is indicated. This schematic does not reproduce the stars' magnitudes correctly, but their plotted positions are accurate to $\sim 2''$. The values of $I(\mathrm{H}\alpha)$ in erg cm^{-2} s^{-1} sr^{-1} are plotted along the vertical axis in Fig. 2, while the horizontal scale is in arcseconds. When the slit was positioned along the direction of right ascension or was tilted, the readings increase from left to right (from east to west). When the slit was positioned along the declination, the readings increase from bottom to top (from south to north). The vertical bars in the figures show the positions of stars through which each section passed. Similarly, Fig. 3 shows the variations of $I(\mathrm{H}\alpha)/I(\mathrm{[NII]}\ \lambda6584)$. The probable errors of these values are $\sim$8–10%. For most sections, the variations of $I(\mathrm{H}\alpha)$ and $I(\mathrm{H}\alpha)/I\mathrm{[NII]}$ relative to a star projected onto the spectrograph slit are essentially identical. However, deviations corresponding to distances of several tens of arcseconds are observed in some sections, since the observations were acquired on different nights, but with the same position of the spectrograph slit.

With the 1.25 m telescope, we obtained about 30 spectrograms for nine sections in the western part of M17, whose Hα emission is weak. All these data were used to determine the visual extinction, A_v. The measured $I(\mathrm{H}\alpha)$ values can be used to derive the emission measure, EM(Hα). The electron temperature $T_e = 10^4$ K corresponds to $\mathrm{EM}(\mathrm{H}\alpha) = 1.21 \times 10^7 I(\mathrm{H}\alpha)$ pc cm^{-6} [8, 9]. Comparing EM(Hα) and the radio emission measure, EM(R), we can determine the visual extinction, A_v: $\log[\mathrm{EM(R)}/\mathrm{EM}(\mathrm{H}\alpha)] = 0.4A(\mathrm{H}\alpha)$, $A_v = 1.33A(\mathrm{H}\alpha)$. We derived EM(R) using the observational results derived from the VLA data of Felli *et al.* [10] (in the 21 cm continuum with an angular resolution of $10.2'' \times 7.5''$).

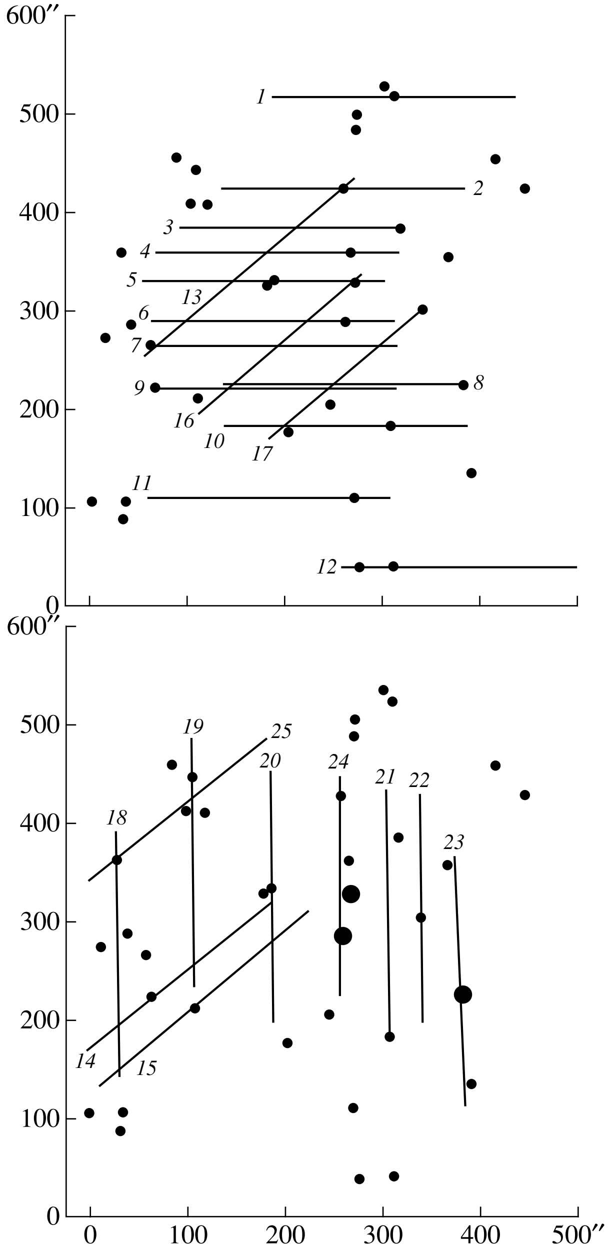


Fig. 1. Schematic of the stellar field in M17 and the projections of the spectrograph slit during the observations with the 70 cm telescope (see text).

We superposed the radio isophotes of [10, Fig. 1] on a red photograph of M17 taken, at our request, by A. Amirkhanyan with the 2.6 m telescope of the Byurakan Astrophysical Observatory (Fig. 4). The stellar coordinates for this purpose were taken from [11]. The same figure shows the sections (spectrograph slit projections) and A_v values (in magnitudes). Also plotted (larger-font numbers) are the A_v values of Glushkov *et al.* [12], based on the EM(Hα) values of [4] and the EM(2 cm) values of Schraml and Mezger [13] (angular resolution $2'$). The cross

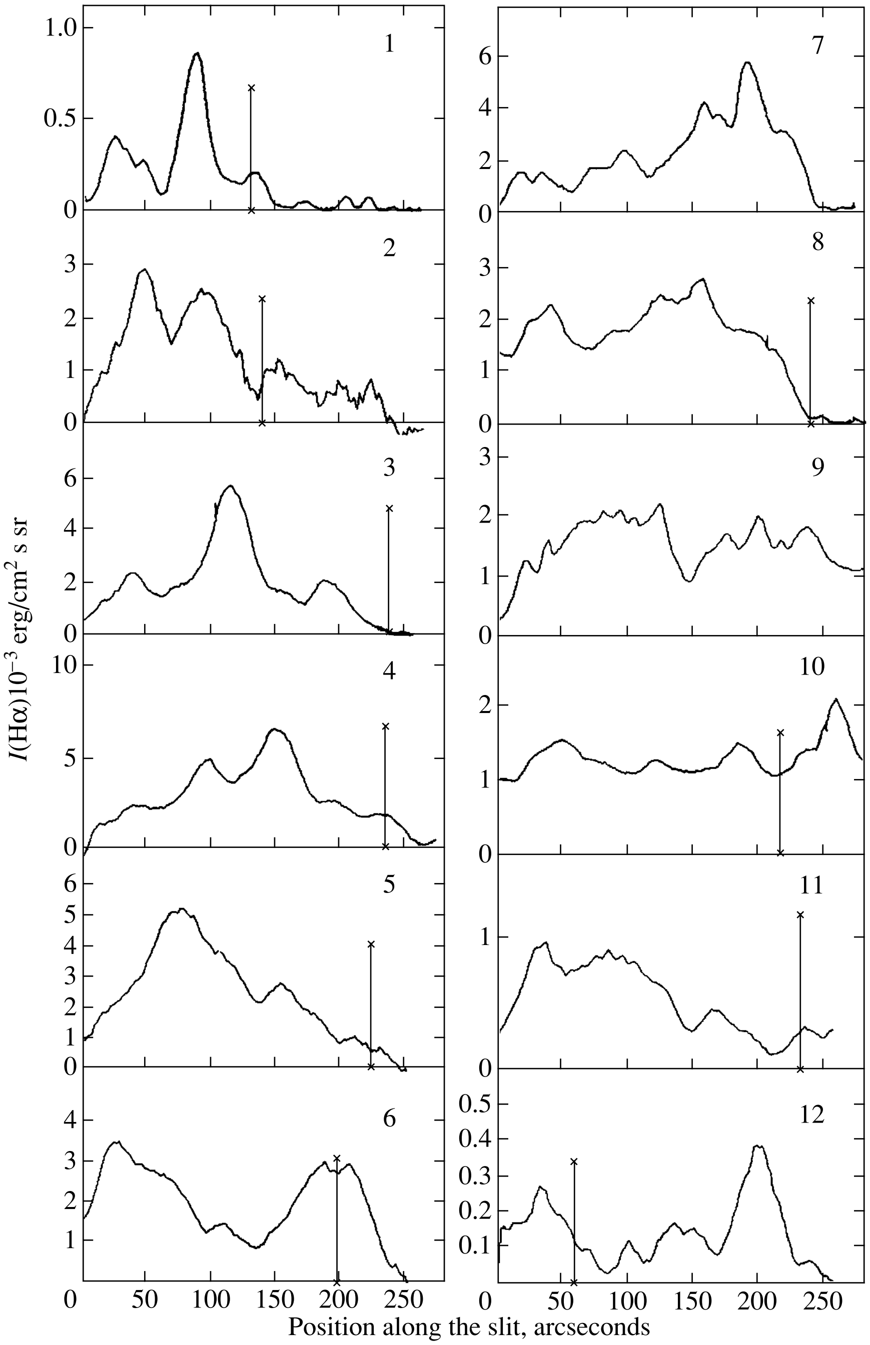


Fig. 2. Absolute intensity of the Hα emission line (see text for details).

shows the position of the ultra-compact HII region M17 UC1 [10].

3. DISCUSSION OF THE RESULTS

3.1. Analysis of the Extinction

When analyzing extinction values, it is desirable to know the distance to the object so that it is possible to estimate the interstellar extinction between the Earth and the nebula. The rough distance estimates obtained prior to 1960 were from 1 to 3 kpc. An analysis of star counts to 19^m made in 1962 in three fields (in the nebula's western part, outside the nebula, and in the Kapteyn area SA134) yielded a lower limit for the distance to M17 of 1200 pc [4]. Nearly the same distance (1300 pc) was derived in [14] from *UBV* observations of stars in M17. After the *UBVRI* observations of Chini *et al.* [15], the distance of 2.2 kpc was accepted for many years. In

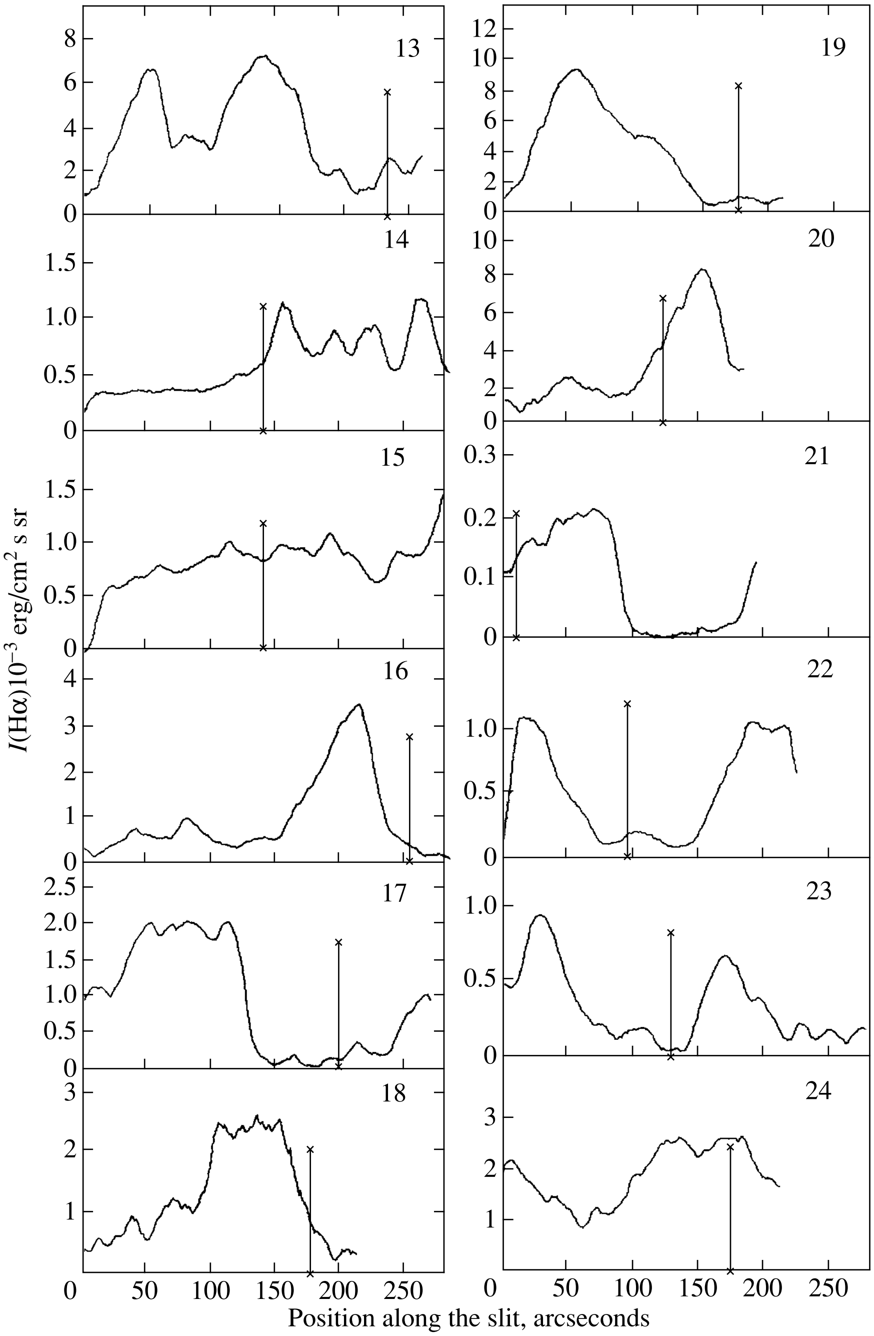


Fig. 2. (Contd.)

1997, Hunson *et al.* [11] used their spectroscopic and photometric data to conclude that the distance to M17 was 1300 (+400, −200) pc. In 2001, the IR and optical observations of massive stars in M17 carried out by Nielbock *et al.* [16] yielded a distance to the Omega Nebula of 1.6 ± 0.3 kpc. Thus, we can be quite certain that the lower limit to the distance is 1200 pc and the upper limit 1800 pc. With the earlier estimates in mind, the currently adopted distance to M17 is ∼1400 pc.

It follows from the work of Sharov [17] that the interstellar extinction within a $1^\circ \times 1^\circ$ field around M17 is 0.6^m/kpc. A similar analysis of stellar extinctions within a $\sim 20'$ radius around M17 yielded the interstellar extinction $\sim 2^m$/kpc [4]; i.e., A_v varies from 0.8^m to 2^m.

It follows from the data in Fig. 4 that the lowest extinction toward M17 is 2–3^m, and occurs along a narrow strip at the southern end of the nebula's optically brightest part (EM(Hα) = (8.5–11.0) ×

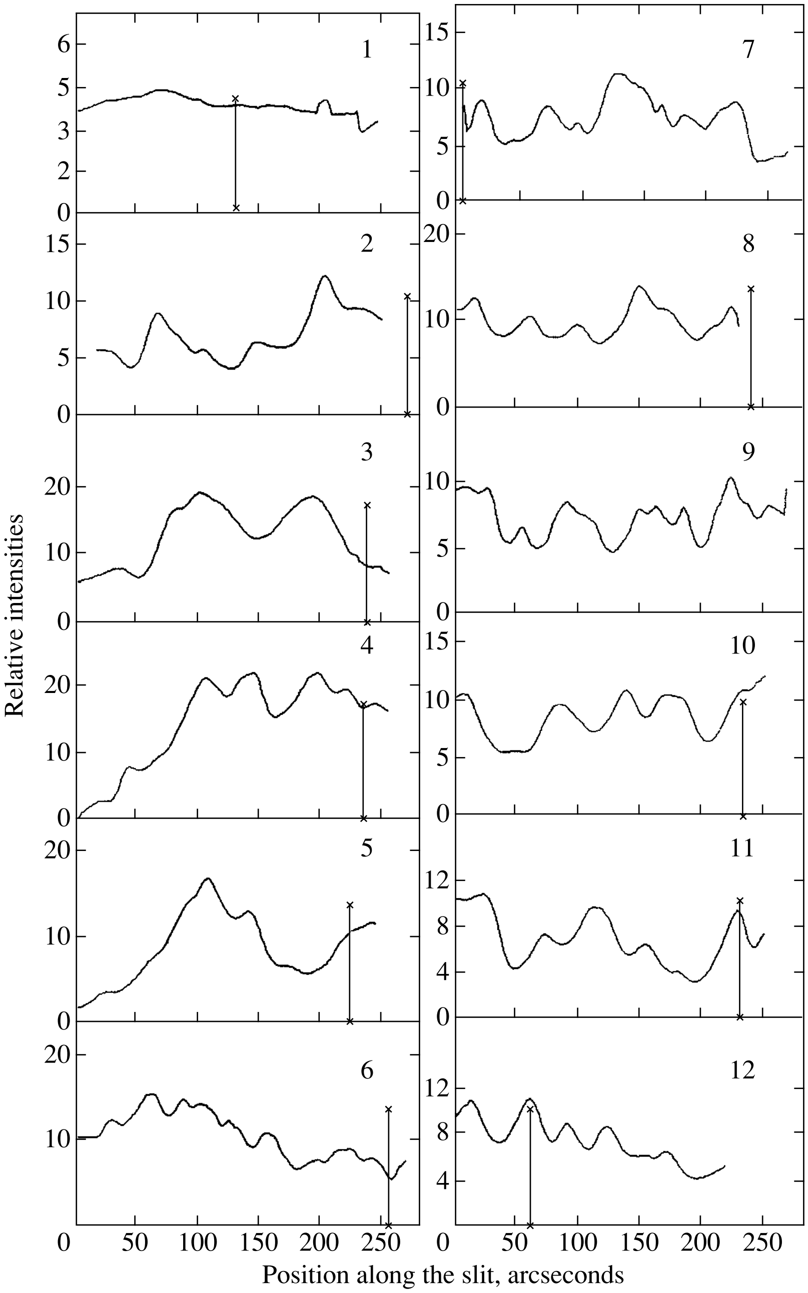


Fig. 3. $I(\mathrm{H}\alpha)/I([\mathrm{NII}]\,\lambda6584)$ ratios (see text for details).

10^4 pc cm^{-6})), coincident with the southern end of the radio emission's northeastern arm (the N bar, λ21 cm). This is a region of high observed ionization. This feature is probably the inner edge of an ionization front associated with the wind from hot O3–O6 stars in the dark bay [1]; the extinction in the nebula itself does not exceed 1^m. The extinction increases in all directions from this edge: reaching 4–4.5^m to the south and 5–6^m to the north, where it coincides with one of the eastern maxima of the molecular cloud [18]. The extinction increases sharply toward the western edge of the nebula's optically bright part: from 5–6^m, it rises to values of 7–8^m, then 9–12^m to the north of the N bar of the radio emission. The extinction throughout the western part of M17 is extremely clumpy, and is lower by 2–3^m at positions where the nebula's patches and filaments are visible through the dark cloud. The highest extinction (13–14.5^m) is observed to the south of the ultra-compact HII region (M17 UC 1). Lower extinction is observed in the compact HII region associated with the star LH7 [1], where its value is $\sim6^m$, in contrast to the

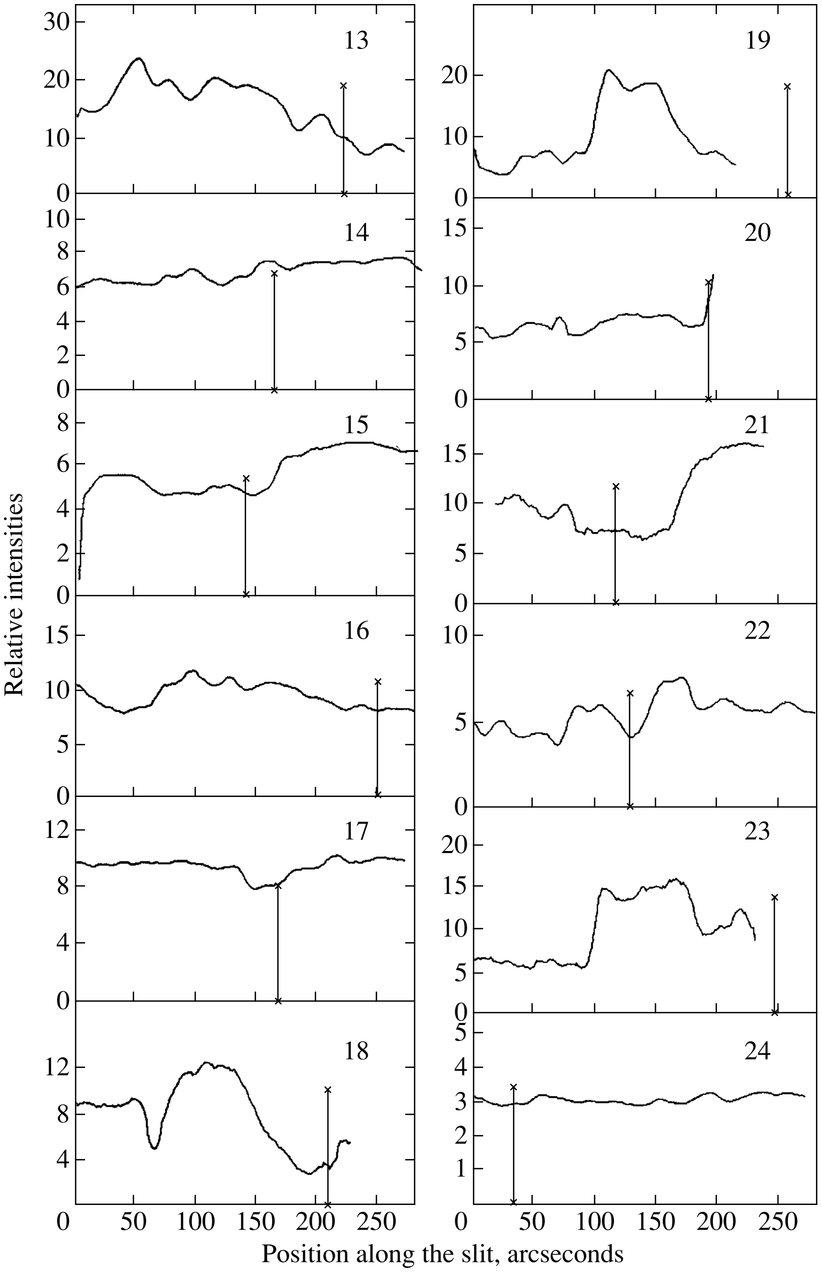


Fig. 3. (Contd.)

value $A_v = 10\text{–}12^m$ for this general region in the nebula.

We compared our results to the data of Chini *et al.* [15], who used *UBVRI* observations to estimate A_v for dozens of stars. Most of our estimates for stars approximately coincide with the A_v estimates of [15], but there are several stars whose extinctions are $2\text{–}3^m$ higher than in the surrounding parts of the nebula, in particular, stars 30, 87, 88, 89, 90, 95, 96, 97, 98, 100, 101. It is quite possible that these stars possess dense circumstellar envelopes.

3.2. Ionization of Nitrogen in M17

Burbidge *et al.* [19] were the first to use the $I(\mathrm{H}\alpha)/I([\mathrm{NII}])$ ratio to determine the electron temperatures (T_e) in spiral and irregular galaxies. This method is not very effective for estimating T_e in diffuse nebulas, since knowledge of elemental abundances and ionization levels is needed.

Computations of $I(\mathrm{H}\alpha)$ and $I([\mathrm{NII}])$ carried out in 1972 led to the formula:

$$f\frac{I(\mathrm{H}\alpha)}{I([\mathrm{NII}]\lambda6584)} = 0.106 b_3 e^{\frac{39\,558}{T_e}} T_e^{-1} \frac{N(\mathrm{H})}{N(\mathrm{N})}.$$

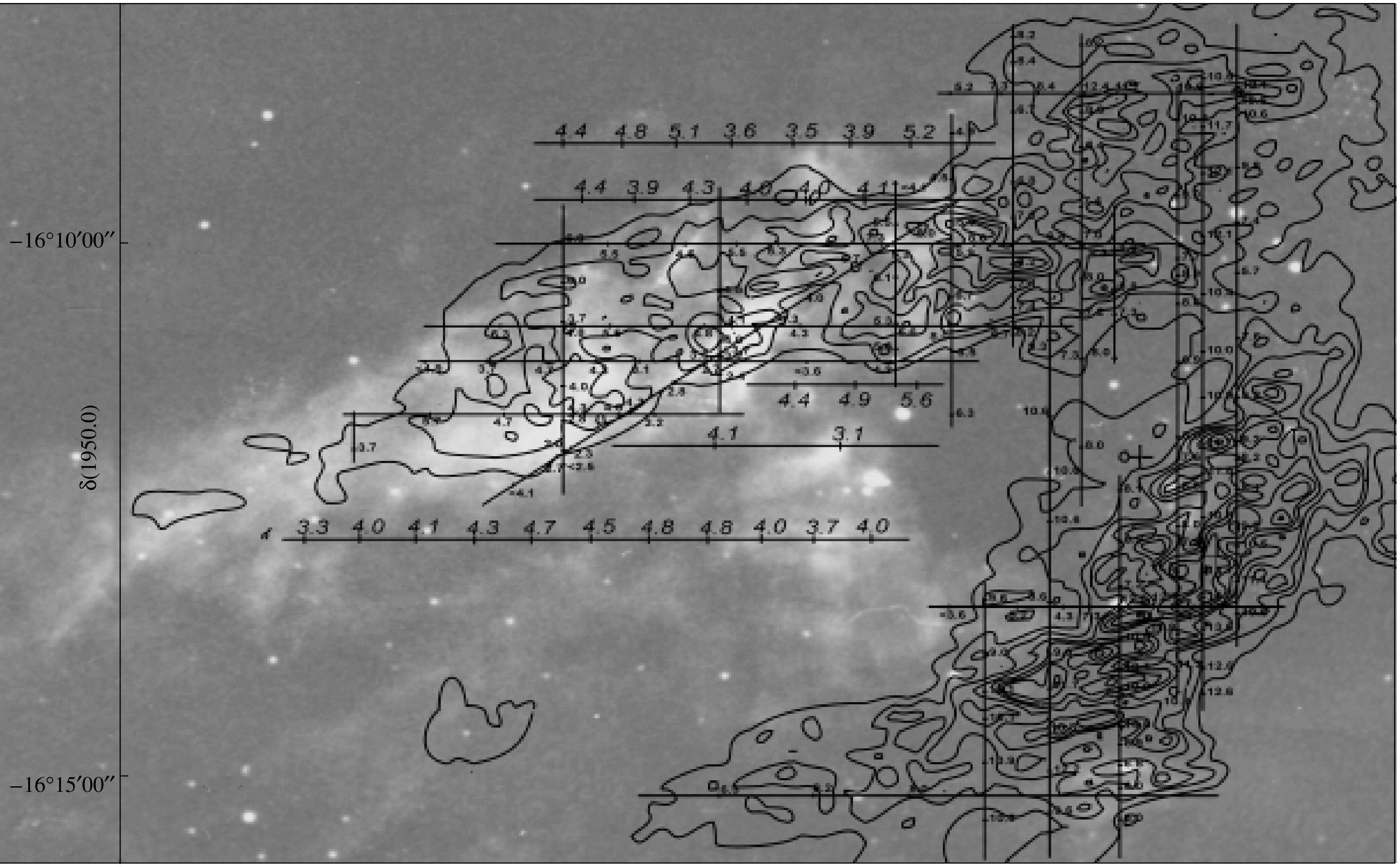


Fig. 4. Photograph of M17 with 21 cm radio contours overlaid and the A_v values indicated (see text for more detail).

Here, $f = N(\mathrm{N}^+)/N(\mathrm{N})$ is the ratio of the number of ions, N^+, to the total number of nitrogen atoms; $N(\mathrm{H})/N(\mathrm{N})$ characterizes the chemical composition; and b_3 is a factor describing deviations from thermodynamic equilibrium. The results of computations using this formula were tabulated for three chemical compositions and a wide range of T_e from 3000 to 20 000 K [8, 9].

A detailed analysis of the chemical composition and electron temperatures of M17 is presented in [20]; on average, $T_e = 9000$ K, and the chemical composition is nearly solar, $N(\mathrm{H})/N(\mathrm{N}) \approx 10^4$. Under such conditions, $fI(\mathrm{H}\alpha)/I([\mathrm{NII}]\lambda6584) = 0.85$.

We can now estimate f using the known ratios. It follows from our analysis of Fig. 3 and from [1] that these ratios vary strongly across the nebula: from 3 at its periphery to 22 in its central regions. Thus, the values of f vary from $\sim$0.3 to 0.04. In the former case, most nitrogen is probably in the N^0 state, and in the latter, in the N^{++} state. The ratio $I(\mathrm{H}\alpha)/I([\mathrm{NII}]\lambda6584) = 16-20$ over large areas of the nebula's eastern, optically brightest, part. In the heavily obscured region in the western part of the nebula, the ratio $I(\mathrm{H}\alpha)/I([\mathrm{NII}]\lambda6584) = 10-16$ only among patches and filaments visible at more transparent places in the dark cloud. The highest values (18–20) are found in the small strip ($\sim 20'' \times 30''$) at the edge of the molecular cloud where the ionization front is located [1]. These ratios for other regions in the western part of the nebula are three to five.

The large observed variations of these ratios imply that both the extinction and ionization in the nebula are clumpy. This is especially clear from spectra taken with the 6 m telescope of the Special Astrophysical Observatory with an angular resolution of 3.6″ [1]; these results remain mostly unpublished. The variations are definitely real; we found a close correlation between the $I(\mathrm{H}\alpha)/I[\mathrm{NII}]$ and $I(\mathrm{HeI})/I[\mathrm{SII}]$ ratios, with the two ratios increasing together. The ratios are sometimes observed to vary within a region 10″ in size.

A forthcoming paper will present an analysis of spectra of the nebula acquired with the long-slit spectrograph of the 6 m telescope (Special Astrophysical Observatory) with angular resolutions of 3.6″ and 4.4″.

4. CONCLUSION

We have presented a complete picture of the visual extinction, A_v, in the Omega Nebula. We estimated A_v in both the nebula's eastern, optically bright part and in its western part, which is obscured by a dense absorbing cloud. Our quantitative estimates of the ionization of nitrogen show that, like the extinction, the ionization is extremely clumpy. This means that the nebula consists of multiple dense knots and filaments that are partially or fully ionized.

ACKNOWLEDGMENTS

The authors are grateful to A.M. Tatarnikova for his assistance during the reduction of the spectra taken with the 1.25 m telescope and to M. Felli for discussions concerning the determination of the emission measure at 21 cm.

REFERENCES

1. Yu. I. Glushkov, Astron. Zh. **75**, 163 (1998) [Astron. Rep. **42**, 137 (1998)].
2. Yu. N. Pariiskii, Astron. Zh. **38**, 483 (1961).
3. R. W. Hobbs, Astron. J. **66**, 517 (1961).
4. D. A. Rozhkovskii, Yu. I. Glushkov, and K. G. Dzhakusheva, Tr. Astrofiz. Inst. Akad. Nauk Kaz. SSR **14**, 19 (1962).
5. K. Ishida and N. Kawajiri, Publ. Astron. Soc. Jpn. **20**, 95 (1968).
6. H. Dickel, Astrophys. J. **152**, 651 (1968).
7. Yu. I. Glushkov and Z. V. Karyagina, Astron. Tsirk., No. 632 (1971).
8. Yu. I. Glushkov, Candidate's Dissertation (Leningrad Gos. Univ., 1973).
9. Yu. I. Glushkov, E. K. Denisyuk, and Z. V. Karyagina, Tr. Astrofiz. Inst. Akad. Nauk Kaz. SSR **34**, 40 (1979).
10. M. Felli, E. Churchwell, and M. Massi, Astron. Astrophys. **136**, 53 (1984).
11. M. M. Hunson, I. D. Howarth, and P. S. Conti, Astrophys. J. **489**, 698 (1997).
12. Yu. I. Glushkov, E. K. Denisyuk, and Z. V. Karyagina, Tr. Astrofiz. Inst. Akad. Nauk Kaz. SSR **31**, 38 (1978).
13. T. Schraml and P. G. Mezger, Astrophys. J. **156**, 269 (1969).
14. K. Ogura and K. Ishida, Publ. Astron. Soc. Jpn. **28**, 35 (1976).
15. R. Chini, H. Elsasser, and T. Neckel, Astron. Astrophys. **91**, 186 (1980).
16. M. Nielbock, R. Chini, M. Jutte, and E. Manthey, Astron. Astrophys. **377**, 273 (2001).
17. A. S. Sharov, Astron. Zh. **40**, 900 (1963) [Sov. Astron. **7**, 689 (1963)].
18. T. Henning, R. Klein, R. Launbardt, *et al.*, Astron. Astrophys. **332**, 1035 (1998).
19. C. R. Burbidge, R. J. Gould, and S. R. Pottasch, Astrophys. J. **138**, 945 (1963).
20. C. Esteban, M. Peimbert, S. Torres-Peimbert, and J. Garcia-Rojas, Rev. Mex. Astron. Astrofis. **35**, 65 (1999).

Translated by N. Samus'

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 44–56. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 50–63.
Original Russian Text Copyright © 2005 by Lekht, Munitsyn, Tolmachev.

Monitoring of the H_2O Maser W31(2) in 1981–2003

E. E. Lekht[1,2], V. A. Munitsyn[3], and A. M. Tolmachev[4]

[1] *Instituto Nacional de Astrofísica, Óptica y Electrónica,*
Luis Enrique Erro No. 1, Apdo Postal 51 y 216, 72840 Tonantzintla, Puebla, México
[2] *Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia*
[3] *Space Research Institute, Russian Academy of Sciences, Profsoyuznaya ul. 84/32, Moscow, 117810 Russia*
[4] *Pushchino Radio Astronomy Observatory, Astro Space Center, Lebedev Physical Institute,*
Russian Academy of Sciences, Pushchino, Moscow oblast, 142290 Russia

Received April 8, 2004; in final form, July 15, 2004

Abstract—A catalog of water-vapor maser spectra at 1.35 cm for the source W31(2), which is associated with an active star-forming region, is presented. The observations were carried out in 1981–2003 on the 22-m antenna of the Pushchino Radio Astronomy Observatory with a spectral resolution of 0.101 km/s. The mean interval between observations was about 1.5 months. The total velocity range in which emission was observed during the monitoring is from −14 to +14 km/s. The spectrum is strongly variable and contains a large number of emission features. Two strong flares with an interval between their emission maxima (integrated flux) of about 12 years (1985–1986 and 1998–1999) were observed, as well as fast variations on a timescale of 0.5–2 years with amplitudes of up to 600 Jy km/s. No long-period component of the variations was found. A drift of the velocity centroid has been detected; it is well approximated by a third-power polynomial corresponding to a period of about 31–33 years. The two strong flares fall on different phases of this curve: the first (1985–1986) is located near the minimum, while the second (1998–1999) is at the maximum. The observed character of the variability of the emission is well explained by the existing model for the region of G10.6–0.4. The drift of the velocity centroid is probably associated with the nonstationary accretion of material onto an HII region formed by a cluster of OB stars.

1. INTRODUCTION

W31 is a composite giant complex of molecular clouds located near the center of our Galaxy. This region was discovered in the radio in the 1950s by Westerhout [1] at a frequency of 1390 MHz. The later 5-GHz observations of Goss and Shaver [2] showed that W31 consists of three isolated regions of radio emission: G10.2–0.3, G10.3–0.1, and G10.6–0.4. Emission in the H109α hydrogen radio recombination line at radial velocities of 13.9, 9.7, and 0.3 km/s, respectively, was detected toward these three regions [3], together with IR emission at 69 μm [4]. There is good coincidence between the coordinates of the sources. The distance to W31, $\approx$6 kpc, was determined by Caswell *et al.* [5] based on interferometric observations in the 21-cm line.

The region G10.6–0.4, which is associated with OH and H_2O maser emission, is of particular interest. As a source of H_2O maser emission, it also bears the name W31(2). The IR source no. 7 detected by Wright *et al.* [4] is observed at the center of this region.

The observations of Ho and Haschick [6, 7] in the continuum at 2 and 6 cm and in the NH_3 line at 1.3 cm, as well as the continuum observations of Ho *et al.* [8] at 6 and 20 cm, revealed fine structure in the emission. The region of ionized gas G10.6–0.4 includes eight components smaller than 0.15 pc in size with masses from 0.3 to 1.08 $M_\odot$. This multiple structure of G10.6–0.4 is associated with the presence of a cluster of forming young OB stars. Analysis of the NH_3, CS, and CO molecular line emissions [9] shows that G10.6–0.4 harbors a warm and dense molecular cloud with typical parameters $T_k = 20–40$ K, $n(H_2) > 10^4$ cm^{-3}, and $M = 10^2–10^4 M_\odot$.

Guilloteau *et al.* [10] found evidence in the H_2CO absorption line for strong fragmentation of the cloud core, which is also rotating around the HII region. The density in the core is 10^6 cm^{-3}, and the mass of each fragment is several solar masses.

Th HII region itself has a core and a halo, as was suggested by the 22-GHz continuum maps of [11]. Ho and Haschick [7] and Keto *et al.* [12], who studied the core of the G10.6–0.4 cloud in the NH_3 absorption line in detail, inferred the probable collapse and

1063-7729/05/4901-0044$26.00 © 2005 Pleiades Publishing Inc.

accretion in the central part of the rotating cloud. The radial velocities of the CS lines are coincident with those for the hydrogen recombination lines [13, 14]. The map of the region obtained by Cesaroni *et al.* [15] in the $C^{34}S$ line confirms the existence of a rotating disk associated with an ultracompact HII region. In addition, they showed that the gas in this region is dense, $\sim 10^6$ cm^{-3}, its mass is $\sim 2000\ M_\odot$, and the size of the region is ~0.3–0.5 pc. Later, based on lines of $C^{34}S$ (the $J = 5$–4 transition) and CS ($J =$ 10–9), the density of the gas in the ultracompact HII regions was found to be still higher, $\sim 10^7$ cm^{-3} [14].

A collapsing and rotating molecular core containing a cluster of forming high-mass OB stars within a region 5″ in size has been adopted as a model for the region [7]. The velocity gradient in the core is ~10 (km/s)/pc. The core is embedded in a slowly rotating envelope with a velocity gradient of ~1 (km/s)/pc. Features with a redshift of ~5 km/s relative to the velocity centroid of the rotating core could be fragments collapsing toward the cluster of OB stars.

Finally, the region G10.6–0.1 is associated with maser sources. OH maser emission was discovered by Turner [16] in 1970, and H_2O maser emission by Turner and Rubin [17] in 1971. The H_2O maser is located near the compact HII region [6], while the OH maser is at the center of the region [11]. The most precise positions of the water maser W31(2) were measured by Genzel and Downes [18] and Fazio *et al.* [19]. According to these data, the maser is located in the immediate vicinity of the ultracompact HII region.

In addition, another H_2O maser source, W31(1), has been detected [18]. This maser does not coincide with any of the three most intense regions of radio emission in W31, but does coincide with an IR source in W31 [4]. The first studies of the time variability of the H_2O maser emission in W31(2) were carried out by White and Macdonald [20]. However, their observations spanned only two years. We began regular observations (monitoring) of W31(2) in the 1.35-cm line in 1981 on the 22 m telescope in Pushchino. Some results for 1981–1992 have been published by Lekht *et al.* [21].

In the current paper, we consider the evolution of the W31(2) H_2O maser emission based on our monitoring in 1981–2003.

2. OBSERVATIONS AND DATA

We obtained observations of the 1.35-cm H_2O maser emission toward the star-forming region G10.6–0.4 and, in particular, W31(2) ($\alpha_{1950} = 18^h07^m30.3^s$, $\delta_{1950} = -19°56'38''$) using the 22-m radio telescope of the Pushchino Radio Astronomy Observatory in 1981–2003, with a mean interval between observations of about 1.5 months.

The receiving equipment has been upgraded several times during the monitoring interval. Before 1993, a liquid-helium-cooled maser amplifier was used. Beginning in September 1993, we have been using a FET amplifier, which lowered the system noise temperature from 200–300 to 150–200 K. After the next upgrade in 2000, the system noise temperature was reduced to 100–150 K. The signals were analyzed first using a 96-channel spectrometer, which was replaced by a 128-channel filter-bank spectrometer with a resolution of 7.5 kHz (0.101 km/s in radial velocity at 22 GHz) in 1997.

The antenna beamwidth at 22 GHz is 2.6′. Symmetrical beam switching was used to reduce the influence of the atmosphere. We mainly used the ON–ON method. The antenna was pointed at the source first using one feed and then the other, with a calibration noise signal being injected for some time during the first cycle. During strong flares of the H_2O maser, we used the ON–OFF method, with the antenna pointed to a comparison point during the first cycle, while a calibration noise generator was activated. For a pointlike source of unpolarized emission, an antenna temperature of 1 K corresponds to a flux density of 25 Jy. The observational technique and instrumentation are described in detail by Sorochenko *et al.* [22].

The reduction of the observations included establishing the radial-velocity scale of the spectrograms, baseline subtraction, and correction for absorption in the Earth's atmosphere. Since the elevation of W31(2) was 10–15° in our observations, correcting the spectra for atmospheric absorption was important. In the summer, the signal was occasionally reduced by as much as a factor of 1.8. The temperature, pressure, and relative humidity were recorded during the observations. The atmospheric attenuation was calculated using the technique described by Zhevakin and Naumov [23] and Tseĭtlin [24].

The absorption-corrected spectra are presented in Fig. 1.

The horizontal axis plots the radial velocity relative to the Local Standard of Rest in km/s, with the same scale for all spectra, while the vertical axis plots the flux density in Jy. In view of the large range of flux variations, the figures are not plotted using the same vertical scale. The vertical arrows show the scales in Jy. The horizontal lines trace the zero levels of each spectrum.

Figure 2 shows the variations of the integrated flux on two different scales together with variations of the

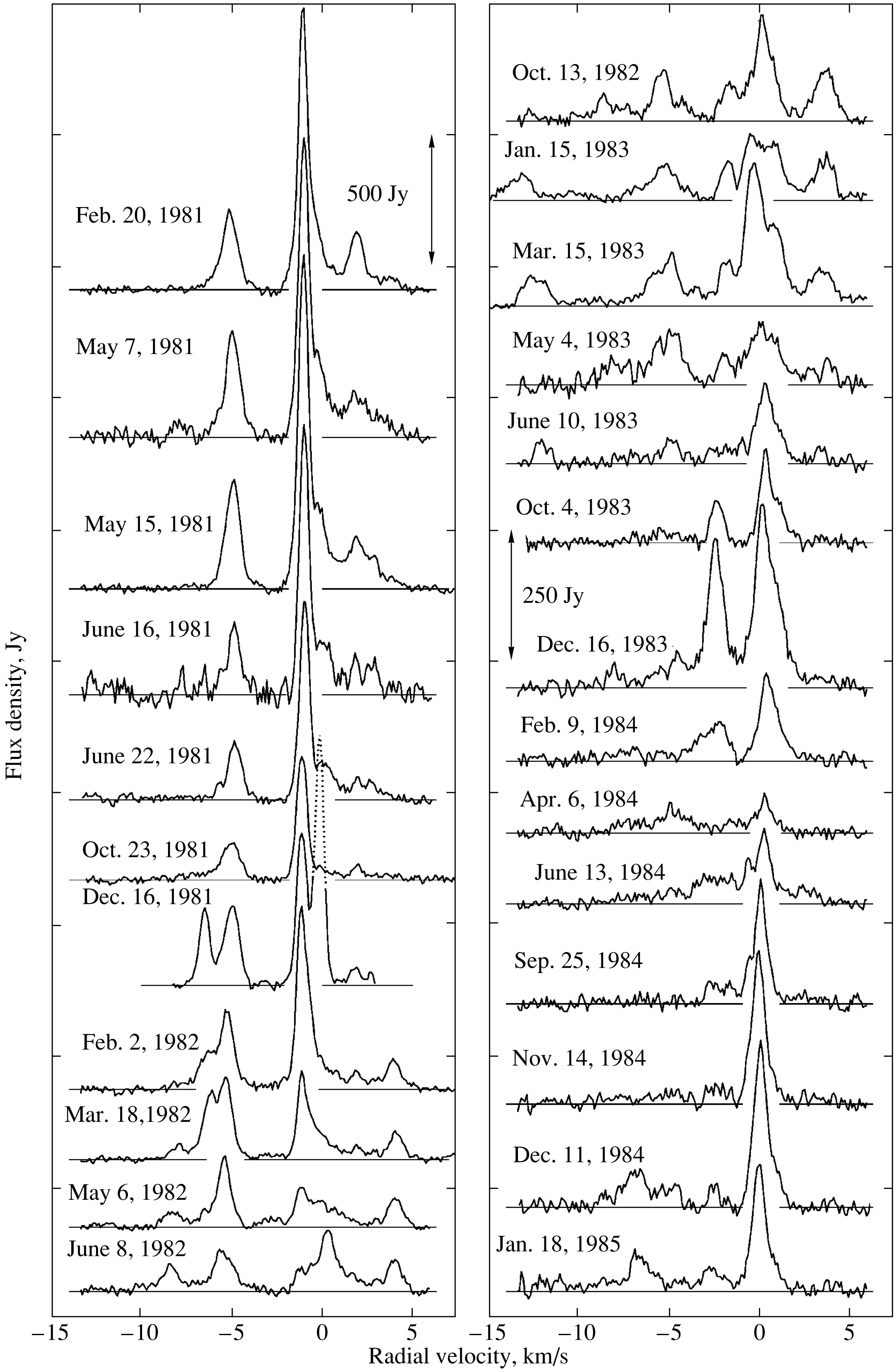


Fig. 1. H_2O maser spectra for W31(2) obtained in 1981–2003. The vertical arrows show the scales in Jy. The radial velocity along the horizontal axis is relative to the Local Standard of Rest.

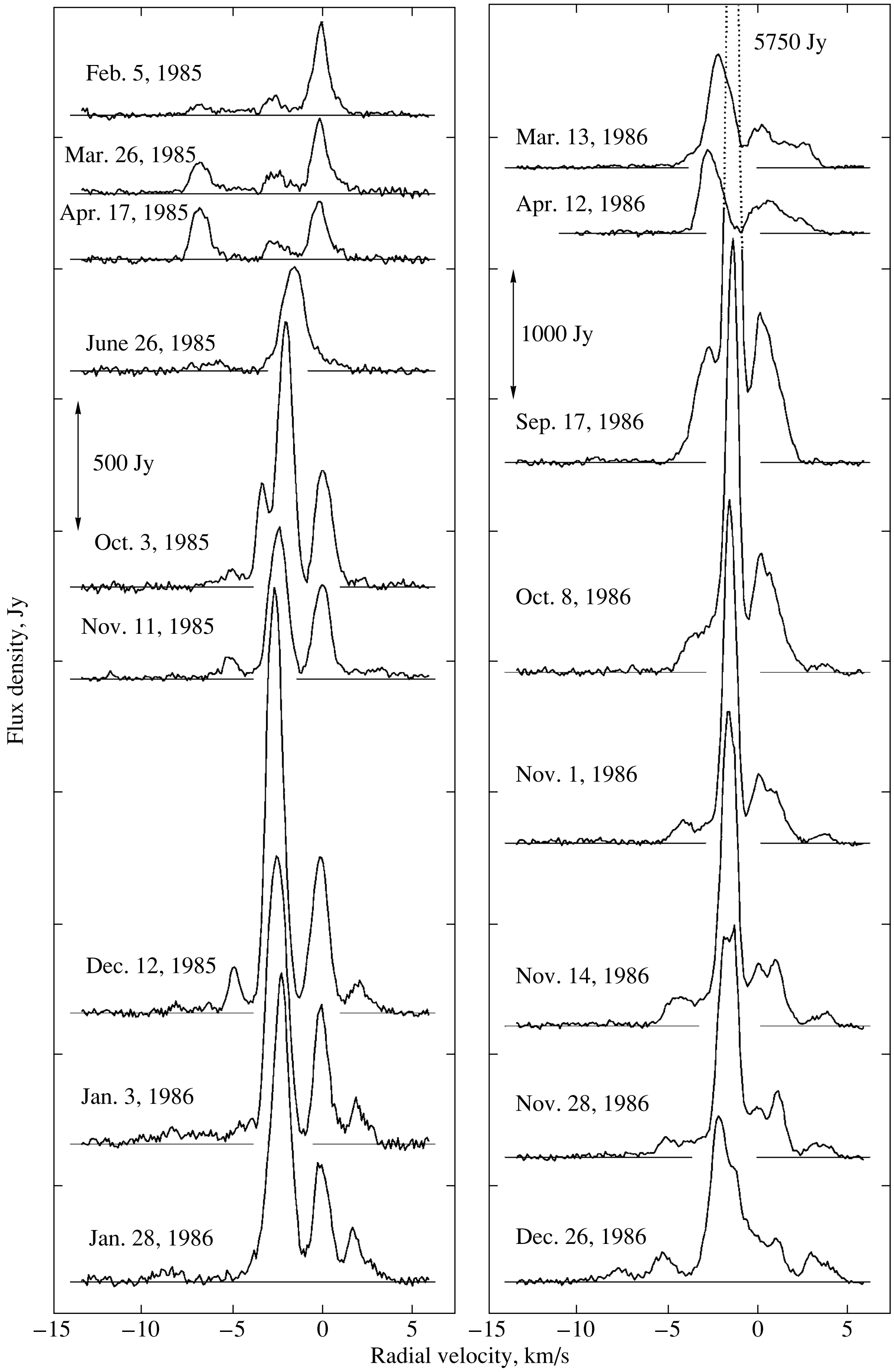


Fig. 1. (Contd.)

velocity centroid (weighted-average radial velocity), which was calculated as

$$V_c = \frac{\sum F_i V_i}{\sum F_i}. \quad (1)$$

We also computed these parameters for spectra obtained on other radio telescopes prior to our monitoring, i.e., before 1981. The cross denotes a point taken from the catalog of Comoretto *et al.* [26]. The dashed and dotted lines in the middle panel show

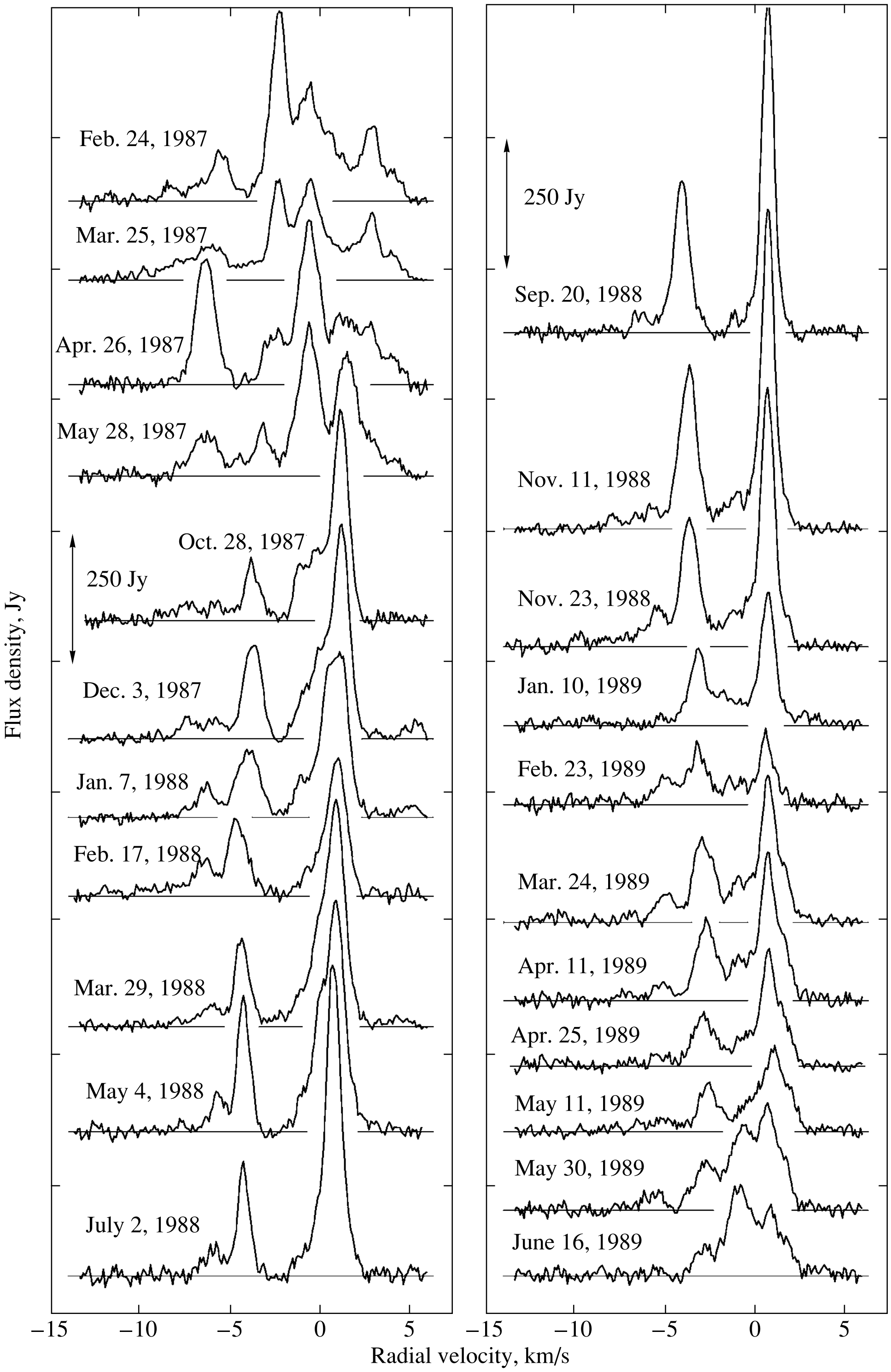


Fig. 1. (Contd.)

the corridor corresponding to the integrated-flux variations, excluding the two powerful flares of 1985–1986 and 1998–1999. The upper (dashed) line is 600 Jy km/s above the lower line.

The dashed curve in the bottom panel is a smooth curve obtained by fitting a third-power polynomial. However, this curve does not describe the centroid variations for 1992–2003 well. Therefore, the dash-and-dot line shows a more complex curve, which reflects faster centroid variations than the smooth

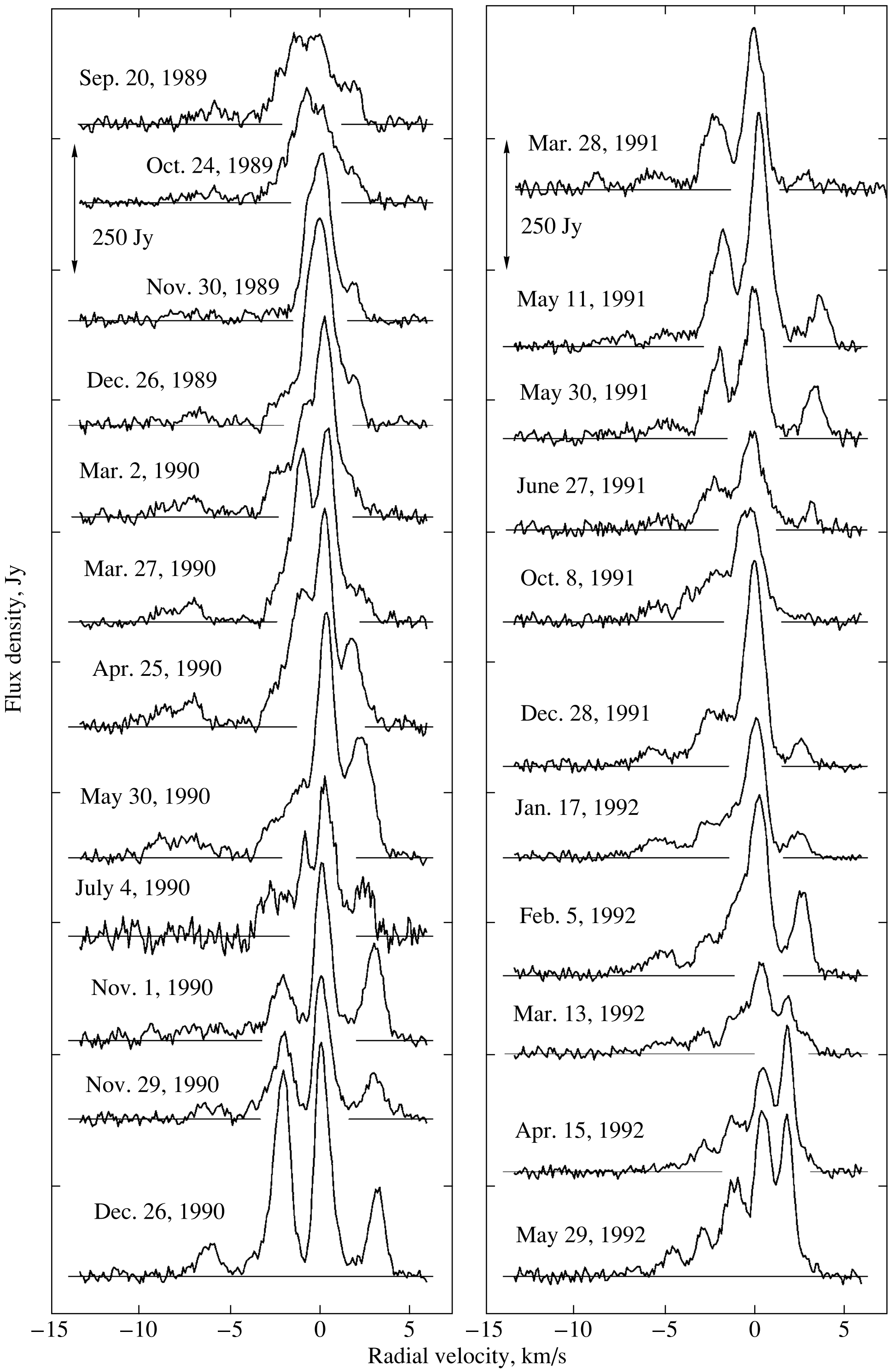


Fig. 1. (Contd.)

curve. The figures at the ends of this curve denote the radial-velocity intervals and velocities of individual features whose emission resulted in variations of the centroid relative to the smooth curve.

In addition, we present another curve for the velocity centroid after 2002 (dotted), which takes into account weak emission at velocities 8.7–11.5 km/s. The upper envelope of the radial velocity for the main curve did not exceed 8.7 km/s. The thin horizontal line shows the mean level of the centroid variations,

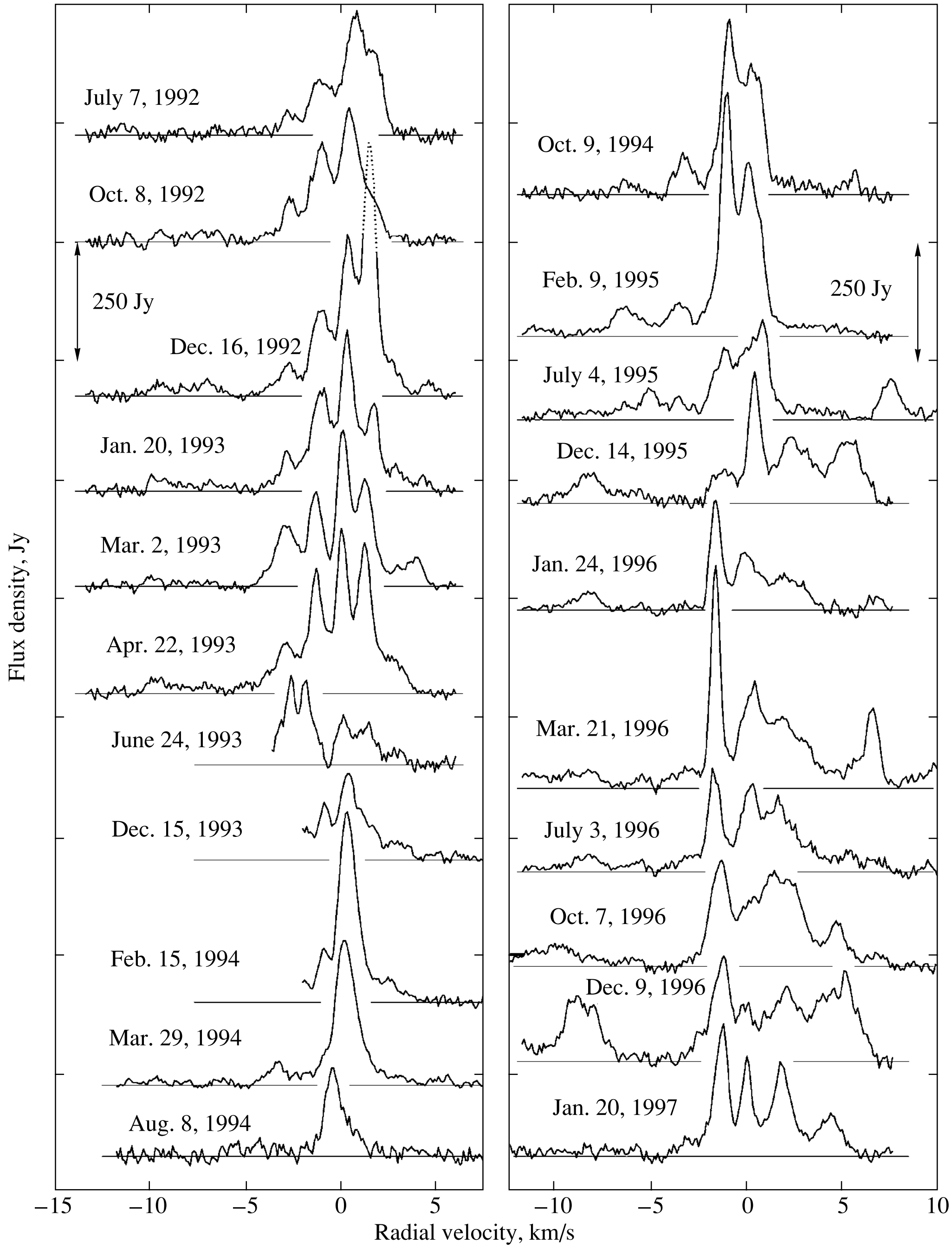


Fig. 1. (Contd.)

equal to −0.8 km/s. The positions of the radio recombination and molecular lines are shown to the left.

3. DISCUSSION

In contrast to most masers in star-forming regions, the W31(2) H_2O maser could be associated with an HII region containing a cluster of forming OB stars. The region G10.6−0.4 has been well studied in the continuum, in many molecular and radio recombination lines, and in the IR. VLA observations of the hydroxyl maser at 1.6 GHz have also been performed [11]. The OH masers are concentrated more in the halo of the ultracompact HII region than toward its core. Unfortunately, similar water-maser observations are lacking, preventing us from localizing the W31(2) H_2O maser spots relative to the ultracompact region G10.6−0.4. For this reason, we attach large importance to the analysis of the variability of the H_2O emission as a whole; i.e., of the integrated flux and velocity centroid (Fig. 2).

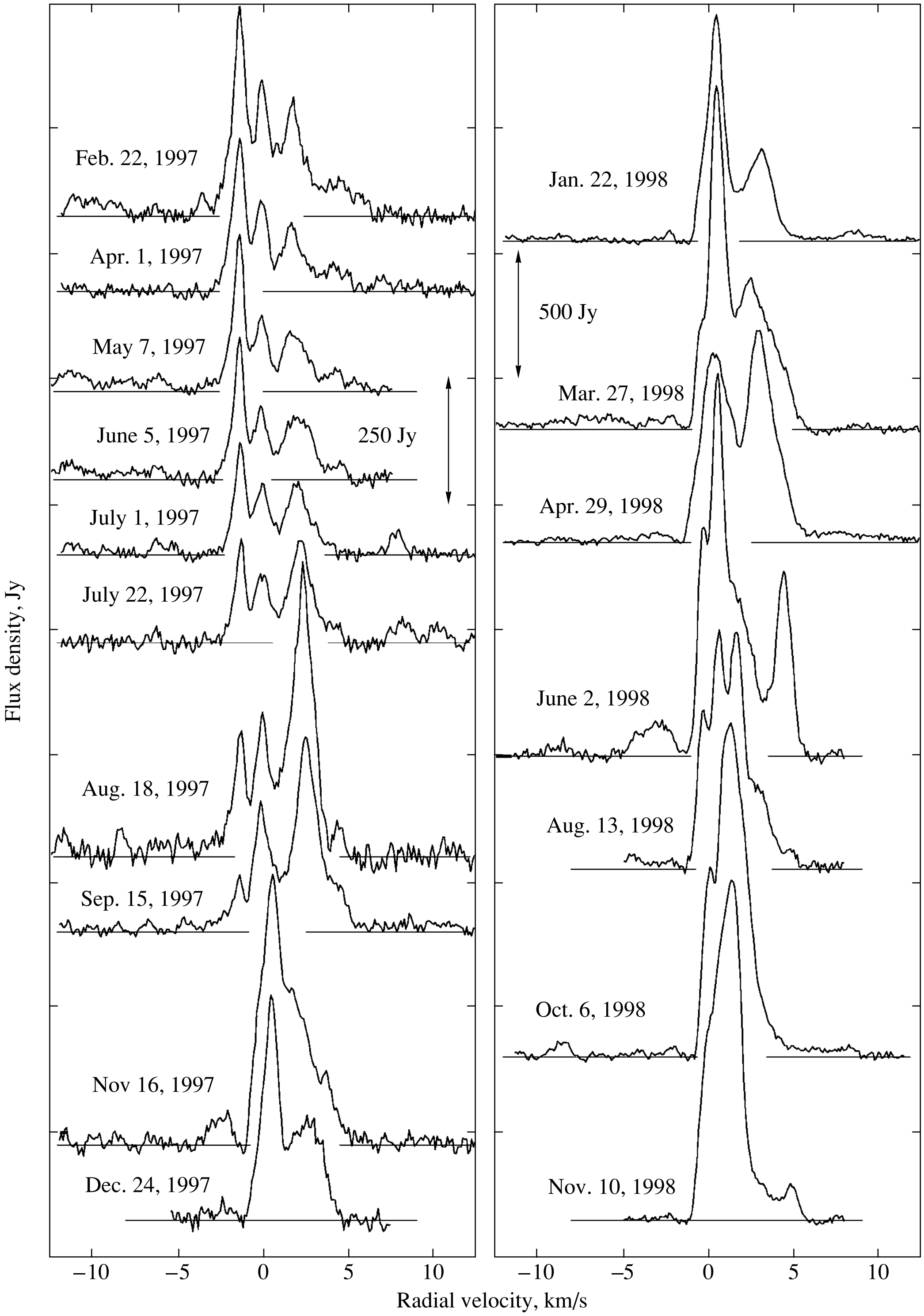


Fig. 1. (Contd.)

3.1. Integrated Flux

The most important events in W31(2) were two powerful flares whose maxima were separated by 12 years: one in 1985–1986 and the other in 1998–1999. The velocity intervals of the flare emission were −5 to 2.5 and −2 to 5 km/s, respectively. The difference in the interval widths was only 0.5 km/s, whereas the shift of the interval center reached 2.75 km/s, toward higher radial velocities. The more powerful flare was in 1985–1986, and was preceded by a very

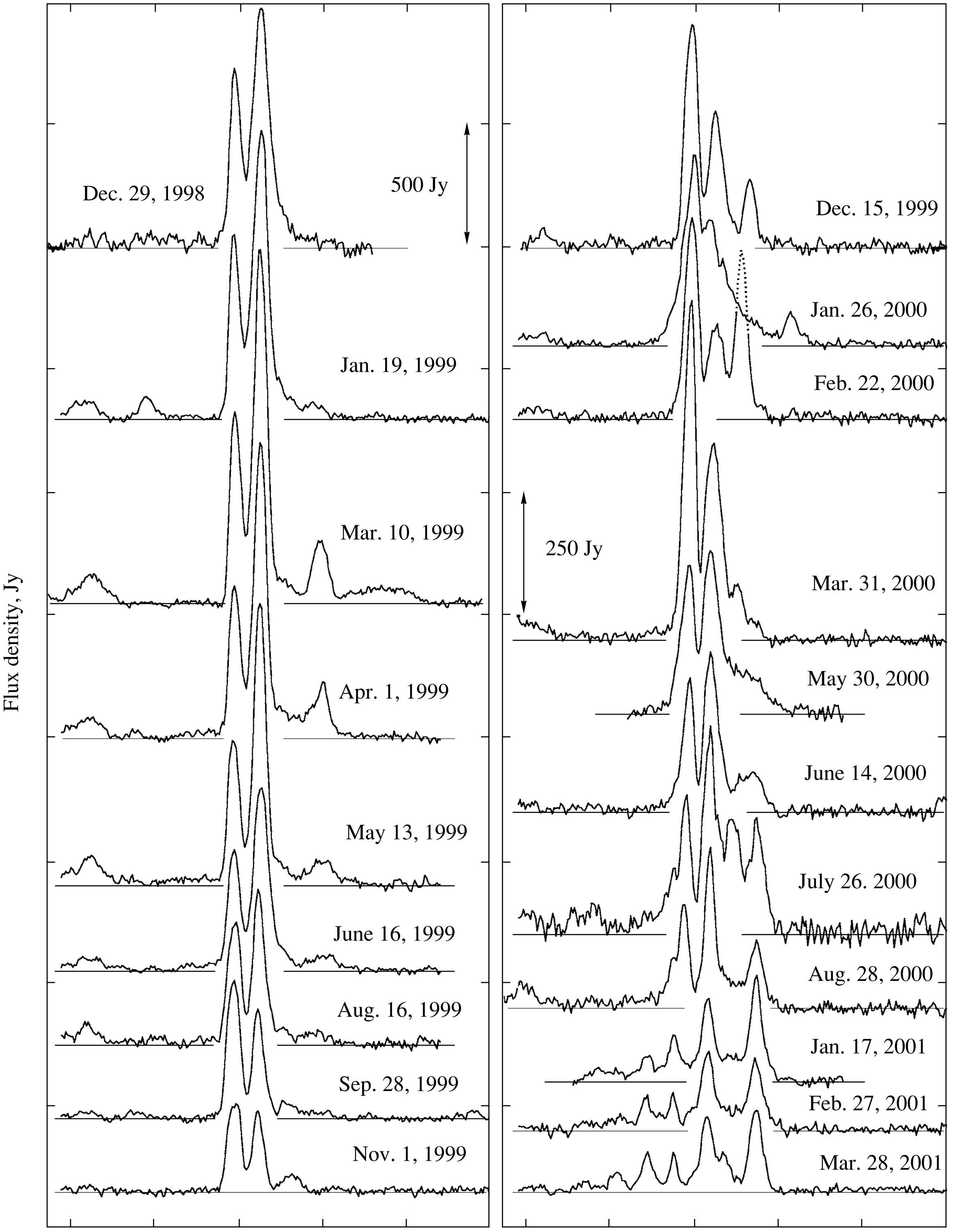


Fig. 1. (Contd.)

deep minimum (in 1984), when the integrated flux fell to 250 Jy km/s. For the maser W31(2), at a distance of 6 kpc, even this is a fairly high level of activity.

Another important result is the presence of a corridor with a width of 600 Jy km/s enclosing fast variations of the emission on a timescale of 0.5–2 years.

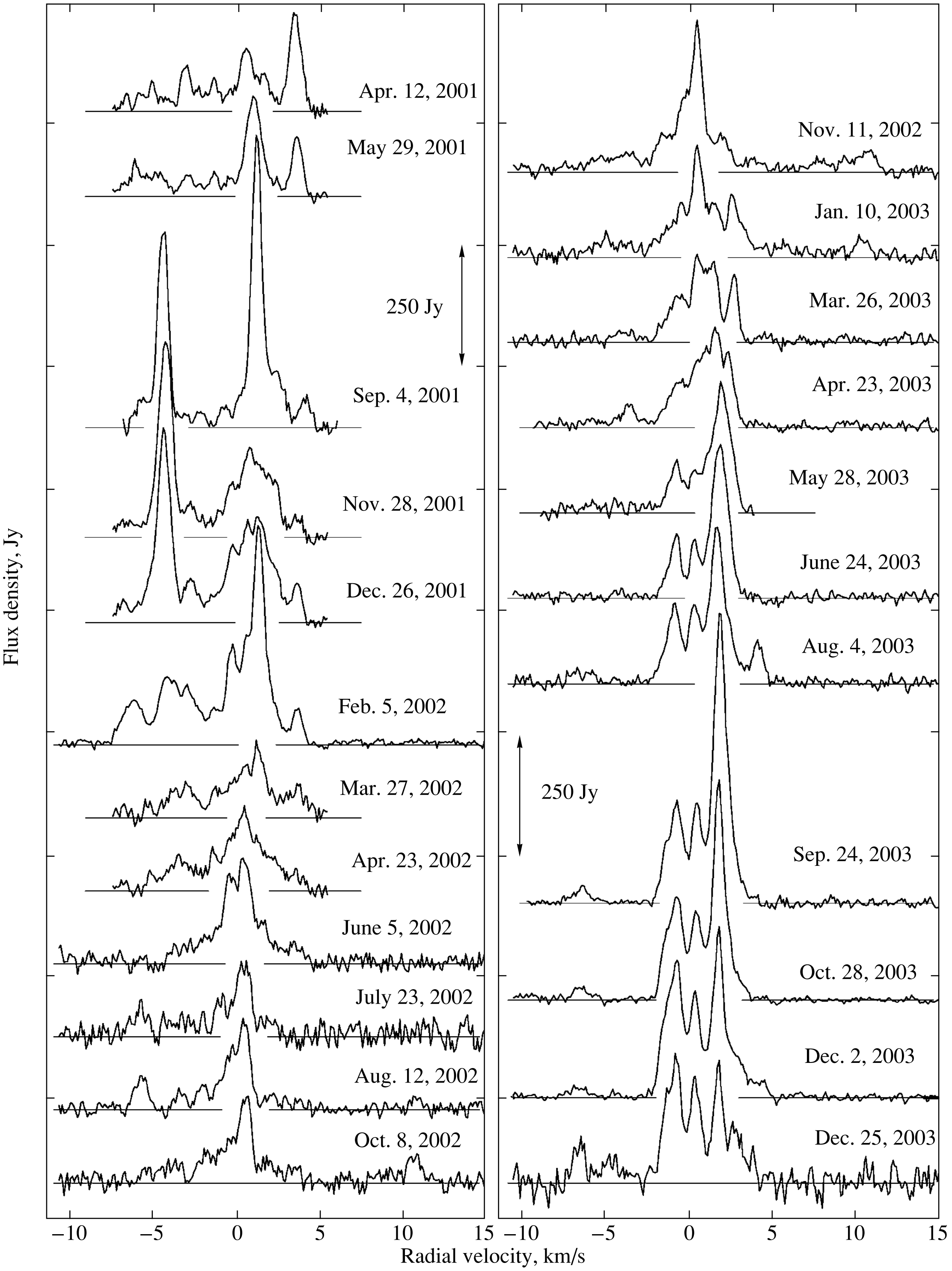


Fig. 1. (Contd.)

Note also that we do not trace any pronounced component of long-period variability in W31(2), such as has been observed in a number of H_2O masers associated with star-forming regions. There is the 12-year interval been the flares, but this is associated only with very powerful flares. If we eliminate these, no other variability components are discernible other than brief flares.

The above results require a somewhat different approach to their interpretation. We note again that the

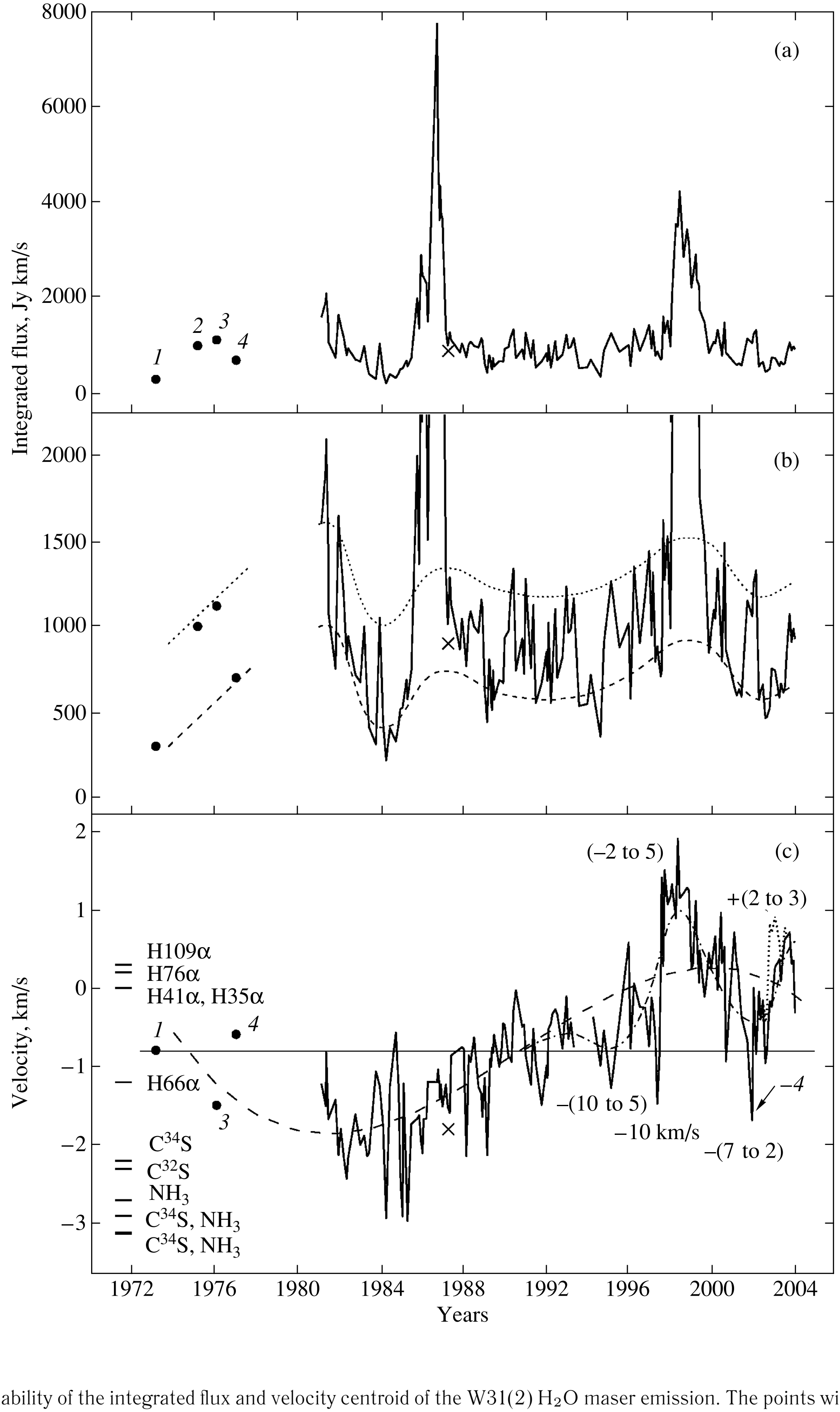


Fig. 2. Variability of the integrated flux and velocity centroid of the W31(2) H_2O maser emission. The points with numbers *1*, *2*, *3*, *4* are taken, respectively, from Yngvesson *et al.* [25], White and Macdonald [20], Fazio *et al.* [19], and Genzel and Downes [18] (see the text for details).

position of the H_2O maser from single-dish observations differs slightly from the position of the ultracompact HII region. No other HII region has been found at the position of the H_2O maser itself. The OH maser is certainly associated with the HII region. There are reasons to believe that the water maser emission is also associated with this HII region, which contains a cluster of OB stars. This suggests a structure in which the H_2O maser pumping is provided by several protostars.

This structure of the region may have resulted in blurring of any long-period maser variations occurring for each protostar separately. The flare components of the luminosity variations of individual protostars are added together, so that we observed continuous short flares of maser emission within the 600-Jy km/s corridor.

3.2. Velocity Centroid

The drift of the velocity centroid followed a complex curve between −3 and 2 km/s. When the data prior to our monitoring are included, the drift curve for 1972–2003 is well approximated by a third-power polynomial. The drift period can be estimated as ∼31–33 years. The total peak-to-peak drift for the smooth curve is 2.2 km/s, and the amplitude is 1.1 km/s relative to the mean value of −0.8 km/s. The two powerful flares fall at different phases of the curve: the first is close to its minimum, and the second is at the maximum (Fig. 2c).

Comparison of the mean velocity centroid ($V_0 = -0.8$ km/s) with the velocities of radio recombination lines of hydrogen and helium shows that these all fall in a narrow interval, from −1.1 to 0.3 km/s, which also includes V_0. This coincidence in velocity may provide an additional argument favoring the association of the H_2O maser with the HII region. The deviation of the centroid from the smooth curve in 1993–2003 (higher modes) was due to the successive appearance of emission at negative and positive velocities. These were essentially short-lived features. The radial velocities of these emission features and the velocity intervals for the appearances of several features are indicated in Fig. 2c.

3.3. Model of the W31(2) Water Maser

Variability of the velocity centroid is not necessarily correlated with the total flux. The total-flux variations are usually due to changes in the pumping level, while regular variations of the velocity centroid may reflect the geometry of the source.

In 1994, Lekht *et al.* [21] carried out a statistical analysis of an ensemble of individual components based on monitoring of W31(2) during 1981–1993. The histogram of the radial-velocity distribution for W31(2) differs strongly from those for other sources, which are usually well described by models with an expanding and rotating envelope. Such envelope (torus) models were considered earlier by Genzel *et al.* [27].

A precessing-disk model is not suitable for several reasons. First, the H_2O maser is associated with the region G10.6−0.4, whose structure differs from that for a precessing-disk model. Second, the amplitude of the velocity variations is small, only 1.1 km/s, while the model requires amplitudes an order of magnitude greater.

In addition, there are problems with interpreting the complex radial-velocity drifts of individual components in W31(2). For instance, the behavior of the feature at −1.5 km/s, associated with the strongest flare, was quite interesting. After the maximum in September 1986 (5750 Jy), this emission feature began to split rapidly into two, and the separation between the components reached 2 km/s in March 1987, with the second component drifting toward negative velocities. The rate of the velocity variations was unusually high. The other feature drifted toward the right, with its velocity changing from −4.7 to −2.5 km/s from February 1988 to June 1989.

The easiest way to explain the observed variability of individual emission features and the 32-year cycle for the variability of the velocity centroid is if these phenomena are associated with the nonstationary accretion of material onto the HII region. By analogy with the OH maser, we can suppose that the H_2O maser spots are located in the halo of the HII region. The possibility that they are associated with fragments of the cloud core that are collapsing toward the cluster of OB stars [7] can also not be ruled out.

A detailed analysis of the two powerful flares and the variability of the emission, and especially the radial velocities, of individual spectral components will be given in forthcoming publications. This is important for the development of the model of the H_2O maser and its connection with the accretion of material onto the HII region.

4. CONCLUSIONS

Let us summarize the main results obtained from our 23-year monitoring of the W31(2) maser source, which is associated with a region of active star formation.

(1) We have presented a catalog of H_2O maser spectra for W31(2), which is associated with the

region G10.6−0.4. The monitoring was carried out on the 22-m radio telescope of the Pushchino Radio Astronomy Observatory in 1981–2003 with a mean interval between observations of about 1.5 months. The spectral resolution was 0.101 km/s. All the spectra were corrected for absorption in the Earth's atmosphere. The spectra contain a large number of emission features. Emission was detected in our monitoring from −14 to +14 km/s.

(2) Two powerful flares whose maxima are separated by 12 years (1985–1986 and 1998–1999) were observed. The rest of the time, the variability of the integrated flux displayed an essentially flaring character, with the amplitudes remaining within a corridor spanning a range of about 600 Jy km/s.

(3) Including observations obtained prior to our monitoring (i.e., considering the interval 1972–2003), the overall drift in the velocity centroid can be approximated by a third-power polynomial that corresponds to a period of 31–33 years. Deviations of the centroid from the smooth curve in 1993–2003 were due to the appearance of emission from groups of features or individual features.

(4) The observed phenomena are explained well by the existing model for the G10.6−0.4 region. The character of the centroid drift may be a consequence of nonstationary accretion, whose cyclic variations have a timescale of ∼31–33 years.

ACKNOWLEDGMENTS

The RT-22 radio telescope is supported by the Ministry of Education, Science, and Technology of the Russian Federation (registration number 01-10). The authors are grateful to the staff of the Pushchino Radio Astronomy Observatory for help with the observations.

REFERENCES

1. G. Westerhout, Bull. Astron. Inst. Netherl. **14**, 215 (1958).
2. W. M. Goss and P. A. Shaver, Australian. J. Phys., Astrophys. Suppl. **14**, 1 (1970).
3. E. C. Reifenstein, T. L. Wilson, B. F. Burke, *et al.*, Astron. Astrophys. **4**, 357 (1970).
4. E. J. Wright, G. G. Fazio, and F. J. Low, Astrophys. J. **217**, 724 (1977).
5. J. L. Caswell, J. D. Murray, R. S. Roger, *et al.*, Astron. Astrophys. **45**, 239 (1975).
6. P. T. P. Ho and A. D. Haschick, Astrophys. J. **248**, 622 (1981).
7. P. T. P. Ho and A. D. Haschick, Astrophys. J. **304**, 501 (1986).
8. P. T. P. Ho, R. I. Klein, and A. D. Haschick, Astrophys. J. **305**, 714 (1986).
9. P. T. P. Ho, R. N. Martin, and A. H. Barret, Astrophys. J. **246**, 761 (1981).
10. S. Guilloteau, T. Forveille, A. Baudry, *et al.*, Astron. Astrophys. **202**, 189 (1988).
11. G. Garay, M. J. Reid, and J. M. Moran, Astrophys. J. **289**, 681 (1985).
12. E. R. Keto, P. T. P. Ho, and M. J. Reid, Astrophys. J. Lett. **323**, L117 (1987).
13. E. Churchwell, C. M. Wamsley, and D. O. S. Wood, Astron. Astrophys. **253**, 541 (1992).
14. H. Hauschildt, R. Gusten, T. G. Phillips, *et al.*, Astron. Astrophys. **273**, L26 (1993).
15. R. Cesaroni, C. M. Walmsley, C. Kompe, and E. Churchwell, Astron. Astrophys. **252**, 278 (1991).
16. B. E. Turner, Astrophys. Lett. **6**, 99 (1970).
17. B. E. Turner and R. H. Rubin, Astrophys. J. Lett. **170**, L113 (1971).
18. R. Genzel and D. Downes, Astron. Astrophys., Suppl. Ser. **30**, 145 (1977).
19. G. G. Fazio, C. J. Lada, D. E. Kleinmann, *et al.*, Astrophys. J. **221**, L77 (1978).
20. G. J. White and G. H. Macdonald, Mon. Not. R. Astron. Soc. **188**, 745 (1979).
21. E. E. Lekht, E. Mendoza-Torres, and R. L. Sorochenko, Astrophys. J. **443**, 222 (1995).
22. R. L. Sorochenko, I. I. Berulism, V. A. Gusev, *et al.*, Tr. Fiz. Inst. im. Lebedeva Akad. Nauk SSSR **159**, 50 (1985).
23. S. A. Zhevakin and A. P. Naumov, Izv. Vyssh. Ucheb. Zaved., Ser. Radiofizika **6**, 674 (1963).
24. N. M. Tseitlin, *Antenna Techniques and Radio Astronomy* (Sov. Radio, Moscow, 1976), p. 49 [in Russian].
25. K. S. Yngvensson, A. G. Cardiasmenos, J. F. Shanley, *et al.*, Astrophys. J. **195**, 91 (1975).
26. G. Comoretto, F. Palagi, R. Cezaroni, *et al.*, Astron. Astrophys., Suppl. Ser. **84**, 179 (1990).
27. R. Genzel, D. Downes, J. M. Moran, *et al.*, Astron. Astrophys. **66**, 13 (1978).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 5–12. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 8–16.
Original Russian Text Copyright © 2005 by Pushkarev, Gabuzda, Vetukhnovskaya, Yakimov.

Blazars with Spine–Sheath Structures

A. B. Pushkarev[1], D. C. Gabuzda[2], Yu. N. Vetukhnovskaya[3], and V. E. Yakimov[4]

[1]*Main (Pulkovo) Astronomical Observatory, St. Petersburg, Russia*
[2]*University College Cork, Cork, Ireland*
[3]*Pushchino Radio Astronomy Observatory, Astro Space Center, Lebedev Physical Institute, Moscow, Russia*
[4]*Astro Space Center, Lebedev Physical Institute, Moscow, Russia*

Received March 17, 2004; in final form, July 15, 2004

Abstract—Multi-frequency (5, 8, 15 GHz) total intensity and polarization VLBA images of three sources from a complete sample of BL Lac objects (0745+241, 1418+546, and 1652+398) are presented together with global VLBI images of the quasar 1055+018 at 1.6 and 5 GHz. These images show signs of a spiral magnetic field in the central channel of the jet, as well as a region of longitudinal magnetic field in "sheaths" at the jet edges, which may indicate an interaction between the jet and the surrounding medium on spatial scales of several parsecs. © *2005 Pleiades Publishing Inc.*

1. INTRODUCTION

BL Lacertae objects are members of an extreme class of active galactic nuclei characterized by a number of peculiar properties: weak (sometimes undetectable) emission lines, rapid and strong variability in all parts of the electromagnetic spectrum, a high degree of linear polarization, and compact radio and pointlike optical structures.

The objects we consider here are members of a complete sample of radio-bright BL Lac objects of the northern sky compiled by Kühr and Schmidt [1] in 1990. The name "complete sample" indicates that we are dealing with a set of sources with homogeneous properties determined via a statistically well defined and fully identified survey. The selection criteria described in [1] are satisfied by 34 sources, which make up the sample of BL Lac objects. Kühr and Schmidt [1] state that the completeness of the sample with regard to the limiting flux density of the survey and the selection criteria adopted is no less than 95%.

Several sources in the sample display polarization structures suggesting the presence of sheaths with longitudinal magnetic fields that form around the jets at some distance from the core (of the order of several parsecs projected onto the plane of the sky). This may indicate that the jet is interacting with a surrounding medium. Here, we present results for three BL Lac objects from the complete sample as the best examples of this behavior (0745+241, 1418+546, and 1652+398), together with new results for the quasar 1055+018.

2. OBSERVATIONS

Multi-frequency polarization observations of 24 of the 34 sources in the complete sample of radio-bright BL Lac objects [1] were carried out in February 1997 (1997.11) using the ten 25-m antennas of the NRAO Very Long Baseline Array: St. Croix, Hancock, North Liberty, Fort Davis, Los Alamos, Pie Town, Kitt Peak, Owens Valley, Brewster, and Mauna Kea. The observations were carried out simultaneously at 5, 8, and 15 GHz in two neighboring frequency channels covering a total bandwidth of 32 MHz. Both right-circular and left-circular polarization were recorded at each of the antennas. The data were correlated at the VLBA correlator in Socorro, New Mexico. The data were reduced in the NRAO AIPS package using standard procedures. The amplitude calibration was performed using system temperatures and gain curves measured at each antenna. After the preliminary calibration, the AIPS tasks CALIB and IMAGR were used iteratively to construct the images.

The strong source 3C 84, which is essentially unpolarized at these frequencies, was used to determine the instrumental polarizations of each of the antennas at each of the three frequencies. The absolute orientation of the polarization angles was calibrated using our VLBA observations of the compact, strongly polarized source 1823+568 together with measurements of its integrated polarization obtained on the Very Large Array one day after the end of the VLBA observations. The polarization-angle calibration assumed that a large fraction of the integrated polarization arises on VLBI scales, so that the total VLBA and VLA polarization angles should coincide.

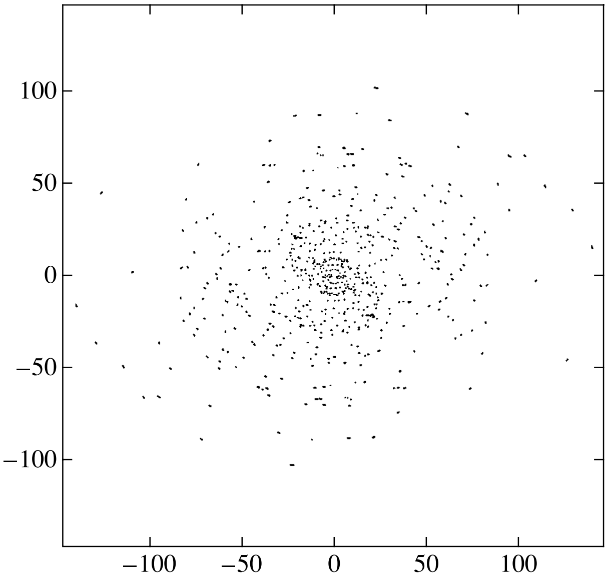


Fig. 1. Coverage of the *uv* plane for the VLBA observations of 1652+398 at 5 GHz. The horizontal and vertical axes are in units of millions of wavelengths.

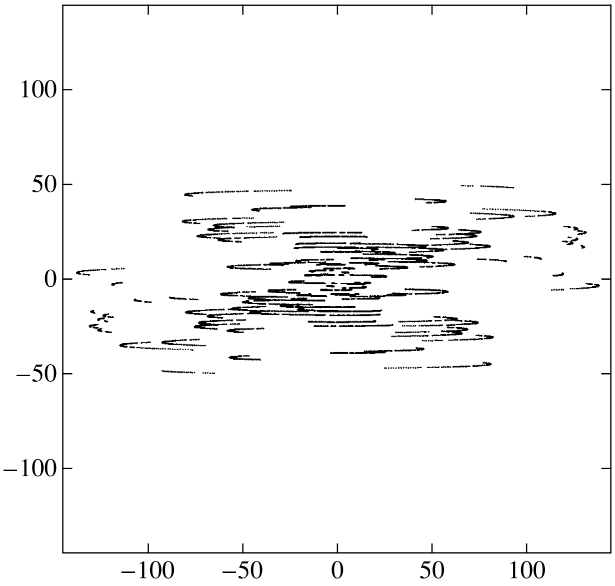


Fig. 2. Same as Fig. 1 for 1055+018.

The final uncertainty in the polarization angles was 2°–3° at all frequencies.

Observations of the quasar 1055+018 were obtained in May 1999 (1999.36) at 5 GHz using the VLBA and the 100 m Effelsberg antenna and at 1.6 GHz using the VLBA, the 100 m Effelsberg antenna, and the 70 m Robledo and Goldstone antennas. The observations were carried out in two neighboring frequency channels with a total bandwidth of 32 MHz, recording both right-circular and left-circular polarization. The data were correlated at the VLBA correlator in Socorro, New Mexico, and the data were calibrated and imaged in the NRAO AIPS package. The amplitude calibration was performed using system temperatures and gain curves measured at each antenna. The alignment of the phases for the two frequency channels was performed by solving for the offsets using a two-minute observation of a strong calibrator with the task FRING and applying the results to the entire set of observations (so-called manual phase calibration). After the preliminary calibration, the AIPS tasks CALIB and IMAGR were used iteratively to construct the images. The amounts of data for the Robledo and Goldstone antennas were relatively small and their quality relatively poor, and we obtained the best final images without using the data for these two antennas.

The coverages of the *uv* plane for the 5 GHz observations of 1652+398 and 1055+018 are shown in Figs. 1 and 2, respectively.

3. RESULTS

The VLBI images presented show contours of total intensity with superimposed linear-polarization vectors. The synthesized beam is shown in the lower left-hand corner of each image. The orientation of the magnetic field is perpendicular to the observed polarization electric vectors if the region emitting the detected synchrotron radiation is optically thin. The observed polarization vectors have been corrected for the known integrated Faraday rotations of these objects [2]. The results for each source are presented and discussed below.

We adopt throughout a Hubble constant of $H_0 = 70h$ km s^{-1} Mpc^{-1} and a deceleration parameter of $q_0 = 0.5$.

3.1. 0745+241

This source has weak [O II] λ 3727 and [OIII] $\lambda\lambda$ 4959, 5007 lines at a redshift of $z = 0.409$ [3]. In previously published VLBI polarization images at 5 GHz [4], the orientation of the magnetic field in the inner jet (close to the core) relative to the jet was unclear, while the magnetic field in the outer jet was longitudinal. Our multi-frequency VLBA observations have enabled us to trace the character of the magnetic field in more detail. The 15 GHz image (Fig. 3, upper panel) shows that the magnetic field in the inner jet is perpendicular to the local jet direction, while the 8- and 5-GHz images (Fig. 3, lower panels) demonstrate that the polarization detected in the outer jet is oriented perpendicular to the jet. Note that the polarization in the outer jet is offset toward the north of the central channel of the jet in total intensity.

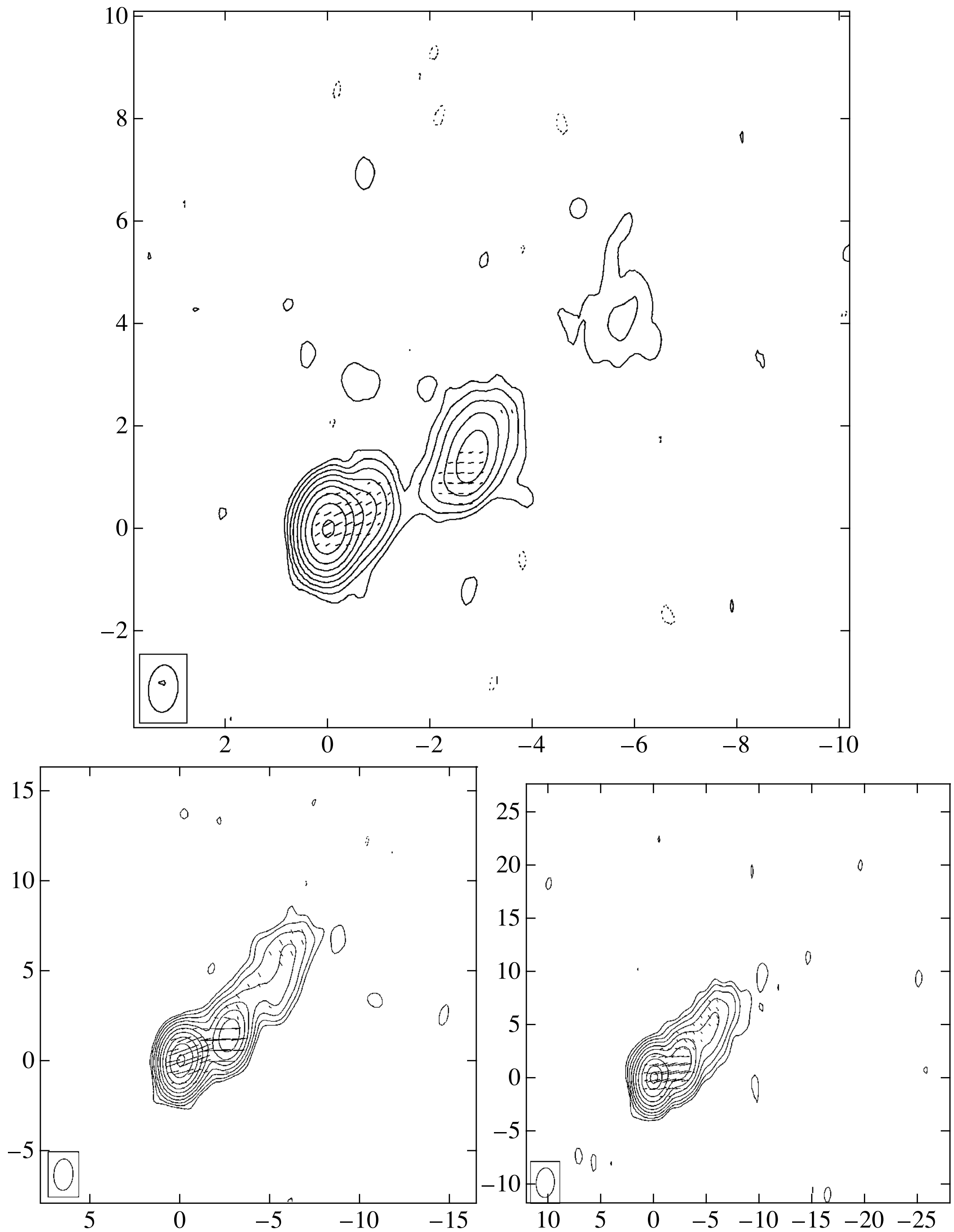


Fig. 3. VLBA images of 0745+241 at 15 GHz (upper), 8 GHz (lower left), and 5 GHz (lower right). The peak flux densities are 378, 354, and 370 mJy/beam, and the lowest contours shown are at 0.35, 0.175, and 0.25% of the peak flux densities. The contour levels increase in steps of a factor of two. The observed polarization electric vectors are superposed, with the length of the vectors being proportional to the polarized flux at the corresponding point. The horizontal and vertical scales are in milliarcseconds.

3.2. 1055+018

The quasar 1055+018 ($z = 0.889$) is a strongly variable, flat-spectrum blazar [5]. The radio spectrum peakd at a frequency of $\sim$15 GHz [6] and remains flat to meter wavelengths.

Monitoring of this source shows that flares occur every one to two years, decaying over roughly 0.9 yrs

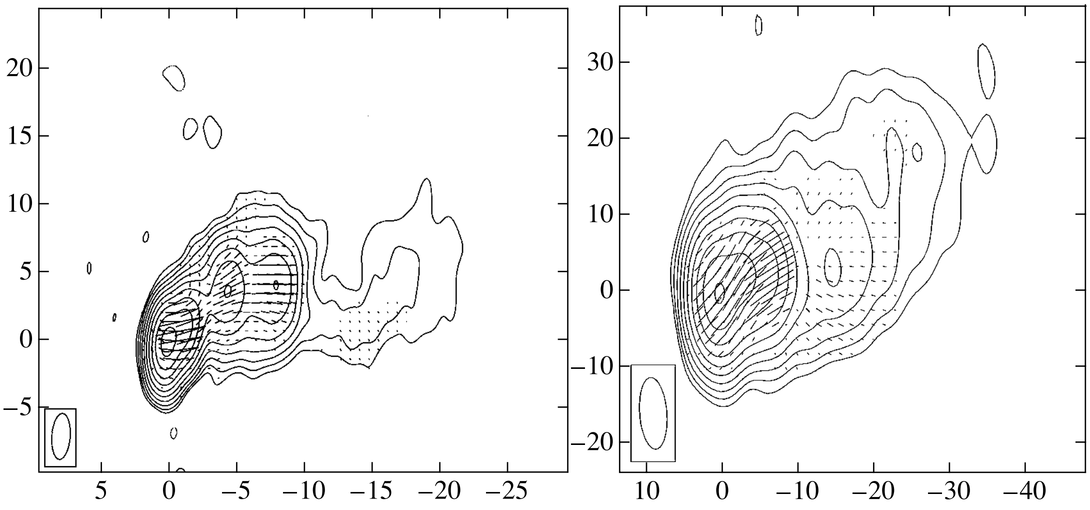


Fig. 4. VLBI images of 1055+018 at 5 GHz (left) and 1.6 GHz (right). The peak flux densities are 999 and 1220 mJy/beam, and the lowest contours shown are at 0.20% of the peak flux densities. The contour levels increase in steps of a factor of two. The observed polarization electric vectors are superposed, with the length of the vectors being proportional to the polarized flux at the corresponding point. The horizontal and vertical scales are in milliarcseconds.

[7]. The low-frequency variability of this object led to observations with European and global VLBI arrays at 1.6 GHz, 608 MHz, and 327 MHz [8–12]. Variations of the 80 and 160 MHz flux density by factors of two to three were reported in 1984 [13]; this was the only source for which such large variations had been observed. Scintillation observations on the Large Phased Array of the Pushchino Radio Astronomy Observatory at 102 MHz showed that the scintillating component, which was $0.1''$ in size, contained only 20% of the total flux, making it difficult to understand the origin of the strong variability [14].

The optical spectrum of the source is not entirely typical for a quasar. The equivalent widths of the optical lines are 5–12 Å [15], suggesting that 1055+018 may be intermediate between a quasar and a BL Lac object.

The source displays a triple FRII structure on kiloparsec scales, with a strong core and two lobes to the north and south of the core; the overall angular size of the structure is $30''$ [16]. The 15 GHz VLBI image [17] shows that the milliarcsecond (mas) jet initially emerges in position angle −47°, then adopts a position angle of −77° about 3 mas from the core. The previously published polarization VLBA image at 6 GHz [18] displayed strong polarization in the core and individual jet components, as well as at the edges of the jet, suggesting that there might be a strong interaction between the jet and surrounding regions of gas. At 43 GHz, the VLBI structure of 1055+018 has a length of 2 mas with significantly polarized components [19]. All these data testified to the presence of activity near the core and demonstrated the need for further studies of this quasar.

Our 5-GHz image (Fig. 4, left) shows that the total-intensity and polarization structure of the quasar had changed appreciably since the observations of Attridge *et al.* [18]. The core is weakly polarized, while the jet components C3, C2, and C1 at distances of 2, 5, and 9 mas from the core are more strongly polarized. The sheath of polarized emission is clearly visible. The flux density of the core decreased from 1.59 Jy at the beginning of 1997 [18] to 1.06 Jy at our observing epoch. Component C3 brightened, while C1 is observed at the same location and with roughly the same brightness. The total flux density of the source in the 5-GHz VLBI image is 2.22 Jy.

The total flux density in the 1.6 GHz image is 2.60 Jy (Fig. 4, right). The jet can be traced to about 30 mas from the core. The direction of the jet changes with distance from the VLBI core, emerging to the northwest near the core, continuing toward the west at distances from 5–15 mas, and then turning toward the north. The polarized emission at 1.6 GHz at a distance of more than 10 mas has components to the north and south of the main channel of the jet. Thus, the sheath of longitudinal magnetic field is also detected at 1.6 GHz.

3.3. 1418+546

The host galaxy of this object was initially classified as an elliptical based on optical observations [20], but later observations provided evidence that it may be a spiral galaxy [21]. The redshift, $z = 0.152$, was obtained from the [O II] λ 3727 and [O III] $\lambda\lambda$ 4959, 5007 lines [3].

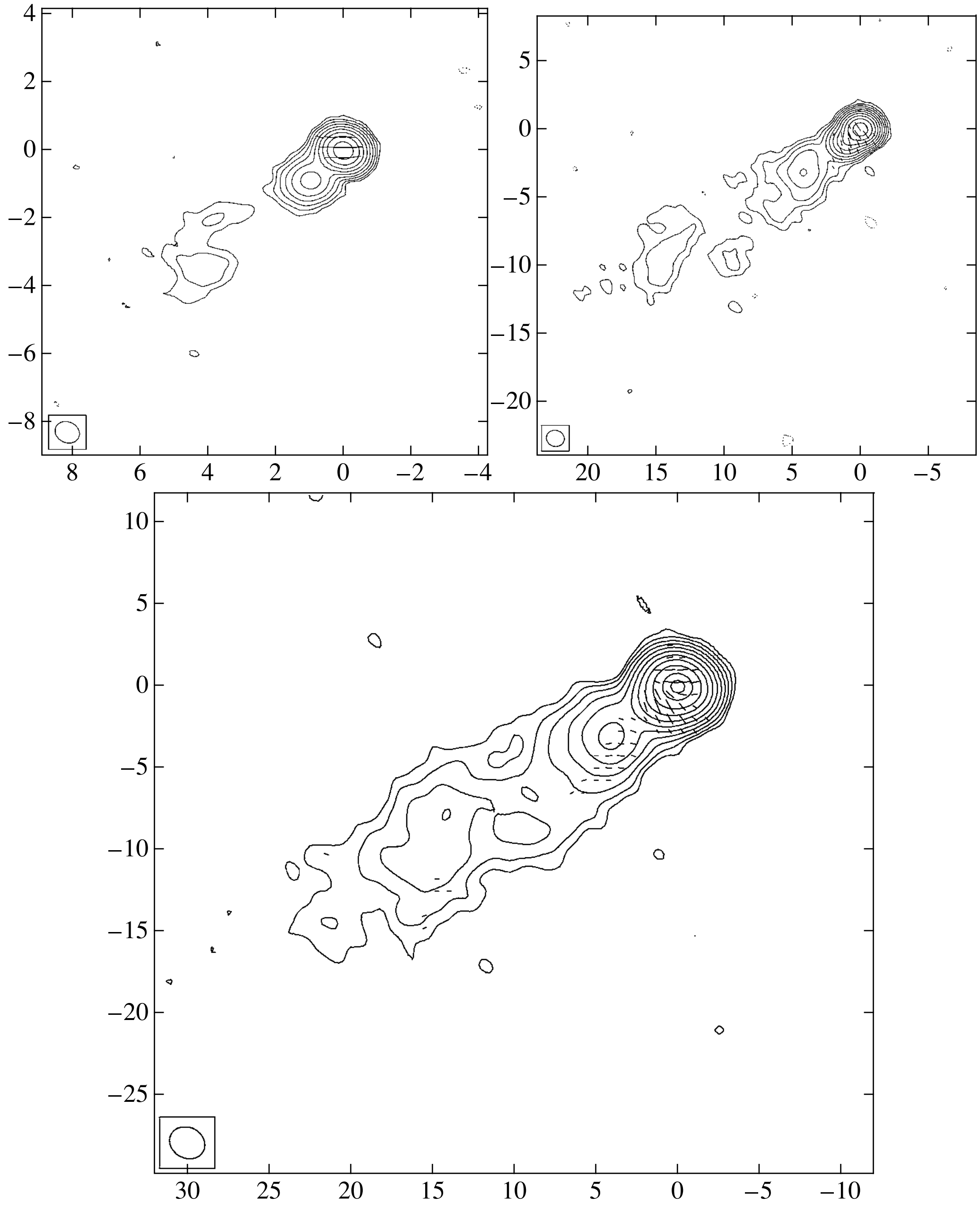


Fig. 5. VLBA images of 1418+546 at 15 GHz (upper left), 8 GHz (upper right), and 5 GHz (lower). The peak flux densities are 565, 583, and 586 mJy/beam, and the lowest contours shown are at 0.35, 0.175, and 0.125% of the peak flux densities. The contour levels increase in steps of a factor of two. The observed polarization electric vectors are superposed, with the length of the vectors being proportional to the polarized flux at the corresponding point. The horizontal and vertical scales are in milliarcseconds.

Our VLBA images of this object (Fig. 5) show a jet extending to the southeast in position angle ∼125°. At 5 GHz, the jet reaches a length of 25 mas, corresponding to a distance of 61 pc projected onto the plane of the sky. At 15 and 8 GHz (Fig. 5, upper panels), polarization is detected only in the core and innermost jet. A richer polarization structure is observed in our 5 GHz image (Fig. 5, lower panel), which displays polarization in both the core region and in the outer, optically thin jet. Like in 0745+241, this polarized emission is offset from the main jet ridge line and the magnetic field in this region is longitudinal. One natural interpretation of this structure is that the jet interacts with the surrounding medium, forming a sheath of longitudinal magnetic field around the VLBI jet.

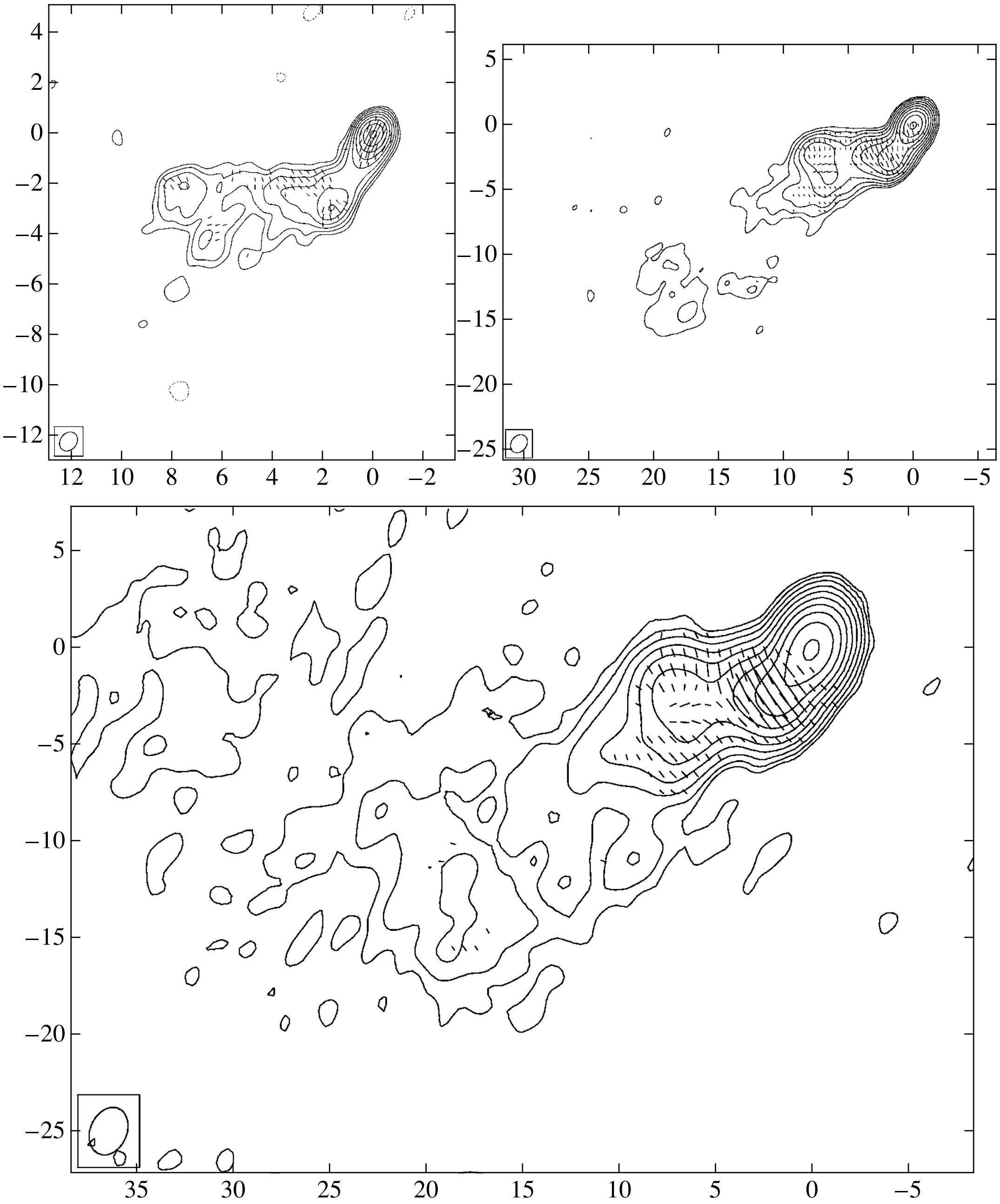


Fig. 6. VLBA images of 1652+398 at 15 GHz (upper left), 8 GHz (upper right), and 5 GHz (lower). The peak flux densities are 361, 440, and 519 mJy/beam, and the lowest contours shown are at 0.35, 0.175, and 0.175% of the peak flux densities. The contour levels increase in steps of a factor of two. The observed polarization electric vectors are superposed, with the length of the vectors being proportional to the polarized flux at the corresponding point. The horizontal and vertical scales are in milliarcseconds.

3.4. 1652+398

This is the most nearby BL Lac object in the sample of Kühr and Schmidt [1], with a redshift of $z = 0.033$ [22]. It is often referred to as Mrk501, and is one of a small number of sources detectable at TeV energies [23].

Our VLBA image of 1652+398 at 5 GHz (Fig. 6, lower panel) shows the very rich total-intensity structure of the source. The jet propagates roughly 25 mas (∼14 pc) toward the southeast in position angle ∼120°, then turns sharply to the northeast, expanding and becoming diffuse.

Our images display a very interesting polarization structure. The image with highest resolution (Fig. 6, upper left) shows that the magnetic field in the inner jet is perpendicular to the jet. In all three maps (at 15,

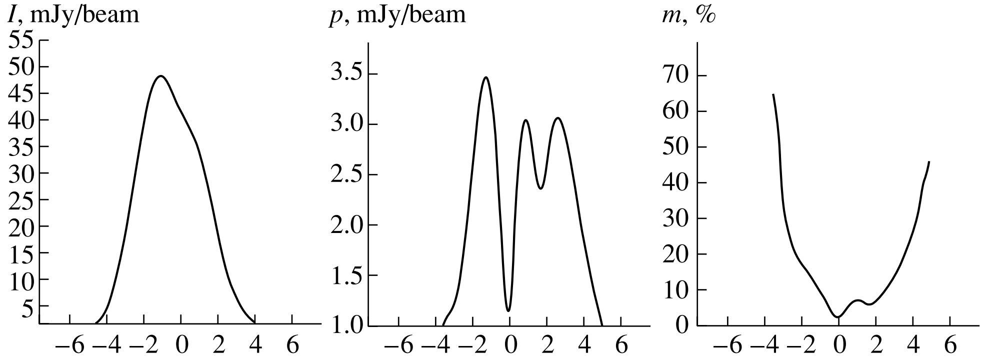


Fig. 7. Transverse profiles of the jet of 1652+398 at 5 GHz 7 mas from the core, where the polarization image shows a change in magnetic-field structure. Shown are profiles for the total intensity I (left), polarized flux p (middle), and degree of polarization m (right). The degree of polarization rises sharply toward the edges of the jet. The horizontal scale is in milliarcseconds.

8, and 5 GHz), we observe a "layered" structure in the polarized emission about 7 mas from the core, with a transverse magnetic field in the central channel of the jet and a longitudinal magnetic field at the edges. The degree of polarization rises sharply toward the jet edges (Fig. 7), indicating a high degree of order for the magnetic field in these regions. This polarization structure suggests the action of shear, as is described in models with a decelerating jet [24]: the plasma in the jet channel moves with relativistic speed, while the material at the edges is slowed due to its interaction with the surrounding medium, which also stretches the magnetic field along the direction of motion of the jet, forming a sheath of longitudinal field. 1652+398 is, to our knowledge, the only example of a source in which a full sheath (i.e., on both sides of the jet) is observed. The transverse magnetic field in the central channel of the jet is mostly likely associated with the toroidal component of the field of the jet itself.

4. CONCLUSION

We have found a number of sources from a complete sample of radio-bright BL Lac objects displaying channel + sheath structures, with a transverse magnetic field along the central spine of the jet surrounded by a sheath with a longitudinal magnetic field. Similar sheathlike structures were first observed in 1055+018 [18] and 0820+225 [25].

These sheaths usually appear at some distance from the core, often in the outer VLBI jet. The possible interpretation of these structures as being consequences of an interaction between the jets and their surrounding medium raises a number of questions. For example, why do we usually observe polarized emission associated with the sheath only on one side of the jet? One possibility is that the sheaths do not always fully surround the jet: they may form only in those regions of the jet where the interaction with the surrounding medium is strongest [18]. Another possibility is that we detect only those parts of the sheath oriented closest to the line of sight. It is likely that both these effects play some role in forming the observed structures.

We believe that the transverse magnetic field observed in the central channels of the jets represents the dominant toroidal component of a helical field associated with the jet. Many theoretical models concerned with the formation and launching of the jets predict that the jets should have a tightly wound helical magnetic field, which develops from a seed field penetrating the accretion disk [26, 27]. The presence of a dominant toroidal field component can explain in a natural way the fact that the jet magnetic fields sometimes remain transverse to the local jet direction even in the presence of appreciably bending of the jet [28, 29].

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16305) and the INTAS foundation (grant 2001-0669). The authors thank the US National Radio Astronomy Observatory for their granting of observing time.

REFERENCES

1. H. Kühr and G. H. Schmidt, Astron. J. **99**, 1 (1990).
2. A. B. Pushkarev, Astron. Zh. **78**, 771 (2001) [Astron. Rep. **45**, 667 (2001)].
3. M. Stickel, J. W. Fried, and H. Kühr, Astron. Astrophys., Suppl. Ser. **80**, 103 (1989).
4. D. C. Gabuzda, A. B. Pushkarev, and T. V. Cawthorne, Mon. Not. R. Astron. Soc. **307**, 725 (1999).
5. C. D. Impey and S. Tapia, Astrophys. J. **354**, 124 (1990).
6. W. K. Gear, Mon. Not. R. Astron. Soc. **267**, 167 (1994).

7. H. Teräsranta and E. Valtaoja, Astron. Astrophys. **116**, 157 (1994).
8. J. Romney, L. Padrielli, N. Bartell, *et al.*, Astron. Astrophys. **135**, 289 (1984).
9. L. Padrielli, J. Romney, N. Bartell, *et al.*, Astron. Astrophys. **165**, 53 (1986).
10. L. Padrielli, F. Mantovani, W. Eastman, *et al.*, Astron. Astrophys. **249**, 351 (1991).
11. D. R. Altschuler, L. I. Gurvits, W. Alef, *et al.*, Astron. Astrophys., Suppl. Ser. **114**, 197 (1995).
12. M. Bondi, L. Padrielli, and R. Fanti, Astron. Astrophys. **308**, 415 (1996).
13. O. B. Slee, Mon. Not. R. Astron. Soc. **209**, 215 (1984).
14. Y. N. Vetukhnovskaya, Astron. Zh. **66**, 480 (1989) [Sov. Astron. **33**, 244, (1989)].
15. R. Falomo, R. Scarpa, and M. Bersanelli, Astrophys. J., Suppl. Ser. **93**, 125 (1994).
16. D. Murphy, I. Browne, and R. Perley, Mon. Not. R. Astron. Soc. **264**, 298 (1993).
17. K. I. Kellermann, R. C. Vermeulen, J. A. Zensus, and M. H. Cohen, Astron. J. **115**, 1295 (1998).
18. J. M. Attridge, D. H. Roberts, and J. F. C. Wardle, Astrophys. J. **518**, L87 (1999).
19. M. L. Lister, A. P. Marscher, and W. K. Gear, Astrophys. J. **504**, 702 (1998).
20. M. Stickel, J. W. Fried, and H. Kühr, Astron. Astrophys., Suppl. Ser. **98**, 393 (1993).
21. R. E. Wurtz, J. T. Stoke, and H. K. C. Yee, Astrophys. J., Suppl. Ser. **103**, 109 (1996).
22. M. H. Ulrich, T. D. Kinman, C. R. Lynds, *et al.*, Astrophys. J. **198**, 261 (1975).
23. E. Pian, G. Vacanti, G. Tagliaferri, *et al.*, Astrophys. J. **492**, 17 (1998).
24. R. A. Laing, *Energy Transport in Radio Galaxies and Quasars*, ASP Conf. Ser. **100**, 241 (1996).
25. D. C. Gabuzda, A. B. Pushkarev, and N. N. Garnich, Mon. Not. R. Astron. Soc. **327**, 1 (2001).
26. G. V. Ustyugova, R. E. V. Lovelace, M. M. Romanova, *et al.*, Astrophys. J. **541**, 21 (2000).
27. M. Nakamura, Y. Uchida, and S. Hirose, New Astron. **6** (2), 61 (2001).
28. D. C. Gabuzda, New Astron. Rev. **43**, 695 (1999).
29. D. C. Gabuzda and A. B. Pushkarev, *Particles and Fields in Radio Galaxies*, ASP Conf. Ser. **250**, 180 (2003).

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 57–70. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 64–78.

Dissipationless Disk Accretion

S. V. Bogovalov and S. R. Kel'ner

Moscow Engineering Power Institute, Moscow, Russia

Received September 13, 2003; in final form, May 27, 2004

Abstract—We consider disk accretion resulting purely from the loss of angular momentum due to the outflow of plasma from a magnetized disk. In this limiting case, the dissipation due to the viscosity and finite electrical conductivity of the plasma can be neglected. We have obtained self-consistent, self-similar solutions for dissipationless disk accretion. Such accretion may result in the formation of objects whose bolometric luminosities are lower than the flux of kinetic energy in the ejected material.

1. INTRODUCTION

Disk-accretion mechanisms can provide clues about a number of processes occurring when matter is accreted onto a gravitating center. It is thus no wonder that such mechanisms have been intensely studied for more than 30 years. The detection of directed ejections from both the cores of active galaxies [1] and Galactic and intergalactic sources, such as SS433 [2], young stellar objects [3], and Galactic superluminal sources [4], has revealed an unambiguous connection between jets and accretion disks. The presence of a directed ejection in a source is essentially always associated with evidence for the presence of an accretion disk [3]. One exception is the jetlike structures in plerions [5], whose nature is obviously different [6]. Thus, the collected observational data indicate that disk accretion is frequently (or even always) accompanied by the narrowly collimated ejection of a substantial fraction of the accreted matter from the source.

This picture gives rise to a number of problems. What is the mechanism that leads to some fraction of the matter falling onto the gravitating center being ejected from the source? How is this matter collimated? What makes the collimated ejection stable enough to propagate to great distances from the source? What accelerates the charged particles that generate the observed radiation in the jets? This list of questions is far from complete. Here, we consider just one of them: what mechanism ejects some of the matter falling onto a gravitating center in the course of disk accretion?

To fall onto a gravitating center, matter with nonzero angular momentum must somehow lose angular momentum. This angular momentum must be carried away by matter, so that disk accretion should inevitably be accompanied by an outflow of matter; the only problem is to determine its rate and direction.

In the classical theory of α accretion disks [7], angular momentum of the accreted matter is lost via viscous, hydrodynamical, turbulent stresses. The classical theory does not include a magnetic field. In [8], however, the magnetic field was used to provide high viscosity. This study also considered the possibility of a matter outflow along the axis of rotation of the accretion disk in a hypercritical regime, when the radiation pressure exceeds the gravitation. A quantitative theory of such an outflow is given in [9], which shows that the outflow already begins in a state with subcritical luminosity. This ejection mechanism may be important in some sources with hyper-Eddington luminosities; however, it is clear that, in many cases (e.g., in young stellar objects), the ejection is driven by another factor, since the radiation pressure in these objects is too small to support this mechanism. Even in SS433, the ejection energy cannot be provided purely by radiation pressure [10].

In standard disk-accretion theory, angular momentum is lost via viscous stresses due to the differential rotation of the disk. Radiation-efficient disks are geometrically thin, with a local temperature that is substantially lower than the virial temperature. It is very difficult to provide an appreciable outflow from the disk under these conditions, since the disk matter is in a very deep potential well. Recently, accretion in radiatively inefficient disks (ADAFs) for which the ion temperature is close to the virial temperature have been widely discussed [13, 14]. Generally speaking, the matter in such disks can flow out vigorously, with the subsequent formation of a jet. However, the existence of ADAF disks is in doubt [15]. Therefore, it remains important to search for alternative mechanisms for the ejection of accreted matter. Ejections due to the presence of a magnetic field in the disks are of particular interest.

1063-7729/05/4901-0057$26.00 © 2005 Pleiades Publishing Inc.

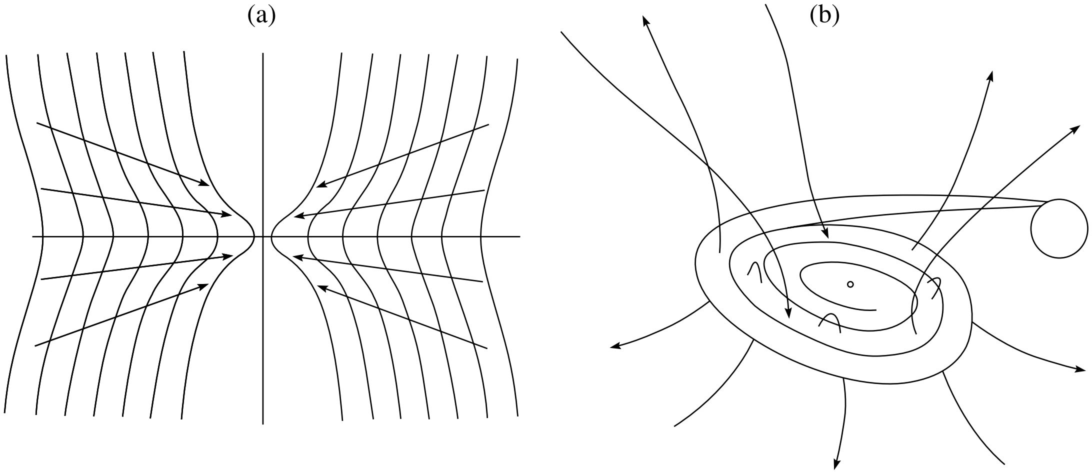


Fig. 1. Role of the magnetic field in different formulations of the problem. Left: plasma diffuses across the magnetic-field lines. Right: plasma falls onto the gravitating center along the magnetic field. The wind flowing from the disk extends some field lines to infinity and carries away some fraction of the angular momentum of the infalling material.

Blandford and Payne [16] have suggested one of the most promising mechanisms for jet outflows. Matter moves in the disk with close to Keplerian velocities. If a magnetic field penetrates the disk, the matter is blown off the disk by centrifugal forces if the inclination of the magnetic-field lines to the disk is smaller than 60°. This mechanism is particularly easy to understand if we imagine the motion of the plasma along a magnetic line to be like that of a bead on a rotating wire. The bead moves in the disk along a circular Keplerian orbit, with the sum of all the forces being zero. If a field line is inclined to the equator at an angle smaller than 60°, any shift of the bead along the magnetic-field line will result in an increase in the centrifugal force and a decrease in the gravitation force. In this case, the position of the bead on the disk is unstable. In addition to forming an ejection, this mechanism can explain the loss of angular momentum, which is carried away from the disk by the wind.

The influence of the matter outflow on the dynamics of the accretion disks themselves was considered in [17–19] in the context of young stellar objects. It was shown that the loss of angular momentum due to the wind from the disk can be more efficient than losses due to turbulent viscosity. This raises the issue that disk accretion onto a gravitating center should take into account the fact that some of the angular momentum is carried away by the disk wind. Ferreira [21, 22] attempted to solve this problem. His formulation is similar to that used previously by Bisnovatyi–Kogan and Ruzmaikin [20] in one of their pioneering studies on the theory of disk accretion in a magnetized plasma. It is assumed that the interstellar plasma is the source of an almost uniform magnetic field. In this case, in order for matter to fall onto the gravitating center, not only must angular momentum be lost, but matter must also diffuse across the magnetic field. In addition to a high turbulent viscosity, this requires a low electrical conductivity of the plasma, which must be associated with a high level of turbulence. In other words, dissipation plays a key role in this case.

We will use another formulation of the problem. Figure 1a presents the difference between our formulation and that used by Ferreira *et al.* [21, 22] (left). Interstellar matter with some initial angular momentum and a relatively uniform magnetic field is accreted onto a gravitating center. The accreted matter diffuses across the field lines, while the forming wind flows out along magnetic lines stretched from interstellar matter.

Figure 1b (right) presents another situation. Matter is accreted onto the compact object along with the frozen-in magnetic field, forming the accretion disk. During the accretion, magnetic lines emerge from the accretion disk and are stretched. A specific mechanism for this stretching of magnetic lines from the disk was proposed and studied numerically in [23]. It results in the formation of a corona and, accordingly, a magnetized disk wind. In this case, the accretion disk itself is a source of the magnetic field of the outflowing wind, just as the Sun is the source of the magnetic field in the solar wind. Note that precisely such a formulation of the problem was used in the study [16] of the centrifugal mechanism for the outflow of matter from an accretion disk.

Here, we study not only the outflow itself, as was done in [16], but also the influence of the plasma outflow from an accretion disk on the accretion. This is not only important for the physics of accretion disks, but is also directly applicable to the interpretation

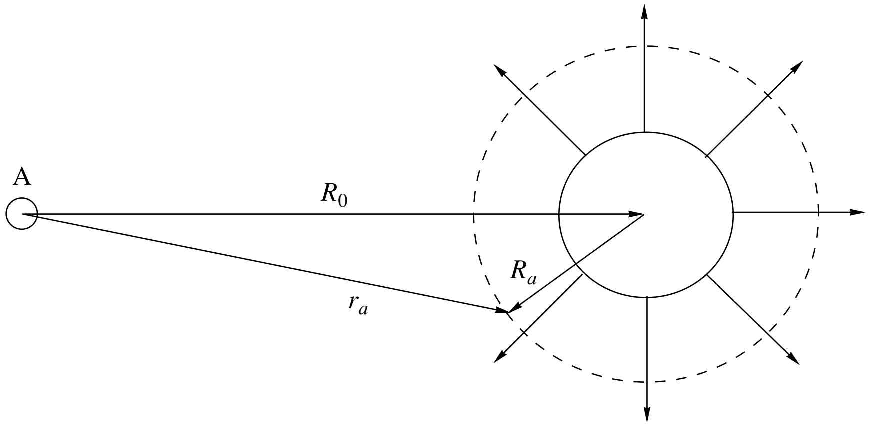


Fig. 2. Loss of angular momentum via a magnetized wind. The gravitating center is at the point A.

of the available observational data. Observations of Galactic superluminal sources [4, 24] are of particular interest, since they provide conclusive evidence that plasma ejections substantially affect the accretion process.

2. FORMULATION OF THE PROBLEM

In general, viscous stresses in accretion disks are undoubtedly important for angular-momentum transport. Here, we primarily consider the influence of the wind on the accretion dynamics. Therefore, we will consider the limiting case when all of the angular momentum is carried away from the disk by the wind with no dissipation. The possibility of such dissipationless accretion can easily be understood based on the following simple consideration. Let one of the components of a binary rotate around a fixed gravitating center (Fig. 2). In the absence of dissipation and wind, this rotation will continue indefinitely.

Let us now suppose that a magnetized wind flows from the rotating object, and that the Alfven radius is R_a. We will assume that the wind is spherically symmetrical and its velocity is v_0. For simplicity, we will use here the mechanical analogy of beads moving along wires for the motion of the plasma in a strong magnetic field. This kind of motion occurs up to the Alfven radius, beyond which the plasma moves essentially freely. We assume that there is no intrinsic rotation of the object, so that its angular momentum is Ml_z, where M is the mass of the object, and $l_z = \Omega R_0^2$ is the specific angular momentum due to its orbital motion.

We denote L to be the total angular momentum contained in a sphere with radius R_a with its center at the point $\mathbf{R}_0(t)$ at time t. At time $t + \delta t$, the angular momentum contained in this sphere (with its center at $\mathbf{R}_0(t)$), will differ from L. The equation for the variation of the angular momentum has the form

$$\frac{dL}{dt} = -\oint \rho[\mathbf{r}_a\mathbf{v}]_z v_n dS. \tag{1}$$

Here, ρ is the density of the plasma, v_n is the velocity normal to the integration surface, the integration is carried out over the Alfven surface, and the z axis is directed along the angular-momentum vector. We can see from Fig. 2 that $\mathbf{r}_a = \mathbf{R}_0 + \mathbf{R}_a$, the velocity of the plasma is $\mathbf{v} = v_0\mathbf{n} + [\boldsymbol{\Omega}\mathbf{r}_a]$, where $\mathbf{n} = \mathbf{R}_a/R_a$ is the unit vector in the direction $\mathbf{R}_a$, and $\boldsymbol{\Omega}$ is the angular-velocity vector for the rotation about the fixed center. After integrating over the directions of the vector $\mathbf{n}$, we obtain

$$\frac{dL}{dt} = -\frac{2}{3}\dot{M}\Omega(2R_0^2 + R_a^2). \tag{2}$$

To interpret this result, note that $\dot{L}_M = -\dot{M}\Omega R_0^2$ is the variation of the angular momentum due to the decrease of the mass of the rotating object. It can be shown that the rate of variation of the angular momentum of the wind is $\dot{L}_V = -\dot{M}\Omega R_0^2/3$. This can be determined as the difference between the angular momenta $L_V(t + \delta t)$ and $L_V(t)$ of the wind particles in a sphere of radius R_a with its center at the fixed point $\mathbf{R}_0(t)$. The sum $\dot{L}_M + \dot{L}_V$ yields the first term in (2). Therefore, the second term describes the rate of variation of the specific angular momentum associated with the orbital motion:

$$\frac{dl_z}{dt} = -\frac{2}{3}\frac{\dot{M}}{M}\Omega R_a^2. \tag{3}$$

Thus, the outflowing wind carries away the angular momentum of the rotating object.

This simple example illustrates an important property of accretion in the presence of a magnetized wind. Since the accretion is not accompanied by heating,

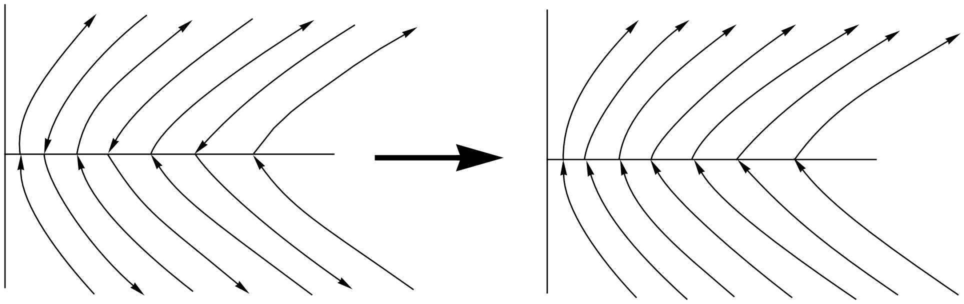


Fig. 3. Independence of the plasma dynamics of the direction of the magnetic-field lines in an ideal MHD approximation. The left panel shows a realistic structure for the flux of the poloidal magnetic field; the total flux of the poloidal magnetic field is zero. In the right panel, the direction of the field lines has changed, while the plasma dynamics remain unaltered.

the wind carries away not only angular momentum, but also rotational energy. Dissipationless accretion can result in the formation of ejections for a relatively low luminosity of the object. This makes this accretion mechanism particularly interesting: it can explain the unusually high efficiency for the transformation of the gravitational energy of the accreted matter into the kinetic energy of the jet in the unusual source SS433. The bolometric luminosity of SS433 is about 3×10^{39} erg/s [25], while the flux of kinetic energy from the object is at least twice this value ($> 6 \times 10^{39}$ erg/s) [12], which is difficult to understand in classical accretion models [10].

It is reasonable to consider dissipationless accretion in an ideal magnetohydrodynamical (MHD) approximation. This leads to the model in the left panel of Fig. 3: the wind stretches magnetic-field lines from the accretion disk. The polarity of the lines varies chaotically, since the total flux of the magnetic field emerging from either side of the accretion disk is, on average, zero. At first glance, it would seem impossible to describe the dynamics of a wind in the presence of a magnetic field with arbitrarily varying polarity at the disk surface. However, there is a circumstance that simplifies the situation radically. As was noted in [26], the dynamics of an ideal plasma are invariant with respect to the direction of the magnetic-field lines. Therefore, if the direction of the field lines varies so that the polarity of the magnetic field is the same on each side of the accretion disk, this does not affect the plasma motion, and we are able to consider a self-consistent outflow of plasma from an accretion disk with an azimuthally symmetrical magnetic field, shown schematically in Fig. 3 (right).

Two features make this approach different from the well-known formulation of Blandford and Payne [16]. First, the outflowing wind must be consistent with the accretion rate in the disk, whereas, in [16], the disk and wind parameters were specified independently. Second, since the plasma moves toward the gravitating center, there is a nonzero azimuthal electrical field, $E_\varphi \neq 0$. Note that a similar formulation was given by Contopoulos [27]. However, his study was not continued, since it was clear that, in this case, magnetic flux should accumulate at the center. If the polarity of the field is the same everywhere, it cannot be annihilated, and the accretion will inevitably be halted by the pressure of the magnetic field. As we noted above, this problem is removed because we are solving for a plasma flow in the presence of a magnetic field of variable polarity. Therefore, the magnetic flux does not accumulate at the center, and is always equal to zero, on average. Since the magnetic-field lines possess different polarities, they can reconnect, so that the magnetic-field pressure at the center will not increase without bound.

Finally, we will assume that the plasma flow in the wind is self-similar. This assumption is natural if we are interested in the motion of plasma directly above a very thin disk away from its edges. Then, the problem has no parameters with the dimension of length, and the flow becomes self-similar. This approximation has been used to describe magnetized winds from accretion disks in numerous studies, starting with [16]. Vlahakis and Tsinganos [28] present a detailed description and classification of all possible types of self-similarity. Note also the study [29], which gives self-similar solutions describing nonstationary accretion.

3. GENERAL RELATIONS

The dimensionless steady-state equations for an ideal, cool plasma (with pressure $P = 0$) can be written (see, for example, [30])

$$\rho(\mathbf{v}\nabla)\mathbf{v} = -\frac{1}{2}\nabla\mathbf{B}^2 + (\mathbf{B}\nabla)\mathbf{B} - \rho\frac{g\mathbf{R}}{R^3}, \qquad (4)$$

$$\mathrm{div}(\rho\mathbf{v}) = 0, \qquad (5)$$

$$\mathrm{div}\mathbf{B} = 0, \qquad (6)$$

$$\mathrm{curl}[\mathbf{v}\mathbf{B}] = 0. \qquad (7)$$

Here, the coordinates, density, velocity, and magnetic field are expressed in units of their characteristic values R_0, ρ_0, v_0, B_0, respectively; in addition,

$$\rho_0 v_0^2 = \frac{B_0^2}{4\pi}. \tag{8}$$

The parameter

$$g = GM/\left(R_0 v_0^2\right), \tag{9}$$

where G is the gravitational constant and M the mass of the central object.

Let us consider azimuthally symmetrical solutions for these equations. The dimensioned variables will be chosen as follows. Let us arbitrarily select a field line, and adopt R_0 as the distance from the center to the point where this line crosses the disk. We will take v_0 to be the Keplerian velocity of a particle at a distance R_0 from the center. We denote $2\pi\Psi_0$ to be the magnetic flux through a circle with radius R_0 located in the disk and with its center on the axis of symmetry. B_0 can then be defined by the relation $B_0 = \Psi_0/R_0^2$. The dimensioned density ρ_0 is defined by relation (8); $g = 1$. Below, we will use only dimensionless variables.

In the azimuthally symmetrical case, the field lines form a two-parameter family of curves, which can be presented in the form

$$\mathbf{R}(\alpha,\psi,\varphi) = \left(r(\alpha,\psi)\cos\phi, r(\alpha,\psi)\sin\phi, z(\alpha,\psi)\right), \tag{10}$$

where $\phi = \varphi + \eta(\alpha,\psi)$. The parameters ψ, φ specify a field line, and α varies along a given line ($\mathbf{B} \sim \partial\mathbf{R}/\partial\alpha$). The substitution $\varphi \to \varphi - \varphi_0$ is equivalent to a rotation by φ_0 of the field line as a whole. The functions of two variables $r(\alpha,\psi)$, $z(\alpha,\psi)$, and $\eta(\alpha,\psi)$ are to be determined from the solution of the MHD equations.

Following [31, 32], we will call the parameters α, ψ, and φ the "frozen-in" coordinates. The cylindrical coordinates (r,ϕ,z) can be expressed in terms of the frozen-in coordinates by the formulas

$$r = r(\alpha,\psi), \quad \phi = \varphi + \eta(\alpha,\psi), \quad z = z(\alpha,\psi).$$

Let us make a transformation from Cartesian to frozen-in coordinates in (4)–(7). J denotes the Jacobian of the transformation:

$$J = \frac{\partial(X,Y,Z)}{\partial(\alpha,\psi,\varphi)} = \mathbf{R}_\alpha[\mathbf{R}_\psi\mathbf{R}_\varphi], \tag{11}$$

where the subscripts denote differentiation with respect to the corresponding variables ($\mathbf{R}_\alpha \equiv \partial\mathbf{R}/\partial\alpha, \ldots$). Since field lines do not cross, there is a one-to-one relation between the frozen-in and Cartesian coordinates, so that the Jacobian J never vanishes. Note the following useful equalities, which can be used to easily change between Cartesian and frozen-in coordinates:

$$\mathbf{R}_\alpha\nabla = \frac{\partial}{\partial\alpha}, \quad \mathbf{R}_\psi\nabla = \frac{\partial}{\partial\psi}, \quad \mathbf{R}_\varphi\nabla = \frac{\partial}{\partial\varphi} \tag{12}$$

and also

$$\frac{\mathbf{R}_\alpha}{J} = [\nabla\psi,\nabla\varphi], \quad \frac{\mathbf{R}_\psi}{J} = [\nabla\varphi,\nabla\alpha], \tag{13}$$
$$\frac{\mathbf{R}_\varphi}{J} = [\nabla\alpha,\nabla\psi],$$

and

$$\frac{[\mathbf{R}_\psi,\mathbf{R}_\varphi]}{J} = \nabla\alpha, \quad \frac{[\mathbf{R}_\varphi,\mathbf{R}_\alpha]}{J} = \nabla\psi, \tag{14}$$
$$\frac{[\mathbf{R}_\alpha,\mathbf{R}_\psi]}{J} = \nabla\varphi.$$

The operator ∇ in frozen-in coordinates has the form

$$\nabla = \frac{1}{J}\left\{[\mathbf{R}_\psi\mathbf{R}_\varphi]\frac{\partial}{\partial\alpha} + [\mathbf{R}_\varphi\mathbf{R}_\alpha]\frac{\partial}{\partial\psi} + [\mathbf{R}_\alpha\mathbf{R}_\psi]\frac{\partial}{\partial\varphi}\right\}. \tag{15}$$

Suppose that the magnetic field strength is specified by the equality

$$\mathbf{B} = \frac{\mathbf{R}_\alpha}{J} = [\nabla\psi,\nabla\varphi]. \tag{16}$$

In this case, (6) becomes an identity (it essentially expresses the field in terms of the Euler potentials). Expression (16) implies that we take the magnetic flux (in units of $2\pi\Psi_0$) through a circle with radius r as the frozen-in coordinate ψ. Note that the ratio $\mathbf{R}_\alpha/J$ is invariant with respect to the reparameterization $\alpha \to \alpha' = g(\alpha,\psi,\varphi)$; i.e.,

$$\frac{\mathbf{R}_\alpha}{J} = \frac{\mathbf{R}_{\alpha'}}{J'},$$

where $g(\alpha,\psi,\varphi)$ is an arbitrary function and $J' = \partial(X,Y,Z)/\partial(\alpha',\psi,\varphi) = J/g_\alpha$. This can be used to assign an arbitrary value to the Jacobian J without varying the field strength. It is convenient to relate J with the plasma density:

$$J = 1/\rho, \tag{17}$$

in which case the magnetic field is

$$\mathbf{B} = \rho\mathbf{R}_\alpha. \tag{18}$$

Let us introduce the three mutually orthogonal unit vectors

$$\mathbf{e}^{(r)} = (\cos\phi, \sin\phi, 0), \tag{19}$$
$$\mathbf{e}^{(\phi)} = (-\sin\phi, \cos\phi, 0), \quad \mathbf{e}^{(z)} = (0,0,1).$$

We obtain from (10)

$$\begin{cases} \mathbf{R}_\alpha = r_\alpha\mathbf{e}^{(r)} + r\eta_\alpha\mathbf{e}^{(\phi)} + z_\alpha\mathbf{e}^{(z)}, \\ \mathbf{R}_\psi = r_\psi\mathbf{e}^{(r)} + r\eta_\psi\mathbf{e}^{(\phi)} + z_\psi\mathbf{e}^{(z)}, \\ \mathbf{R}_\varphi = r\mathbf{e}^{(\phi)}. \end{cases} \tag{20}$$

Therefore, the Jacobian of the transformation is

$$J = r(z_\alpha r_\psi - z_\psi r_\alpha), \tag{21}$$

and, in accordance with (17), the plasma density is

$$\rho = \frac{1}{r(z_\alpha r_\psi - z_\psi r_\alpha)}. \tag{22}$$

Let us expand the plasma velocity into three noncoplanar vectors:

$$\begin{aligned}\mathbf{v} &= f\mathbf{R}_\alpha + \Omega\mathbf{R}_\varphi + \varepsilon\mathbf{R}_\psi \qquad (23)\\ &= (fr_\alpha + \varepsilon r_\psi)\mathbf{e}^{(r)} + (fz_\alpha + \varepsilon z_\psi)\mathbf{e}^{(z)}\\ &\quad + r(f\eta_\alpha + \varepsilon\eta_\psi + \Omega)\mathbf{e}^{(\phi)},\end{aligned}$$

where the functions f, Ω, ε may depend on the frozen-in coordinates.

It turns out that, in both the stationary and nonstationary cases [31, 32], Eqs. (5) and (7) in frozen-in coordinates can be integrated in general form. Substituting (23) and (17) into the continuity equation (5) and transforming to frozen-in coordinates, we obtain

$$\frac{\partial f}{\partial\alpha} + \frac{\partial\Omega}{\partial\varphi} + \frac{\partial\varepsilon}{\partial\psi} = 0. \tag{24}$$

The electrical-field strength is

$$\mathbf{E} = -\frac{1}{c}[\mathbf{vB}] = \frac{1}{c}(\varepsilon\nabla\varphi - \Omega\nabla\psi). \tag{25}$$

It follows that (7) in frozen-in coordinates has the form

$$\begin{aligned}&\mathrm{curl}[\mathbf{vB}] = [\nabla\Omega\nabla\psi] - [\nabla\varepsilon\nabla\varphi] \qquad (26)\\ &= \rho\left(\frac{\partial\Omega}{\partial\alpha}\mathbf{R}_\varphi + \frac{\partial\varepsilon}{\partial\alpha}\mathbf{R}_\psi - \left(\frac{\partial\Omega}{\partial\varphi} + \frac{\partial\varepsilon}{\partial\psi}\right)\mathbf{R}_\alpha\right) = 0.\end{aligned}$$

The derivative $\partial\Omega/\partial\varphi = 0$ due to the assumed azimuthal symmetry. Solving (24) and (26), we find that $\varepsilon = \mathrm{const}$, while the functions f and Ω may depend only on the single variable ψ. In this case, (5) and (7) are identically satisfied.

The gradient $\nabla\psi$ has only r and z components; the first term in (25) specifies the azimuthal component of the electrical field:

$$E_\phi = \frac{\varepsilon}{cr}. \tag{27}$$

The velocity of the accretion disk $\mathbf{u}$ is also expressed in frozen-in coordinates. Let us suppose that the disk surface corresponds to the parameter $\alpha = 0$; i.e., $z(0, \psi) = 0$. Then, the derivative $z_\psi(0, \psi) = 0$. The plasma density above the disk surface is

$$\rho_0 = \frac{1}{rz_\alpha r_\psi}\Big|_0. \tag{28}$$

The z component of the magnetic field is

$$B_{z0} = \rho z_\alpha\big|_0 = \frac{1}{rr_\psi}\Big|_0. \tag{29}$$

The radial component of the electrical field at the disk surface is

$$E_{r0} = -\frac{1}{c}[\mathbf{vB}]_r\big|_0 = -\frac{1}{r_\psi c}(\Omega + \varepsilon\eta_\psi)\big|_0, \tag{30}$$

while the azimuthal component is defined by (27). Using the frozen-in condition

$$\mathbf{E} = -\frac{1}{c}[\mathbf{uB}],$$

the continuity of the z component of the magnetic field at the boundary between the plasma and disk, and the continuity of the r and ϕ components of the electrical field, we obtain the components of the disk velocity

$$u_r = \varepsilon r_\psi\big|_0, \quad u_\phi = r(\Omega + \varepsilon\eta_\psi)\big|_0. \tag{31}$$

After transforming to frozen-in coordinates, the Euler equation (4) assumes the form

$$\begin{aligned}&\left(f\frac{\partial}{\partial\alpha} + \Omega\frac{\partial}{\partial\varphi} + \varepsilon\frac{\partial}{\partial\psi}\right)(f\mathbf{R}_\alpha + \Omega\mathbf{R}_\varphi + \varepsilon\mathbf{R}_\psi) \qquad (32)\\ &- \frac{\partial}{\partial\alpha}(\rho\mathbf{R}_\alpha) = -\frac{1}{2}\Big\{[\mathbf{R}_\psi\mathbf{R}_\varphi]\frac{\partial}{\partial\alpha} + [\mathbf{R}_\varphi\mathbf{R}_\alpha]\frac{\partial}{\partial\psi}\\ &+ [\mathbf{R}_\alpha\mathbf{R}_\psi]\frac{\partial}{\partial\varphi}\Big\}(\rho\mathbf{R}_\alpha)^2 - g\frac{\mathbf{R}}{R^3}.\end{aligned}$$

Substituting $\mathbf{R}$ in the form (10) and ρ in the form (22), we can obtain equations for the functions $r(\alpha, \psi)$, $z(\alpha, \psi)$, and $\eta(\alpha, \psi)$.

4. BOUNDARY CONDITIONS

Let us study the relation between the field components and the parameters of the accretion disk, assuming the disk is infinitely thin (which is valid for a cool plasma). We will specify the volume density of the disk in the form

$$\rho_d(r, z) = \sigma(r)\delta(z), \tag{33}$$

where $\delta(z)$ is a delta function and $\sigma(r)$ is the mass surface density. Equation (4) is valid for $z > 0$ (or $z < 0$); i.e., beyond the disk. The dynamics of the plasma in the disk are specified by the equation

$$\begin{aligned}&\rho_d(\mathbf{u}\nabla)\mathbf{u} + \rho(\mathbf{v}\nabla)\mathbf{v} \qquad (34)\\ &= -\frac{1}{2}\nabla\mathbf{B}^2 + (\mathbf{B}\nabla)\mathbf{B} - (\rho_d + \rho)\frac{g\mathbf{R}}{R^3}.\end{aligned}$$

We will integrate (34) over z in the small interval $(-z_0, z_0)$ and take the limit $z_0 \to +0$. This operation will be applied to each term of the equation. In cylindrical coordinates,

$$\nabla = \mathbf{e}^{(r)}\frac{\partial}{\partial r} + \mathbf{e}^{(z)}\frac{\partial}{\partial z} + \mathbf{e}^{(\varphi)}\frac{\partial}{r\partial\phi}, \tag{35}$$

so that

$$\rho_d(\mathbf{u}\nabla)\mathbf{u} = \sigma(r)\delta(z) \tag{36}$$
$$\times\left(u_r\frac{\partial}{\partial r} + u_z\frac{\partial}{\partial z} + u_\phi\frac{\partial}{r\partial\phi}\right)(\mathbf{e}^{(r)}u_r + \mathbf{e}^{(z)}u_z$$
$$+\,\mathbf{e}^{(\phi)}u_\phi) = \sigma(r)\delta(z)$$
$$\times\left[\left(u_r\frac{\partial u_r}{\partial r} - \frac{u_\phi^2}{r}\right)\mathbf{e}^{(r)} + u_r\left(\frac{\partial u_\phi}{\partial r} + \frac{u_\phi}{r}\right)\mathbf{e}^{(\phi)}\right].$$

When deriving (36), we took into account azimuthal symmetry, the equality $u_z = 0$, and the relations $\partial\mathbf{e}^{(r)}/\partial\phi = \mathbf{e}^{(\phi)}$ and $\partial\mathbf{e}^{(\phi)}/\partial\phi = -\mathbf{e}^{(r)}$. Integrating (36), we obtain

$$\int\limits_{-z_0}^{z_0}\rho_d(\mathbf{u}\nabla)\mathbf{u}dz = \sigma(r) \tag{37}$$
$$\times\left[\left(u_r\frac{\partial u_r}{\partial r} - \frac{u_\phi^2}{r}\right)\mathbf{e}^{(r)} + u_r\left(\frac{\partial u_\phi}{\partial r} + \frac{u_\phi}{r}\right)\mathbf{e}^{(\phi)}\right].$$

If the magnetic field and the plasma velocity above the accretion disk ($z > 0$) are known, then the solution in the region under the disk continues in accordance with the rules

$$B_z \to B_z, \quad B_r \to -B_r, \quad B_\phi \to -B_\phi. \tag{38}$$

$$v_z \to -v_z, \quad v_r \to v_r, \quad v_\phi \to v_\phi. \tag{39}$$

Since the field components $B_r, B_\phi \neq 0$ at the equator (for $z = +0$), this implies that the field line is discontinuous there.

Using relations (39), it can be shown that, in the limit $z_0 \to +0$

$$\int\limits_{-z_0}^{z_0}\rho(\mathbf{v}\nabla)\mathbf{v}dz = 0.$$

Next, we obtain

$$\int\limits_{-z_0}^{z_0}\nabla B^2 dz = 0, \tag{40}$$

since B^2 is a continuous function of z.

The term $(\mathbf{B}\nabla)\mathbf{B}$ will be written in the form

$$(\mathbf{B}\nabla)\mathbf{B} = \left(B_r\frac{\partial}{\partial r} + B_z\frac{\partial}{\partial z} + B_\phi\frac{\partial}{r\partial\phi}\right) \tag{41}$$
$$\times\left(B_r\mathbf{e}^{(r)} + B_z\mathbf{e}^{(z)} + B_\phi\mathbf{e}^{(\phi)}\right).$$

If the integration is taken over an infinitely small interval, a nonzero contribution will be obtained only from terms containing the derivative $\partial/\partial z$ (a derivative of a discontinuous function yields a δ function). After the integration, we obtain

$$\int\limits_{-z_0}^{z_0}(\mathbf{B}\nabla)\mathbf{B}dz = 2B_z\left(B_r\mathbf{e}_r + B_\phi\mathbf{e}_\phi\right). \tag{42}$$

The integration of the last term in (34) is trivial due to the presence of the δ function:

$$\int\limits_{-z_0}^{z_0}\frac{\rho_d g\mathbf{R}}{R^3}dz = \frac{\sigma g}{r^2}\mathbf{e}^{(r)}. \tag{43}$$

Collecting these various relations, we obtain two equalities (for the r and ϕ components in the equation), which are a direct consequence of (34) in the model with an infinitely thin disk:

$$u_r\frac{\partial u_r}{\partial r} - \frac{u_\phi^2}{r} = \frac{2}{\sigma}B_rB_z - \frac{g}{r^2}, \tag{44}$$

$$\frac{u_r}{r}\frac{\partial}{\partial r}(ru_\phi) = \frac{2}{\sigma}B_\phi B_z. \tag{45}$$

The conservation of mass yields another relation:

$$\frac{1}{r}\frac{\partial}{\partial r}(r\sigma u_r) + 2j_z = 0, \tag{46}$$

where $j_z = \rho v_z$ is the density of the z component of the mass flux. In (44)–(46), the magnetic-field components and plasma parameters are taken for the disk surface ($z = +0$).

Equations (44)–(46) play the role of boundary conditions, which must be satisfied by the solution. They were derived from the equations of motion in differential form. Obviously, the same conditions can be obtained using the integral form of the equations of motion.

5. SELF-SIMILAR SOLUTIONS

We are primarily concerned here with the plasma dynamics immediately above the disk, at distances z much smaller than the radius of the disk R_{disk}. In the limit $z \ll R_{\text{disk}}$, only two parameters with the dimensions of length remain in the problem: z and r. Therefore, the solution will be self-similar in this limit [11]. Let us underscore an important feature of these solutions: they describe flows only at small distances from the disk, and are not applicable at distances comparable to or exceeding the size of the disk.

Equations (4)–(7) have self-similar solutions of the form

$$\begin{cases}\mathbf{v}(r,z,\phi) = r^{-\delta_v}\tilde{\mathbf{v}}(z/r,\phi),\\ \rho(r,z) = r^{-\delta_\rho}\tilde{\rho}(z/r),\\ \mathbf{B}(r,z,\phi) = r^{-\delta_B}\tilde{\mathbf{B}}(z/r,\phi).\end{cases} \tag{47}$$

The superscripts δ_v, δ_ρ, and δ_B are determined from the following conditions. Substituting (47) into (4) leads to the equations

$$2\delta_B - \delta_\rho = 2\delta_v = 1. \tag{48}$$

Equality (27) is satisfied under the condition

$$\delta_v + \delta_B = 1. \tag{49}$$

Thus,

$$\delta_v = \delta_B = \frac{1}{2}, \quad \delta_\rho = 0 \tag{50}$$

and hence,

$$\begin{cases} \mathbf{v}(r, z, \phi) = r^{-1/2}\tilde{\mathbf{v}}(z/r, \phi), \\ \rho(r, z) = \tilde{\rho}(z/r), \\ \mathbf{B}(r, z, \phi) = r^{-1/2}\tilde{\mathbf{B}}(z/r, \phi). \end{cases} \tag{51}$$

Note that it is possible to determine the subscripts unambiguously only for $u_r \neq 0$; i.e., when there is advection of the plasma toward the center of gravitation. When $u_r = 0$, one of the subscripts is arbitrary, so that there is a single-parameter family of self-similar solutions. A large number of such solutions have been studied previously (see, for example, [28] and references therein). For clarity, we will compare our results with those of [16], which is similar to our study in terms of the formulation for the wind. The case $\delta_\rho = 3/2$, $\delta_B = 5/4$ was considered in [16].

In frozen-in coordinates, the law of similarity (51) is described by the relation

$$\mathbf{R}(\alpha, \psi, \varphi) = \psi^{2/3}\tilde{\mathbf{R}}\left(\frac{\alpha}{\psi}, \varphi\right). \tag{52}$$

Therefore, the two-variable functions in (10) are expressed in terms of one-variable functions. Denoting $s = \alpha/\psi$, we can write

$$r(\alpha, \psi) = \psi^{2/3}\tilde{r}(s), \quad z(\alpha, \psi) = \psi^{2/3}\tilde{z}(s), \tag{53}$$
$$\eta(\alpha, \psi) = \tilde{\eta}(s).$$

Let us verify the equivalence of (51) and (53). Substituting (53) into (22), we derive the plasma density:

$$\rho(s) = \frac{3}{2}\frac{1}{\tilde{r}(\tilde{r}\tilde{z}_s - \tilde{z}\tilde{r}_s)}, \tag{54}$$

where the subscript s denotes differentiation with respect to s. This function depends only on s, or, if we switch to cylindrical coordinates, on the ratio z/r, as in (51). The components of the magnetic-field strength are

$$B_r = \psi^{-1/3}\rho\tilde{r}_s, \quad B_z = \psi^{-1/3}\rho\tilde{z}_s, \tag{55}$$
$$B_\phi = \psi^{-1/3}\rho\tilde{r}\tilde{\eta}_s.$$

We find for the components of the plasma velocity

$$v_r = \psi^{-1/3}\left(\frac{2}{3}\varepsilon\tilde{r} + (f - \varepsilon s)\tilde{r}_s\right), \tag{56}$$

$$v_z = \psi^{-1/3}\left(\frac{2}{3}\varepsilon\tilde{z} + (f - \varepsilon s)\tilde{z}_s\right), \tag{57}$$

$$v_\phi = \psi^{-1/3}\tilde{r}\left(\psi\Omega + (f - \varepsilon s)\tilde{\eta}_s\right). \tag{58}$$

The law of similarity $\mathbf{v} \sim r^{-1/2}$ will be satisfied if $f =$ const, while Ω depends on ψ as follows:

$$\Omega(\psi) = \tilde{\Omega}/\psi. \tag{59}$$

$\tilde{\Omega}$ represents the angular velocity of the selected field line in units of the Keplerian angular velocity.

Substituting (52) into (32), we obtain a system of three ordinary second-order differential equations[1] for the functions $\tilde{r}(s)$, $\tilde{z}(s)$, and $\tilde{\eta}(s)$. Since second derivatives appear in the equations in linear form, the system can be solved for the second derivatives. For $\varepsilon \neq 0$, the right-hand sides of the equations depend explicitly on s; f and s appear in the equations only in the combination $f - \varepsilon s$. The parameter s has an arbitrary zero point. Indeed, shifting the reference point for s via the substitution $s = s' + s_0$ is equivalent to redefining the constant f: $f \to f' = f - \varepsilon s_0$. We will take the surface of the accretion disk as the reference point for s: $\tilde{z}(s = 0) = 0$.

The resulting equations can be written as a system of six ordinary first-order differential equations. We introduce the six-dimensional vector ξ_i with the components

$$\xi_1 = \tilde{r}_s, \quad \xi_2 = \tilde{z}_s, \quad \xi_3 = \tilde{\eta}_s, \tag{60}$$

$$\xi_4 = \tilde{r}, \quad \xi_5 = \tilde{z}, \quad \xi_6 = \tilde{\eta}. \tag{61}$$

Then, for $i = 1, 2, 3$, the equations have the structure

$$\frac{d\xi_{1,2,3}}{ds} = \frac{N_{1,2,3}}{D}, \tag{62}$$

while for $i = 4, 5, 6$

$$\frac{d\xi_{4,5,6}}{ds} = \xi_{1,2,3}. \tag{63}$$

The denominator D is the determinant of a certain 3×3 matrix; it appears when the initial system of equations is solved relative to the highest derivatives. The functions N_i and D depend on $\xi_1, \ldots, \xi_5$; ξ_6 does not appear in the equations due to the azimuthal symmetry of the problem. In addition, the right-hand sides of (62) depend on s.

For fixed ψ and φ, the curve $\mathbf{R} = \psi^{2/3}\tilde{\mathbf{R}}(s, \varphi)$ specifies a magnetic-field line in parametric form. In

---

[1] The equations are cumbersome, and we do not present them here. The analytical manipulations were performed using the Maple package.

the case of self-similar solutions, it will suffice to find one field line; others can be derived using the similarity transformation (multiplication by $\psi^2/3$) and rotation about the z axis by some angle φ. The field line corresponding to the frozen-in coordinate $\psi = 1$ crosses the disk at the distance $r(0) = 1$. The value $\tilde\eta(0)$ can be arbitrary, for example, zero. The solution of the system of differential equations should then satisfy the initial conditions

$$\tilde r(0) = 1, \quad \tilde z(0) = 0, \quad \tilde\eta(0) = 0. \tag{64}$$

The frozen-in coordinate ψ and cylindrical coordinate r are related by the simple expression

$$r = \psi^{2/3}. \tag{65}$$

To obtain an unambiguous solution of the system of equations, we must have another three conditions.

It follows from (54) that the plasma density at the disk surface is $\rho = 3/(2\tilde z_s^{(0)})$. Therefore, the components of the magnetic field in the disk are

$$B_r^{(0)} = \frac{3}{2}\frac{\tilde r_s^{(0)}}{\tilde z_s^{(0)}}\psi^{-1/3}, \quad B_z^{(0)} = \frac{3}{2}\psi^{-1/3}, \tag{66}$$

$$B_\phi^{(0)} = \frac{3}{2}\frac{\tilde\eta_s^{(0)}}{\tilde z_s^{(0)}}\psi^{-1/3},$$

where the superscript 0 denotes values for $z = +0$.

Let us write the boundary conditions (44)–(46) for the self-similar solution. Substituting (65) into (31) and applying the initial conditions, we obtain the velocity of the accretion disk:

$$u_r = \frac{2}{3}\varepsilon\psi^{-1/3}, \quad u_\phi = \tilde\Omega\psi^{-1/3}. \tag{67}$$

The velocity component $u_r < 0$, so that ε should also be negative. Together with the plasma velocity, u_r and u_ϕ depend on r as $r^{-1/2}$.

It follows from (44) or (45) that the mass surface density of the disk is a linear function of r:

$$\sigma = \mu r, \tag{68}$$

where $\mu = \text{const}$. In (46), j_z is expressed using (54) and (56) as follows:

$$j_z = \frac{3}{2}f\psi^{-1/3}. \tag{69}$$

We derive from the continuity equation (46)

$$\mu = 3f/|\varepsilon|. \tag{70}$$

Substituting (66), (67), and (68) into the boundary conditions (44) and (45), we obtain

$$\tilde\Omega^2 = 1 - \frac{2}{9}\varepsilon^2 + \frac{3\varepsilon}{2f}\frac{\tilde r_s^{(0)}}{\tilde z_s^{(0)}}, \tag{71}$$

$$\tilde\eta_s^{(0)} = -\frac{2}{9}f\tilde\Omega\tilde z_s^{(0)}. \tag{72}$$

This last expression can be used to calculate the azimuthal component of the magnetic field at the equator:

$$B_\phi^{(0)} = -\frac{1}{3}f\tilde\Omega\psi^{-1/3}. \tag{73}$$

We can see from these conditions that the rotational velocity of a field line is lower than the Keplerian velocity. Only in the limit $\varepsilon \to 0$ do we obtain $\tilde\Omega = 1$, which corresponds to a Keplerian accretion disk. Note that the boundary conditions in the disk fully specify the azimuthal magnetic field above the disk. Since ε does not appear in (72) and (73), these equalities should be satisfied even in the limit $\varepsilon \to 0$.

The conditions (44) and (46) make it possible to relate in a self-consistent way all the parameters of the accretion disk and of the outflowing plasma, and, in particular, to take into account the decrease in the surface density of the disk due to the expansion of the plasma. The equality (45) leads to the additional condition (72), which must be satisfied by the solution.

6. SOLUTION IN THE VICINITY OF THE DISK

Let us consider an important limiting case, when the radial velocity of the disk is small compared to the Keplerian velocity, $|\varepsilon| \ll 1$. Let us also assume that the velocity of the plasma outflow from the disk is fairly low. Then, the following equations are valid in the vicinity of the disk:

$$\frac{d\tilde r_s}{ds} = \tilde r_s G, \quad \frac{d\tilde z_s}{ds} = \tilde z_s G, \quad \frac{d\tilde\eta_s}{ds} = \tilde\eta_s G. \tag{74}$$

Here,

$$G = \frac{1}{6f^3}\frac{6f\tilde z_s(3\tilde r_s\delta\tilde r - \tilde z_s\tilde z) + 9\varepsilon\tilde r_s^2 - 2\varepsilon f\tilde z_s\tilde\eta_s}{\tilde z_s(\tilde r_s^2 + \tilde z_s^2 + \tilde\eta_s^2)}, \tag{75}$$

where $\delta\tilde r = \tilde r - 1$. To derive (74), $\tilde\Omega$ in the form (71) is substituted into the exact equations (62) and expanded in small parameters. Equations (74) are valid when $\delta\tilde r$ and $\tilde z$ are small compared to unity and

$$f^2\tilde r_s \ll 1, \quad f^2\tilde z_s \ll 1, \quad f^2\tilde\eta_s \ll 1. \tag{76}$$

These conditions follow from the requirement that subsequent terms of the expansion be small. It follows from (74) that

$$\xi_s = \xi_s^{(0)}\exp\left(\int\limits_0^s G ds\right), \tag{77}$$

where ξ_s is any one of the three functions $\tilde r_s$, $\tilde z_s$, or $\tilde\eta_s$. After integrating (77) with the initial conditions

$\tilde{z}(0) = 0$, $\delta\tilde{r}(0) = 0$, $\tilde{\eta}(0) = 0$, we find that a field line is straight in the vicinity of the disk:

$$\delta\tilde{r}(s) = k_r\tilde{z}(s), \quad \tilde{\eta}(s) = k_\eta\tilde{z}(s), \tag{78}$$

where k_r and k_η are constant. As we can see from (72), $k_\eta = -2f/9$ in the approximation considered. k_r soecifies the direction of the poloidal field in the vicinity of the disk.

Using equality (78), the second of Eqs. (74) can be presented in the form

$$\frac{d^2\tilde{z}}{ds^2} = w^2\tilde{z} - |\varepsilon|\zeta, \tag{79}$$

where

$$w^2 = \frac{1}{f^2}\frac{3k_r^2 - 1}{k_r^2 + k_\eta^2 + 1}, \tag{80}$$

$$\zeta = \frac{1}{6f^3}\frac{9k_r^2 - 2fk_\eta}{k_r^2 + k_\eta^2 + 1}. \tag{81}$$

When $w^2 < 0$, (79) has an oscillating solution, which is physically unacceptable. The inequality $k_r^2 > 1/3$, obtained in [16], must be satisfied by field lines moving away from the disk. This implies that a magnetic-field line must be inclined to the disk at an angle smaller than 60°.

Unlike [16], we also obtained another constraint. When $w^2 > 0$, the solution (79) has the form

$$\tilde{z} = \frac{|\varepsilon|\zeta}{w^2}(1 - \cosh(ws)) + \frac{\tilde{z}_s^{(0)}}{w}\sinh(ws), \tag{82}$$

where $\tilde{z}_s^{(0)}$ is an arbitrary constant equal to the derivative $\tilde{z}_s$ in the disk. It follows from (82) that solutions cannot exist for arbitrarily small $\tilde{z}_s^{(0)}$. The field line will move away from the disk only when

$$\tilde{z}_s^{(0)} > \tilde{z}_{smin} = \frac{|\varepsilon|\zeta}{w}. \tag{83}$$

This is equivalent to a constraint on the z component of the plasma outflow velocity:

$$v_z\big|_{z=0} > v_{zmin} = f\frac{|\varepsilon|\zeta}{w}. \tag{84}$$

The nature of this constraint, which appears when $\varepsilon \neq 0$, is easy to understand. In this case, each particle of the plasma moves along its orbit with a velocity that is smaller than the Keplerian velocity. In the z direction, the particle is in the effective potential

$$U(\tilde{z}) = -\frac{1}{2}w^2\tilde{z}^2 + |\varepsilon|\zeta\tilde{z}.$$

In order to be ejected from the disk, the particle must overcome the potential barrier corresponding to the difference between the gravitational and centrifugal forces.

The plasma density in the vicinity of the disk is

$$\rho = \frac{3}{2\tilde{z}_s} = \frac{3}{2\left(z_s^{(0)}\cosh(ws) - \dfrac{|\varepsilon|\zeta}{w}\sinh(ws)\right)}. \tag{85}$$

In the disk itself, it is finite:

$$\rho\big|_{z=0} = \frac{3}{2\tilde{z}_s^{(0)}} < \frac{3w}{2|\varepsilon|\zeta}. \tag{86}$$

In the limit $\varepsilon \to 0$, $\tilde{z}_s^{(0)} \to 0$, the density in the disk increases without bound, as in [16].

7. PASSAGE THROUGH CRITICAL SURFACES

The general solutions of the equations describing stationary MHD flows are singular on certain surfaces, usually called critical surfaces. These have already been found for one-dimensional flows, as was proposed by Parker for the solar wind [33]. In one-dimensional solutions, they are manifest as critical points, while, in the three-dimensional case, a set of critical points forms a surface. In the one-dimensional solutions, the critical points coincide with sonic points, where the local velocity of weak perturbations becomes comparable to the flow velocity, while the type of equation changes from elliptical to hyperbolic and back. Even the earliest solutions obtained for axisymmetrical flows indicated that this is not the case in general [16]. The nature of critical surfaces and their role in the formation of stationary MHD solutions was investigated in [34, 35]. It turns out that critical surfaces divide regions of solutions with different cause–effect relations, and these surfaces do not, in general, coincide with the surfaces where the form of the equation changes.

Critical surfaces as applied to self-similar solutions were studied in [36]. The parameters of critical surfaces for self-similar flows with $E_\varphi \neq 0$ were studied in detail in [27]. In these solutions, the locations of the critical surfaces are specified by the zeros of the denominator D in (62). In the general case (with a nonzero wind temperature), the denominator vanishes at the slow magnetoacoustic separatrice surface, Alfven surface, and fast magnetoacoustic surface. Regularizing the solution on these surfaces makes it possible to determine the density of the plasma flux from the disk and the two tangential components of the magnetic field at the disk surface (the normal component is specified by the conditions adopted for the problem). In the case of a cool plasma, the slow magnetoacoustic velocity is zero, and one of the regularization conditions disappears. Therefore, the plasma flux from the disk surface must be specified for a cool plasma. This means that the ratio f of the density of the plasma flux from the disk and the z component of the magnetic field is also specified. A similar approach was used in [16] (where k is analogous to f).

Thus, in self-similar flows of cool plasma, regularizing the solution on the Alfven and fast magnetoacoustic critical surfaces specifies both the azimuthal and tangential components of the magnetic field based on the flow. One characteristic feature of self-similar solutions is that, as a rule, it is not possible to regularize them on the fast magnetoacoustic critical surface, only at the Alfven surface. Recently, Vlahakis *et al.* [37] were able to regularize a self-similar solution on the fast magnetoacoustic surface, but, even in this case, the solution disappears very rapidly behind this surface. Therefore, as in [16], we will not try to regularize the solution on the fast magnetoacoustic surface. As a result, one of the components of the magnetic field (the azimuthal component in [16]) remains free and must be specified *a priori*. The other component is determined from the solution and the regularization condition at the Alfven surface.

The denominator D in (62) contains the factor

$$e_1 \equiv \rho - (f - \varepsilon s)^2. \tag{87}$$

In the disk, $e_1 > 0$ when $f^2 \tilde{z}_s^{(0)} < 3/2$. Since the density decreases along a field line, e_1 vanishes at some point on the Alfven surface. To keep the solution finite, the functions N_i must also vanish at this point. After substituting $\rho = (f - \varepsilon s)^2$ into N_i, all the functions N_i possess the common factor

$$\begin{aligned} e_2 \equiv & \left(-108\tilde{\Omega} b + \left(12\tilde{r}^2\varepsilon\tilde{\Omega}\tilde{z} - 27\tilde{\eta}_s\right) b^2\right) \tilde{r}_s^2 \\ & + \Bigg((8\tilde{r}^3\tilde{z}\varepsilon^2\tilde{\eta}_s + 8\tilde{r}\tilde{z}^3\varepsilon^2\tilde{\eta}_s + 12\tilde{r}\tilde{z}^2\varepsilon\tilde{\Omega}\tilde{z}_s \\ & - 12\tilde{r}^3\varepsilon\tilde{\Omega}\tilde{z}_s + 36\tilde{r}^3\tilde{z}\tilde{\Omega}^2\tilde{\eta}_s)b^2 - 108\frac{\tilde{\Omega} b\tilde{z}\tilde{z}_s}{\tilde{r}}\Bigg)\tilde{r}_s \\ & + \big(-36\tilde{r}^4\tilde{z}_s\tilde{\Omega}^2\tilde{\eta}_s - 8\tilde{r}^4\tilde{z}_s\tilde{\varepsilon}^2\tilde{\eta}_s - 12\tilde{r}^2\tilde{z}\varepsilon\tilde{\Omega}\tilde{z}_s^2 \\ & - 8\tilde{r}^2\tilde{z}_s\varepsilon^2\tilde{\eta}_s\tilde{z}^2 - 27\tilde{r}^2\tilde{\eta}_s^3 - 27\tilde{\eta}_s\tilde{z}_s^2\big)b^2 \\ & + 54\frac{\tilde{\eta}_s}{\sqrt{\tilde{r}^2 + \tilde{z}^2}} - 108 b\tilde{\Omega}\tilde{r}^2\tilde{\eta}_s^2, \end{aligned} \tag{88}$$

where $b = f - \varepsilon s$. Therefore, the condition that the denominator and three functions N_i in (62) simultaneously vanish results in the system of two equations

$$e_1 = 0, \quad e_2 = 0, \tag{89}$$

which can be solved analytically for $\tilde{r}_s$ and $\tilde{z}_s$. Consequently, the derivatives $(\tilde{r}_s, \tilde{z}_s)$ can be taken to be specified at the critical point $(\tilde{r}, \tilde{z})$.

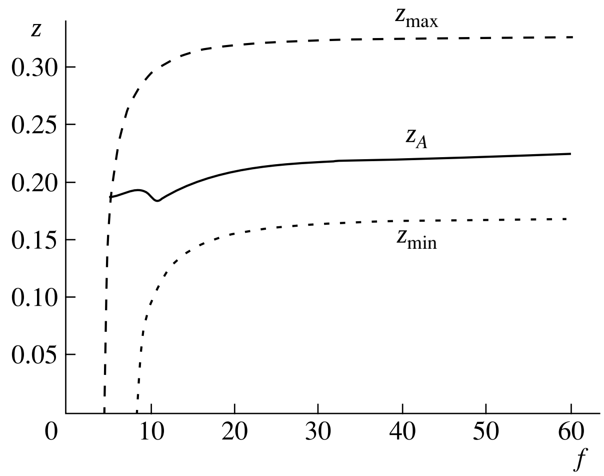


Fig. 4. Values z_A of $\tilde{z}(f)$ for which a field line emerging from the Alfven point crosses the equator when $\tilde{r} = 1$. z_{max} and z_{min}, which appear in (95), are also indicated.

8. SOLUTION IN THE LIMIT $\varepsilon \to 0$

It is of interest to consider the solution in the limit of very slow accretion, when $\varepsilon \to 0$. This case is of interest because the disk becomes Keplerian in this limit ($\tilde{\Omega} = 1$), so that our results can easily be compared to those of [16], and the physics of accretion made possible by a wind outflow becomes particularly simple.

In the case of small ε, εs can be neglected compared to f. The parameter s then disappears from the right-hand sides of (62), and the simplified equations have two integrals of motion: the momentum L and the energy W. For the self-similar solution,

$$L = \tilde{r}^2 \left(f\tilde{\Omega} - \tilde{\eta}_s(\rho - f^2)\right), \tag{90}$$

$$W = \frac{1}{2} f^2 \left(\tilde{r}_s^2 + \tilde{r}^2\tilde{\eta}_s^2 + \tilde{z}_s^2\right) - \frac{1}{2}\tilde{\Omega}^2\tilde{r}^2 - \frac{1}{\sqrt{\tilde{r}^2 + \tilde{z}^2}}, \tag{91}$$

where we can assume $\tilde{\Omega} = 1$ on the right-hand sides of the equality.

Substituting the values at the disk surface into (90) and using (64) and (72), we find

$$L = \frac{4}{3} f \left(1 - \frac{1}{6} f^2 \tilde{z}_s^{(0)}\right). \tag{92}$$

Since $\tilde{z}_s^{(0)} \geq 0$, $L \leq L_{max} = \frac{4}{3} f$. This new important constraint on the angular-momentum flux carried away from the disk follows from the boundary conditions at the disk. It means that a relatively small flux of angular momentum must be carried away by the wind to enable accretion. In addition, the boundary conditions at the disk specify the azimuthal component of

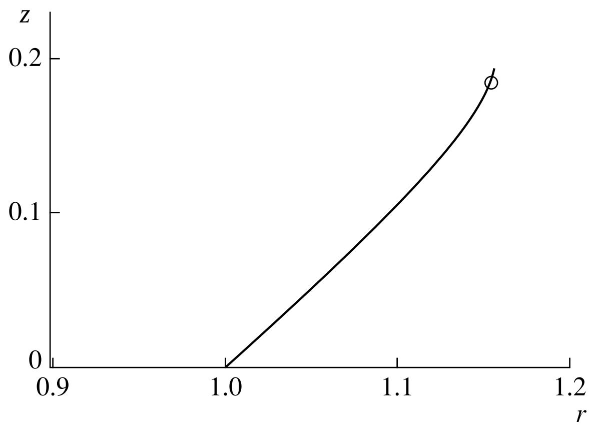


Fig. 5. A line of the poloidal field for $f = 6$. The circle marks the position of the Alfven singularity.

the magnetic field, and the problem is fully specified even when the solution is regularized at one Alfven point.

As we can see from (84), in the limit $\varepsilon \to 0$, the initial velocity of the plasma outflow can be arbitrarily small. In this case, the energy tends to the minimum $W_{min} = -3/2$ [see (91)]. Assuming that the conditions (76) are satisfied at the disk surface, we obtain $L = 4/3f$, $W = -3/2$. Therefore, at the Alfven point,

$$\tilde{r}^2 = \frac{L}{f} = \frac{4}{3}. \qquad (93)$$

The energy integral and Eqs. (89) can be used to express analytically the derivatives $\tilde{r}_s$, $\tilde{z}_s$, and $\tilde{\eta}_s$ at the Alfven point in terms of $\tilde{z}$. A straightforward but fairly cumbersome analysis shows that the solutions (89) exist for $L = 4/3f$, $W = -3/2$ if

$$f > \frac{9}{16}\sqrt{6(5 + 3\sqrt{3})} \approx 4.4. \qquad (94)$$

$\tilde{z}$ must be confined within the interval

$$z_{min} \le \tilde{z} \le z_{max}, \qquad (95)$$

where

$$z_{max} = \left[\frac{216f\sqrt{16f^2 - 45} - 736f^2 - 1215}{1200f^2}\right]^{1/2}, \qquad (96)$$

$z_{min} = 0$ when $f < 8.27$, and

$$z_{min} = \frac{2}{\sqrt{3}}\frac{\sqrt{2351f^4 - 159894f^2 - 59049}}{329f^2 + 243} \qquad (97)$$

when $f > 8.27$. It follows from (96) that $\tilde{z} < 0.327$ at the Alfven point; this means that the Alfven point cannot be far from the disk surface.

We solved (62) and (63) in the limiting case $\varepsilon \to 0$ numerically using standard techniques. We started from the Alfven point, with f as the input parameter. As the initial condition, we took the Alfven point with $\tilde{r} = 2/\sqrt{3}$ and $\tilde{z}$ from the interval (95); the derivatives were selected using the described procedure. At the Alfven point, the right-hand sides of (62) were determined using l'Hopitale's rule. The initial value for $\tilde{z}$ was selected so that the field line crossed the disk at the point $\tilde{r} = 1$. We adopted a value slightly larger than $-3/2$ for W. The reason for this is that, when $W = -3/2$, the plasma velocity at the disk surface vanishes and the plasma density at the disk displays a nonintegrable singularity: $\rho = 3/(2w\tilde{z})$. Therefore, we chose W so that $\tilde{z}_s^{(0)} \sim 10^{-3}$ in the disk, or, in other words, so that the velocity of the plasma outflow relative to the disk was of the order of $f \times 10^{-3}$ of the Keplerian velocity. The requirement that $\tilde{z}_s^{(0)}$ be nonzero also results from the fact that the inequality (84) should be rigorous.

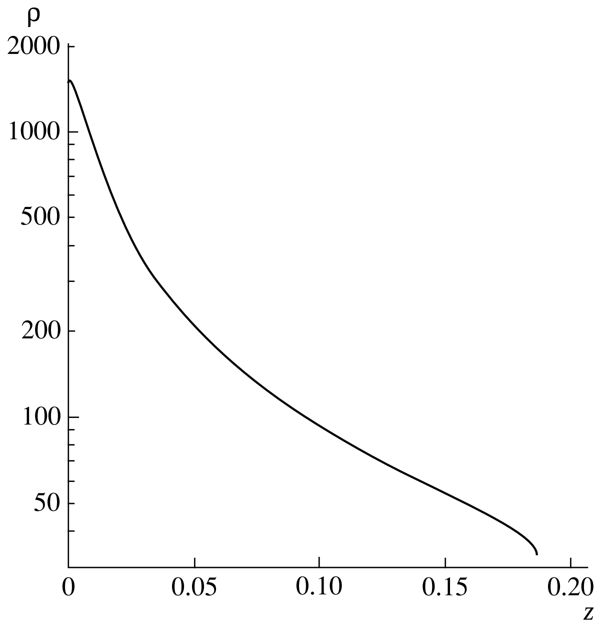


Fig. 6. z dependence of the plasma density on the field line shown in Fig. 5.

Figure 4 presents the f dependence of $\tilde{z}$ at the Alfven point together with z_{min} and z_{max}. The minimum f for which the solution exists is $f_{min} = 5.28$. $z_{max} = z_A$ when $f = f_{min}$.

Figure 5 presents a line of the poloidal magnetic field for $f = 6$, which does not differ much from f_{min}. $k_r = 0.96$, so that the line is inclined to the equator at an angle that is close to $\pi/4$; at the disk surface, $\tilde{z}_s = 10^{-3}$. There is no significant acceleration of the plasma; at the Alfven point, $\tilde{z}_s = 0.033$ and $\tilde{r}_s = 0.010$, so that the components of the plasma velocity v_z and v_r are small compared to the Keplerian velocity. Figure 6 presents the plasma density on a field line for various heights. In the case of motion from the disk toward the Alfven point, s varies from 0 to $s_{max} \approx 30$. Therefore, εs can be neglected compared to f when

$$|\varepsilon| \ll f/s_{max}. \tag{98}$$

This condition, as well as the inequality (76), ensures the applicability of the approximation used in the case considered. The analytical solution (82) coincides with the numerical solution along nearly the entire field line, with the exception of the immediate vicinity of the Alfven point.

9. DISCUSSION

The conclusion that the plasma outflow from an accretion disk can exert a substantial effect on the dynamics of the accretion disk is not new [19]. The originality of our work is our consideration of the case when plasma falls onto the gravitating center along with the frozen-in magnetic field, rather than seeping across magnetic-field lines (see, for example, [21] and references therein). Our results indicate that dissipationless disk accretion of the type considered is, indeed, possible, as is demonstrated by the derived self-consistent solution.

The mechanism for carrying away angular momentum that we have considered should be taken into account in studies of accretion in sources displaying violent ejections of matter. This is particularly important for our understanding of the processes occurring in a number of peculiar sources, prominently represented by SS433 [2]. Objects in which disk accretion occurs primarily due to the carrying away of angular momentum by an outflowing wind can have bolometric luminosities that are appreciably lower than the kinetic-energy flux in the outflowing wind. This is what leads to the most important difficulties in explaining the plasma-ejection mechanism in both SS433 and young stellar objects.

The derived self-consistent solution possesses a number of interesting properties. The boundary conditions (71), (72) fully specify the azimuthal magnetic field in the disk. In this case, with fixed f and L, the toroidal and tangential components of the magnetic field are determined unambiguously, unlike in the study of Blandford and Payne [16]. The flow turns out to be fully specified. This begs the question of whether there exists the fundamental possibility of passing through the fast magnetoacoustic critical point, since, at first glance, no free parameters remain. In fact, there is still one free parameter: ε, which specifies the radial velocity of the plasma in the disk. This parameter can be used to regularize the solution at the fast magnetoacoustic critical point. If this is possible, the entire solution will prove to be unambiguously specified.

An unexpected feature of the obtained solutions is that, in the limiting case of dissipationless accretion, 100% of the infalling matter is ejected back out from the disk. In this respect, our results are similar to those of Lery *et al.* [38], who also considered dissipationless accretion, but without the formation of a disk. Both of these results indicate that, in any case, dissipation is probably unavoidable, to provide for the infall of some fraction of the matter onto the gravitating center. On the other hand, this feature of the flow may well be due to its self-similarity. Additional studies are required before we will be able to draw final conclusions.

A natural extension of our study is the consideration of accretion due simultaneously to matter outflow from the disk and dissipation. Even at this stage, it should be possible to determine the relationship between the disk luminosity and the flux of outflowing matter, which can be compared to observations.

ACKNOWLEDGMENTS

The authors thank G.S. Bisnovatyĭ-Kogan, V.M. Chechetkin, and G. Pelletierre for useful discussions. We also acknowledge Johnatan Ferreira, whose numerous remarks on the connection between disk accretion and jets initiated this study. This work was supported by the Russian Foundation for Basic Research (project no. 03-02-17098), joint INTAS-ESA grant 99-120 and the program of the Presidium of RAS "Non-stationary phenomena in astronomy."

REFERENCES

1. C. M. Urry and P. Padovani, Publ. Astron. Soc. Pac. **107**, 803 (1995).
2. A. M. Cherepashchuk, Itogi Nauki Tekh. **38**, 60 (1988).
3. M. Livio, Phys. Rep. **311**, 225 (1999).
4. I. F. Mirabel and L. F. Rodrigues, Nature **392**, 673 (1998).
5. M. C. Weisskopf, J. J. Hester, F. A. Tenant, *et al.*, Astrophys. J. **536**, L81 (2000).
6. S. V. Bogovalov and D. V. Khangoulian, Mon. Not. R. Astron. Soc. **336**, L53 (2002).
7. N. I. Shakura, Astron. Zh. **49**, 921 (1972) [Sov. Astron. **16**, 756 (1972)].
8. N. I. Shakura and R. A. Sunyaev, Astron. Astrophys. **24**, 337 (1973).
9. G. S. Bisnovatyi-Kogan and S. I. Blinnikov, Astron. Astrophys. **59**, 111 (1977).
10. T. Okuda and M. Fujita, Publ. Astron. Soc. Jpn. **52**, L5 (2000).
11. G. I. Barenblatt, *Similarity, Self-Similarity, and Intermediate Asymptotics* (Gidrometeoizdat, Leningrad, 1978, 1982; Consultants Bureau, New York, 1979).

12. W. Brinkman and N. Kawai, Astron. Astrophys. **363**, 640 (2000).
13. H. C. Spruit, *The Neutron Star—Black Hole Connection*, Ed. by C. Kouveliotou, J. Ventura, and Ed van den Heuvel (Kluwer Acad., 2001), p. 141.
14. I. Yi, *Astrophysical Disks*, Ed. by J. A. Sellwood and J. Goodman; Astron. Soc. Pac. Conf. Ser. **160**, 279 (1999).
15. G. S. Bisnovatyi-Kogan, Discrete Dyn. Nature Soc. **6**, 247 (2001).
16. R. D. Blandford and D. G. Payne, Mon. Not. R. Astron. Soc. **199**, 883 (1982).
17. R. E. Pudritz and C. A. Norman, Astrophys. J. **274**, 677 (1983).
18. R. E. Pudritz and C. A. Norman, Astrophys. J. **301**, 571 (1986).
19. G. Pelleter and R. E. Pudritz, Astrophys. J. **394**, 117 (1992).
20. G. S. Bisnovatyĭ-Kogan and A. A. Ruzmaikin, Astrophys. Space Sci. **28**, 45 (1974).
21. J. Ferreira and G. Pelletier, Astron. Astrophys. **295**, 807 (1995).
22. F. Casse and J. Ferreira, Astron. Astrophys. **353**, 1115 (2000).
23. M. M. Romanova, G. V. Ustyugova, V. M. Chechetkin, and R. V. E. Lovelace, Astrophys. J. **500**, 703 (1998).
24. I. F. Mirabel, V. Dhawan, S. Chaty, *et al.*, Astron. Astrophys. **330**, L9 (1998).
25. P. Murdin, D. H. Clark, and P. G. Martin, Mon. Not. R. Astron. Soc. **193**, 135 (1980).
26. S. V. Bogovalov and K. Tsinganos, Astron. Astrophys. **356**, 989 (2000).
27. J. Contopoulos, Astrophys. J. **460**, 185 (1996).
28. N. Vlahakis and K. Tsinganos, Mon. Not. R. Astron. Soc. **307**, 279 (1999).
29. Z. Y. Li, Astrophys. J. **444**, 848 (1995).
30. V. S. Beskin, Usp. Fiz. Nauk **169**, 689 (1997).
31. M. I. Pudovkin and V. S. Semenov, Ann. Geophys. **33**, 429 (1977).
32. V. S. Semenov, in *Geomagnetic Investigations* (Sov. Radio, Moscow, 1979) Vol. 24, p. 32 [in Russian].
33. E. Parker, *Interplanetary Dynamical Processes* (Interscience, New York, 1963), p. 51.
34. S. V. Bogovalov, Mon. Not. R. Astron. Soc. **270**, 721 (1994).
35. S. V. Bogovalov, Astron. Astrophys. **323**, 634 (1997).
36. K. Tsinganos, C. Sauty, G. Surlantzis, *et al.*, Mon. Not. R. Astron. Soc. **283**, 811 (1996).
37. N. Vlahakis, K. Tsinganos, C. Sauty, and E. Trussoni, Astrophys. J. **545**, 758 (2000).
38. T. Lery, R. Henriksen, and J. D. Fiege, Astron. Astrophys. **350**, 254 (1999).

Translated by K. Maslennikov

Astronomy Reports, Vol. 49, No. 1, 2005, pp. 71–78. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 1, 2005, pp. 79–87.
Original Russian Text Copyright © 2005 by M. Eselevich, V. Eselevich.

The Structure of the Coronal-Streamer Belt

M. V. Eselevich and V. G. Eselevich

Institute of Solar–Terrestrial Physics, Siberian Division, Russian Academy of Sciences,
P.O. Box 4026, Irkutsk, 664033 Russia

Received May 25, 2004; in final form, July 15, 2004

Abstract—We analyze calibrated white-light coronal images from the LASCO-C2/SOHO experiment (processing level L1), focusing on quasistationary events without coronal mass ejections or their manifestations in the solar wind. The previous result that the streamer belt forms a set of rays of increased brightness is confirmed. The cross section of the streamer belt is frequently observed as two closely spaced rays differing in brightness. It is difficult to explain this in terms of ordinary bending of the belt. We suggest that the belt is normally a set of pairs of rays with enhanced brightness (or two close rows of rays). The distance between the rays in each pair is comparable to the ray size. The ray brightnesses in any pair can, in general, be different. The magnetic field has opposite directions in the rays forming a pair, so that the neutral line of the radial component of the solar magnetic field probably runs along the strip between the pairs of rays. © *2005 Pleiades Publishing Inc.*

1. INTRODUCTION

Observations of the white-light solar corona during eclipses, in which numerous Russian and foreign investigators have been deeply involved since the first half of the past century, revealed bright structures that came to be known as streamers. A helmet—one form of arch system—is located at the base of a streamer, with a radial ray stretching above the helmet [1, 2]. Individual photographs taken during fairly rare solar eclipses before the 1970s suggested that streamers were local features with similar sizes in latitude and longitude [2, 3]. However, continuous daily observations from October 11, 1971 to January 15, 1973 carried out on the OSO-7 spacecraft [4] showed that streamers form in a series, or a so-called streamer belt, with a neutral line of the radial component of the global solar magnetic field running along it [5, 6]. Gulyaev [7] suggested that the observational difference in the appearance of the white-light corona at solar-activity minima and maxima could be due to different orientations of the streamer belt relative to the plane of the sky, and likely reflects the presence of spatial plasma-density inhomogeneities along the belt.

An analysis of uncalibrated LASCO data (processing level L0.5) showed [8] that, in the absence of coronal mass ejections (CMEs), the streamer belt is inhomogeneous at heliocentric distances $R > 3R_\odot$ (where $R_\odot$ is the radius of the Sun), and is formed by a set of radial rays of enhanced brightness. The minimum angular size is nearly the same for all the rays, $d \approx 2.0^\circ$–3.0°, and the minimum distance between rays is ≈5°–10°. Here and below, we measure angular sizes in units of the solar-disk arc, i.e., the circumference of the Sun corresponds to 360°. At heliocentric distances of $R > (4$–$5)R_\odot$, the rays are directed radially to within $\approx \pm 1.5^\circ$. Both the quasistationary solar wind, whose parameters vary relatively slowly, on characteristic time scales of 10 days and more, and sporadic (nonstationary) plasma motions with characteristic lifetimes of several hours are present in the rays [9]. Such sporadic plasma flows include (a) matter inhomogeneities ("blobs") that move away from the Sun [10] and (b) recently discovered sporadic plasma streams directed either away from or toward the Sun [11, 12]. In some cases, the nonstationary solar wind is associated with the formation of the quasistationary solar wind, when an additional plasma stream with enhanced density directed away from the Sun fills a separate ray [13]. The quasistationary ray structure of the streamer belt (which can be revealed by averaging over intensity fluctuations of the white-light corona with periods shorter than a day) is omnipresent; it is sometimes disrupted by the action of CMEs, but recovers within one to two days [14]. The streamer belt encircles the entire Sun as a wavy surface ("skirt"). At an activity minimum, this surface is close to the solar equatorial plane and intersects it in two or four places. The neutral line of the radial component of the global solar magnetic field runs along the streamer belt [6]. Recent studies based on calibrated LASCO data (processing level L1) [13, 15] strongly suggest that the belt is made up of a set of pairs of rays (with the rays in a pair, in general, having different brightnesses) rather than a bundle of single rays. In this case, the neutral line probably runs along the belt between the rays of each pair. In other words, this structure would consist of two closely spaced (separated by a distance comparable to the size of a ray, d) rows of rays of enhanced brightness with opposite signs of the radial component of the global solar magnetic field. The question of whether this exotic structure of the streamer belt exists is of paramount importance both for the nature of the slow solar wind in the belt and from the standpoint of the possibility of collective processes in the rarefied coronal plasma, which could be responsible for the formation of such structures.

Our aim here is to apply extensive data and a refined analysis technique to demonstrate that the streamer belt is indeed a set of ray pairs rather than a set of single rays, with the rays in a pair, in general, differing in brightness.

2. SOURCE DATA

We used white-light images of the solar corona obtained with the SOHO LASCO-C2 experiment as the source data for our analysis, specifically, calibrated 1024 × 1024 LASCO-C2 images (L1 processing level) available from http://lasco-www.nrl.navy.mil/. These data are free of effects such as scattering, vignetting, etc., and the intensity values are presented in units of the mean solar brightness, P_{msb}. The technique described in [8] was used to investigate the structure of streamer-belt segments in the absence of CMEs, which occur in two simple and characteristic situations: (i) when the belt segment is normal to the plane of the sky (the events of April 27 through May 11, 1996) and (ii) when the belt segment is nearly in the plane of the sky (individual events in 2000). The essence of this technique is briefly described in the next section.

3. IDENTIFICATION OF RAYS IN THE STREAMER BELT. ANALYSIS TECHNIQUE

A narrow ray ($d \approx 2^\circ - 3^\circ$) in the streamer belt can easily be visually identified in coronal images within the field of view of the coronagraph if the ray is located in a meridionally stretched segment of the streamer belt [8]. The possibility of identifying and tracing individual rays in other segments of the streamer belt over relatively long periods depends on an important property of the projection of such rays onto the plane of the sky: the apparent latitude of the ray Λ in the plane of the sky (the angle between the observed projection of the ray onto the plane of the sky and the plane of the solar equator) varies as the Sun rotates and the ray moves out of the plane of the sky (i.e., as the longitude of the ray varies). This variation depends on the actual solar latitude of the ray, λ (the angle between the radial ray and the plane of the solar equator), and the heliographic latitude of the visible center of the solar disk (or the heliographic latitude of the Earth), $B_\circ$. As a result, the ray will describe a curve in a synoptic map (usually an arc with its bulge facing the equator), whose equation for a given λ (in terms of the latitude and Carrington longitude) was obtained by Hundhausen [16] for the particular case of $B_\circ = 0^\circ$ and by us [8] for $B_\circ \neq 0^\circ$. If $\lambda \neq 0^\circ$, this effect can be used to resolve rays that are aligned with a parallel, and therefore merge into a single ray as they deviate from the plane of the sky, since they have different latitudes Λ projected onto the plane of the sky.

A visual analysis of LASCO synoptic maps of the coronal-brightness distribution at a fixed heliocentric distance (http://lasco-www.nrl.navy.mil/) shows that rays located at the tip of the bend in the streamer belt that is furthest from the solar equator in the northern or southern directions are most pronounced [15]. It is such rays that are most suitable for the analysis, in particular, for determining the plasma density in the ray [17]. In general, the brightness of an individual ray can be time dependent due to the solar rotation and nonstationary processes in the plasma flowing in the ray. Because of the solar rotation, the brightness of an individual ray decreases as the ray deviates from the plane of the sky, or increases as the ray approaches this plane [9, 15]. The characteristic time scale for such variations is about two to three days. Relatively rapid, nonstationary intensity variations with characteristic time scales of several hours or less (for example, so-called blobs and CMEs) can be identified against the background of these variations without considerable difficulties. Therefore, comparisons of the time variations in the ray brightness observed under certain conditions and the corresponding calculated dependences enable us to reliably identify quasistationary rays and investigate their properties [8, 9].

Our technique for extracting quantitative information from white-light coronal images consists of the following steps. For each white-light LACSO-C2 image of the corona, we constructed the coronal-brightness distribution $P(\Lambda)$ or $P(R)$ for a fixed apparent latitude Λ or a fixed heliocentric distance R separately for the eastern and western limbs, using P_{msb} units for the P values.

We characterized the brightness of an individual ray using the ray-brightness amplitude P_R, which we defined in [17]. The concept of ray brightness P_R can be introduced because a ray is characterized in the brightness profile $P(\Lambda)$ by the slope of the two lines forming the ray, which can be considered to be straight from the tip P_M—the intensity maximum in the ray—to the inflection points A and B. This makes it possible to define the ray-brightness as $\Pi(\Lambda) = P(\Lambda) - P_S(\Lambda)$ and, accordingly, the amplitude of the ray brightness at a given time t and a given distance R as $P_R = \max\Pi(\Lambda) = (P_M - P_S)$. The angular size $d \approx 2^\circ{-}3^\circ$ is defined as the angular width of the function $\Pi(\Lambda)$ at a height of $1/3P_R$ [17]. The quantity $P_S(\Lambda)$ is the background intensity—the result of averaging $P(\Lambda)$ over angles $\Lambda \leq 10^\circ$ at a given t and R. We also constructed the two-dimensional ray-brightness functions $\Pi(\Lambda, R) = P(\Lambda, R) - P_S(\Lambda, R)$ for a chosen range of angles Λ (which do not normally exceed 60°) $\Delta\Lambda$ for sequential times t using histogram equalization. The step in R was $0.007R_\odot$. This procedure is needed to enhance the brightness of faint ray segments located far from the Sun. This obviously leads to a loss in quantitative information about the brightness.

We studied the variations of the distributions $P(\Lambda)$, $P(R)$, and $P(\Lambda, R)$ with time t. To identify rays and study the dynamics of the processes occurring in them, we also used another parameter—the time-differenced ray brightness $P(t - t_\circ)$. We employed the following procedure to determine the time-differenced brightness functions $\Pi(\Lambda, t - t_\circ)$ and $\Pi(R, t - t_\circ)$. The intensity profiles $P(\Lambda, t_\circ)$ and $P(R, t_\circ)$ for some initial time $t_\circ < t$ were subtracted from the profiles $P(\Lambda, t)$ and $P(R, t)$ to obtain the time-differenced brightness functions $\Pi(\Lambda, t - t_\circ) = P(\Lambda, t) - P(\Lambda, t_\circ)$ and $\Pi(R, t - t_\circ) = P(R, t) - P(R, t_\circ)$, respectively. These were used to determine the time-differenced amplitude functions of the ray brightness for individual rays in accordance with the relationships $P_R(\Lambda, t - t_\circ) = P_M(\Lambda, t) - P(\Lambda, t_\circ)$ and $P_R(R, t - t_\circ) = P_M(R, t) - P(R, t_\circ)$, where P_M is the intensity maximum in the ray. If the ray-brightness function $\Pi(\Lambda, R)$ and the time-difference brightness function $\Pi(t - t_\circ, \Lambda, R)$ are defined in this way, the contribution of the F corona is eliminated. This enables us to determine the plasma density N_R averaged over the angular diameter of the ray d at a given distance R without using data on the polarization intensity, which are mainly intended for eliminating the effects of the F corona.

4. GENERAL APPEARANCE OF THE STREAMER-BELT STRUCTURE

To determine the structure of the streamer belt, it is sufficient to investigate its structure along two mutually orthogonal directions: along and normal to the surface of the belt. In limb observations of the white-light corona, the former situation takes place when the surface of the belt is located in (or near) the plane of the sky and the latter when this surface is normal to the plane of the sky. Figure 1 shows examples of the two-dimensional brightness functions $\Pi(\Lambda, R) = P(\Lambda, R) - P_S(\Lambda, R)$ for the latitude range $\Delta\Lambda = 50^\circ{-}55^\circ$ for the cases when the streamer belt is nearly in the plane of the sky (top) and normal to the plane of the sky (bottom). We can see in the upper plot that the structure of the belt along its surface is formed by a set of enhanced-brightness rays. The following features can be noted:

(a) the minimum distance between rays is 5°–10°, while the maximum distance can be more than 25°;

(b) the rays are essentially radial at $R > (4{-}5)R_\odot$ and are not radial at $2 < R_\odot < (4{-}5)R_\odot$, deviating toward the poles with heliocentric distance (recall that the rays in both hemispheres at latitudes of less than 60° deviate toward the equator at $R < (4{-}5)R_\odot$) in the plane normal to the belt) [12];

(c) the brightness decline with R can be markedly different in different rays.

An examination of a time sequence of the ray patterns in the plane of the sky shows that the behavior of these patterns is quite dynamic. As we have shown [17], the observed ray dynamics are associated with the motion of additional streams of solar-wind plasma with steep leading fronts within the rays, away from the Sun.

The cross section of the streamer belt that is observed at the limb when the streamer belt is normal to the plane of the sky (Fig. 1, bottom) generally appears as two rays of enhanced brightness, "Ray+" and "Ray−," which are essentially radial at $R > (4{-}5)R_\odot$ and approach the Sun at $R < (4{-}5)R_\odot$, passing around the helmet on either side. The rays of enhanced brightness are magnetic-flux tubes in which plasma with enhanced density moves away from the Sun, and the helmet is a system of arcades (loops) of magnetic-field lines filled with plasma moving along them. The radial component of the magnetic field is oppositely directed on either side of the helmet. The apex of the helmet is located between Ray+ and Ray− (Fig. 1, bottom), and corresponds to the position of the neutral line separating regions of opposite polarity of the global solar magnetic field. In the case under consideration, the polarity is positive (field pointing away from the Sun) north of the neutral line (above this line in the figure) and negative south of the neutral line. The rays are denoted Ray+ and Ray− in accordance with the expected sign of the magnetic field. We first drew this tentative conclusion in [18]. However, as we noted then, the cross section of the streamer belt is frequently observed as a single ray, while the helmet may be completely invisible. This is due to the dynamics of the pattern; i.e., the brightnesses in individual rays and in the helmet vary independently and with time.

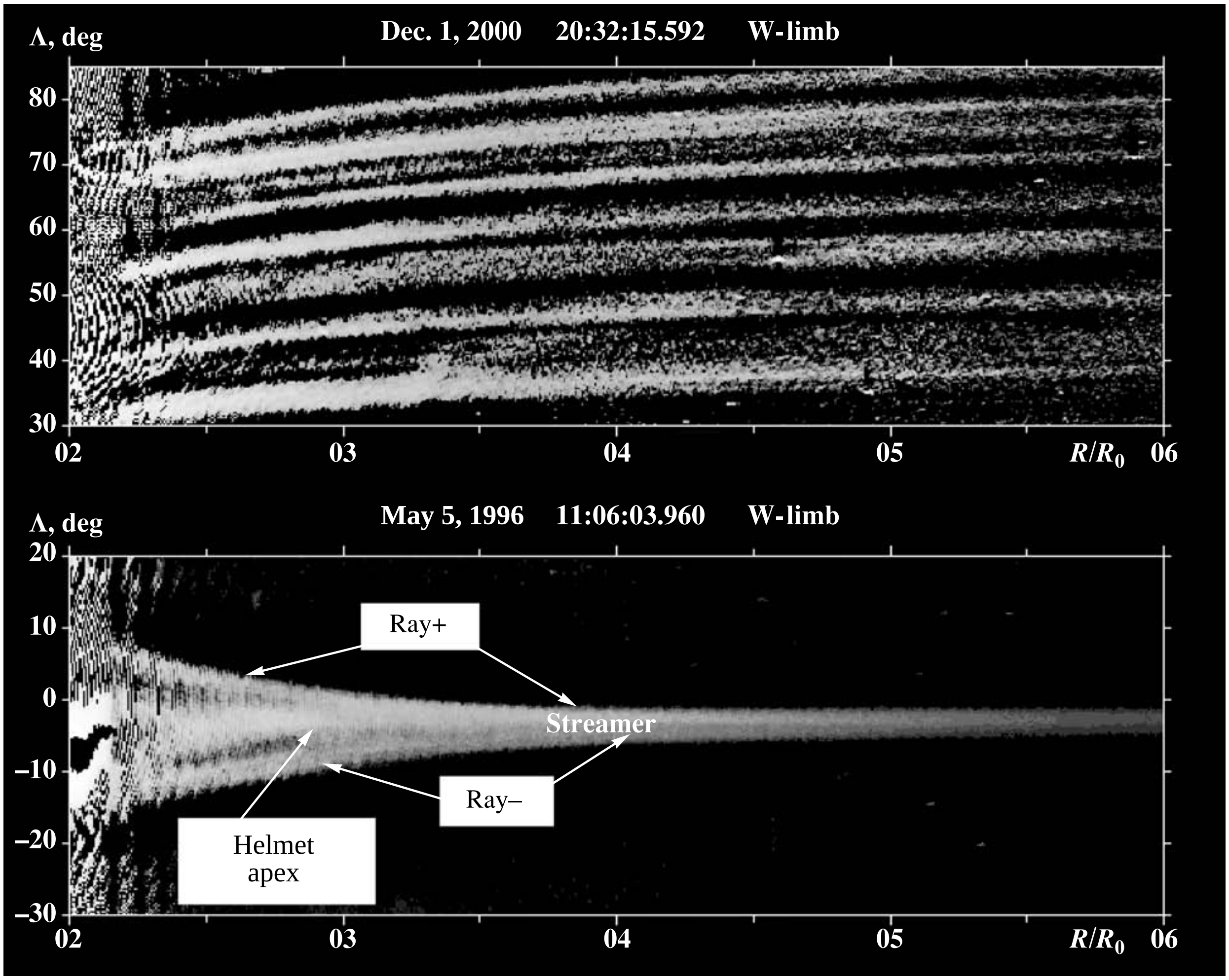


Fig. 1. Two-dimensional representations of the ray-brightness function $\Pi(\Lambda, R) = P(\Lambda, R) - P_S(\Lambda, R)$ for the two most simple and characteristic situations according to the LASCO-C2 data. In the top panel, a segment of the streamer belt is situated nearly in the plane of the sky (December 1, 2000, 20:31 UT; W limb); in the bottom panel, a segment of the streamer belt is normal to the plane of the sky (May 5, 1996, 11:06 UT; W limb).

5. STRUCTURE OF THE CROSS SECTION OF THE STREAMER BELT

To investigate the structure of the cross section of the streamer belt in more detail, we consider a fairly extended segment of the streamer belt (about 15 days, or $\delta\Psi_L \approx 180°$) typical of the solar-activity minimum, normal to the plane of the sky, near the equator (a latitude of $\lambda \approx -2° \ldots -5°$), at the western limb during April 27–May 11, 1996, in the absence of the action of coronal transients (i.e., CMEs). Whenever we need to separate time variations in the brightness and latitude of the rays due to nonstationary processes from variations due to the solar rotation, we will compare the measured dependences of the ray-brightness amplitude $P_R(\Psi_L)$ on the apparent latitude of the ray position $\Lambda(\Psi_L)$ at a heliocentric distance of $R = 4R_\odot$ with the corresponding dependences calculated for the same case using formulas (3) and (4) from [17] with $\alpha \approx 5°$ and formula (5) from [8].

Figures 2 and 3 show the measured profiles at the western limb, $\Pi(\Lambda)$, in units of P_{msb} for successive times from April 27 to May 11, 1996 and for $R = 4R_\odot$. The dependence $\Pi(\Lambda)$ occupying the smallest angular size is that for April 27, 1996 (21:14 UT) in Fig. 2. This angular size is $d \approx 2°$ at a level of one-third of the amplitude; i.e., nearly equal to the characteristic minimum size of a ray of enhanced brightness in the streamer belt. We denote this distribution $\Pi_\circ(\Lambda)$ and take this form of the distribution to correspond to the brightness profile of an individual ray, in order to represent all other profiles $\Pi(\Lambda)$ as combinations of two functions $\Pi_\circ(\Lambda)$ with possibly different brightness amplitudes P_R. These functions are shown in Figs. 2 and 3 by the light dashed curves.

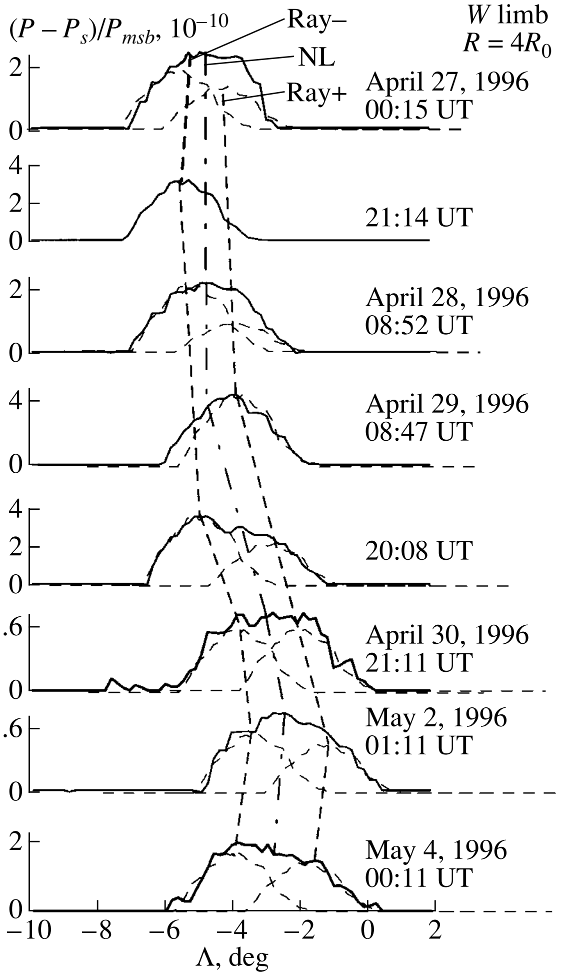


Fig. 2. One-dimensional brightness functions $\Pi(\Lambda)$ at $R = 4R_\odot$ for successive times from April 27 to May 4, 1996 (W limb; LASCO-C2 data; solid lines). The streamer belt is normal to the plane of the sky. The dot—dashed line in the vertical direction and shows the position of the apex of the helmet (the neutral line) the dashed lines the positions of the apices of the two rays.

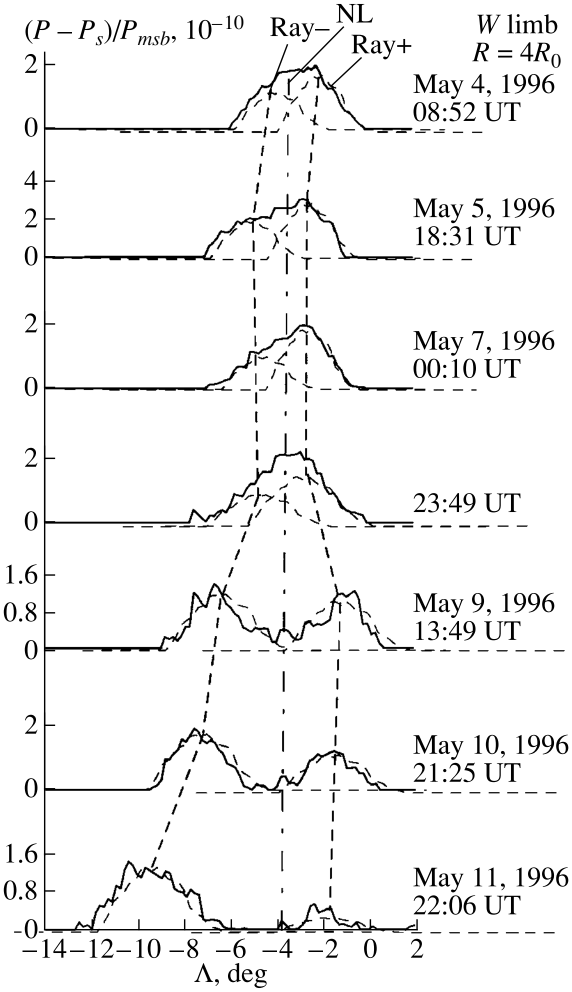
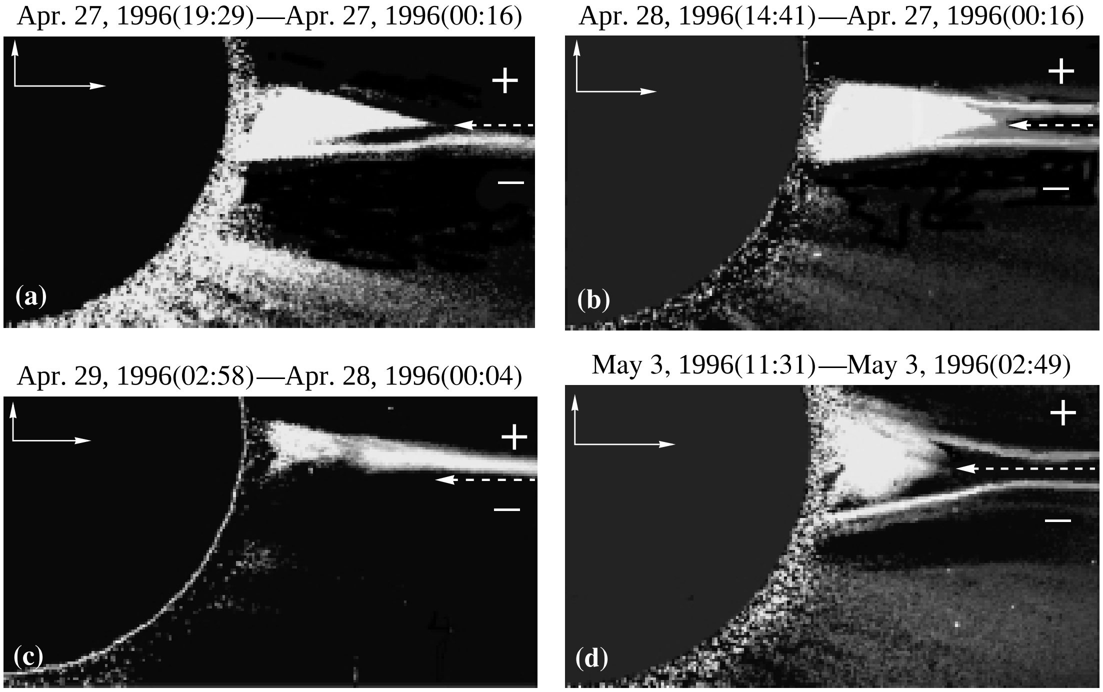


Fig. 3. Same as Fig. 2 for May 4–11, 1996.

We can see that most of the distributions $\Pi(\Lambda)$ correspond fairly well to a sum of two closely spaced functions $\Pi_\circ(\Lambda)$ with different brightness amplitudes, which we denote Ray+ and Ray–. The peaks of the distributions $\Pi_\circ(\Lambda)$ for these two rays at successive times are connected by the two bold dashed curves. The dot—dashed line between them indicates the approximate position of the apex of the helmet, or the neutral line (to within $\approx \pm(1°–2°)$).

The dynamics of the ray and helmet brightnesses are quite pronounced in the differenced images of the corona for the time intervals $t - t_\circ$ labeled on each plot.

The helmet is clearly visible in Figs. 4a, 4b, and 4d. The position of the helmet apex (neutral line) is shown by the white dashed arrow. The helmet is virtually invisible in Fig. 4c. However, this does not mean it is absent; its weakness is due to the fact that the brightness variations in the rays during the considered period, $t - t_\circ$, substantially exceeded the brightness variations in the helmet. The absence of one ray in Figs. 4a and 4c (April 27 and 29, 1996) is also due to the strong difference in the ray-brightness variations during the intervals $t - t_\circ$ labeled in these plots. Thus, Figs. 4c and 4a show the most substantial brightness variations (enhancement of the additional plasma streams) in Ray+ and Ray–. The brightness variations are nearly the same in both rays in Fig. 4b, and the largest brightness changes take place in the helmet, which has the highest brightness. According to [17], a brightness increase or decrease in the rays reflects the enhancement or weakening of the additional plasma streams from the Sun. Starting from May 8, 1996 (Fig. 3), the angular distance between Ray+ and Ray– increases. The reason for this


Apr. 27, 1996(19:29)—Apr. 27, 1996(00:16)

Apr. 28, 1996(14:41)—Apr. 27, 1996(00:16)

Apr. 29, 1996(02:58)—Apr. 28, 1996(00:04)

May 3, 1996(11:31)—May 3, 1996(02:49)

Fig. 4. Differenced brightness $P(t) - P(t_\circ)$ of a fragment of the coronal image at successive times from April 27 to May 3, 1996. The time differences $t - t_o$ are labelled by the date and UT above the images. The arrows begin at the center of the Sun and point toward the north (up) and the western limb (to the right).

can be understood from the latitudes Λ of the brightness maxima in Ray+ (solid circles) and Ray− (open circles) and their brightness amplitudes P_R/P_{msb} as functions of time (or longitude $\Psi_L(\Lambda)$), which are shown in the upper and lower graphs in Fig. 5, respectively.

The latitude of the curve for Ray+ from April 30 to May 11, 1996 in the upper graph in Fig. 5 (solid circles) is nearly constant: $\Lambda \approx -2^\circ \pm 1^\circ$. It experiences only minor fluctuations $\approx\pm1^\circ$ (smaller than the angular size of the ray d), in disagreement with the calculated curve for an individual ray whose latitude in the plane of the sky is $\lambda = -2^\circ$ (the dashed curves in the upper graph in Fig. 5 are computed for an isolated ray for $\lambda = 1^\circ, 2^\circ, 3^\circ, 4^\circ, 5^\circ$). Therefore, in this case, the streamer belt is formed by a set of rays with similar brightness amplitudes P_R separated by $\approx 5^\circ$, probably of the type presented in Fig. 1. Ray+ is missing during April 27–29, 1996. It is likely that no bright rays were present at that time within several tens of degrees of the plane of the sky along the belt.

The situation for the Ray− curve (open circles) from April 27 to May 7, 1996 is similar to that for the Ray+ curve—the latitude of the ray experiences only small fluctuations $\approx\pm1^\circ$ about $\Lambda \approx -5^\circ$. However, the behavior of this curve during May 8–11, 1996 is in fairly good agreement with the calculated curve for an isolated ray with $\lambda = -5^\circ \ldots -4^\circ$ (to within $\approx\pm1^\circ$). This may indicate that, starting from May 8, essentially only one bright ray was present within several tens of degrees of the plane of the sky along the belt. The segment of the Ray− brightness-amplitude curve for May 8–11, 1996 (open circles in the lower graph) differs considerably from the calculated (dashed) curve for an isolated ray with $\lambda = -5^\circ$. Therefore, the brightness of the isolated Ray− was time dependent during May 8–11, 1996. It is noteworthy that the apparent latitude Λ increased by about 2° at the end of April 30, 1996 simultaneously for Ray+ (solid circles) and Ray− (open circles); see upper graph in Fig. 5. To all appearances, this can be considered a manifestation of a real latitude bend of the streamer belt as a whole. Thus, our analysis allows us to identify:

(a) two rows of rays (Ray+ and Ray−) throughout most of the belt (from April 30 to May 11, 1996),

(b) a latitude bend of the streamer belt as a whole (at the end of April 30, 1996; upper graph in Fig. 5).

Based on comparisons of calculations with observations, Wang *et al.* [19, 20] suggested that the observed two-ray structure of the streamer belt could result from small-scale bends of the belt. In this case, the neutral line should follow these bends and run through the ray-brightness maxima. However, in our view, the results presented above enable us to distinguish the fairly rare situations in which the belt bends

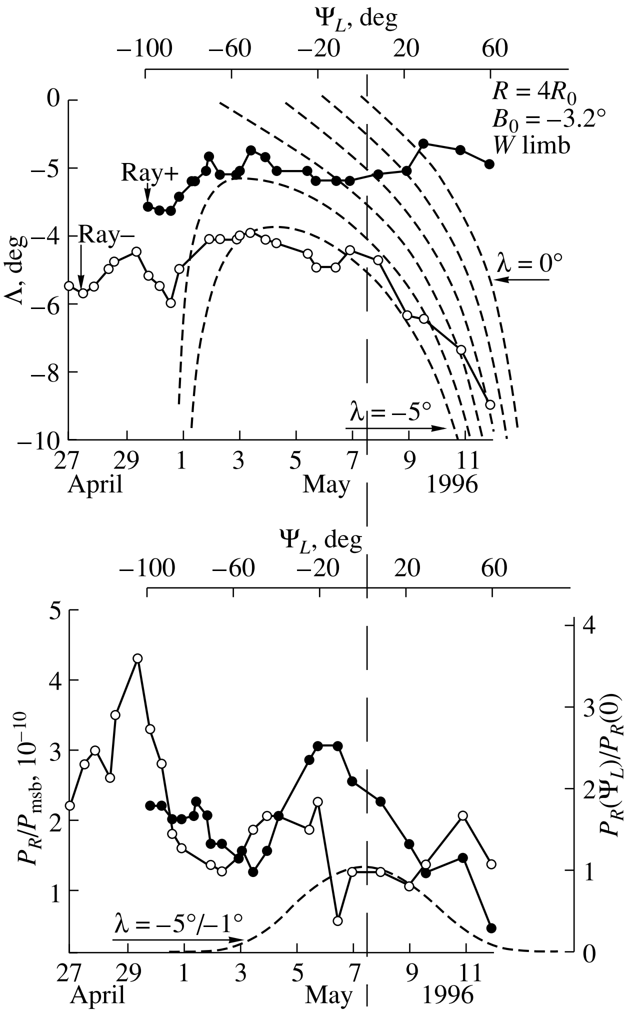


Fig. 5. Top: observed apparent latitudes Λ of the positions of the Ray+ and Ray– brightness maxima (solid and open circles, respectively) as a function of time (lower horizontal axis) or angular deviation from the plane of the sky Ψ_L (upper horizontal axis) at $R = 4.5R_\odot$ (Ψ_L is positive in the direction of solar rotation). The dashed curves were computed for $\lambda = 0°, 1°, 2°, 3°, 4°, 5°$, $B_\circ = -3.2°$, and $R = 4.5R_\odot$, and are close to the observational points. Bottom: same for the ray-brightness amplitude. The dashed curve was computed for $\lambda = -5°$, $B_\circ = -3.2°$, $R = 4.5R_\odot$.

from the more typical situation when the cross section of the streamer belt is formed by two rays of enhanced brightness, with the neutral line running between the two closely located rows of these rays.

The spatial ray pattern of the streamer belt suggested by the above results is shown schematically in Fig. 6.

6. DISCUSSION

The quasistationary structure of the streamer belt in the form of a set of pairs of enhanced-brightness rays, as is shown in Fig. 6, is rather surprising and requires discussion. First of all, recall that the rays remain essentially radial, so that their angular size, $d \approx 2°–3°$, is preserved at least in the range of heliocentric distances from $R \approx (5–6)R_\odot$ to $R \approx 30R_\odot$. At distances of $R < (3–4)R_\odot$, i.e., below the helmet apex, the adjacent rays of a pair pass around either side of the helmet as they approach the solar surface. The angular sizes of the rays in this region also remain constant, $d \approx (2°–3°)$ [15]. Since the coronal plasma has a high conductivity, the observed brightness distribution within each ray and the difference between adjacent rays in the time variations of their brightnesses provide evidence that each ray is isolated from its neighbors, which can be brought about only by the magnetic field. Thus, the rays are magnetic tubes, which may originate in the boundaries between supergranules [8]. This possibility is supported by the coincidence of the ray angular size d with the characteristic angular size of the supergranules, $\approx 2.0°–2.5°$. Note that the plumes forming the flows in coronal holes are also magnetic tubes whose bases have diameters of $d \approx 2°–3°$. However, in contrast to the magnetic tubes in the streamer belts, the brushes exhibit a super-radial divergence [21, 22], and their angular size varies from $\approx 2.5°$ at $R \approx 1.05R_\odot$ to $\approx 7.0°$ at $R \approx 5R_\odot$, and further to $\approx 15°$ at $R \approx 15R_\odot$ (where it exceeds the size of a supergranule by more than a factor of six [22]). It is quite plausible that the nature of the flows in both the streamer belts and coronal holes is controlled by the character of the streams in the magnetic tubes forming these flows. It also follows that the nature of the ray pattern of the streamer belt may be closely related to the nature of the supergranules at the solar surface. At the same time, the two-ray structure of the streamer belt may result from the development of an instability. In particular, Gubchenko *et al.* [23] have shown in a kinetic approach that, in streamer-belt-type current systems, the development of the so-called tearing and stratification instability can lead to the formation of a set of magnetic-tube (ray) pairs along the belt, which resemble the observed rays. If this is correct, we are dealing with collective properties of rarefied plasmas manifest via the formation of structures on cosmic scales.

7. CONCLUSIONS

(1) We have confirmed the previous result that the streamer belt is formed of a set of rays of enhanced brightness.

(2) The cross section of the streamer belt is frequently observed in the form of two closely spaced rays with different brightnesses. It is difficult to explain this phenomenon as a consequence of an ordinary bend of the belt. We suggest that, in the general

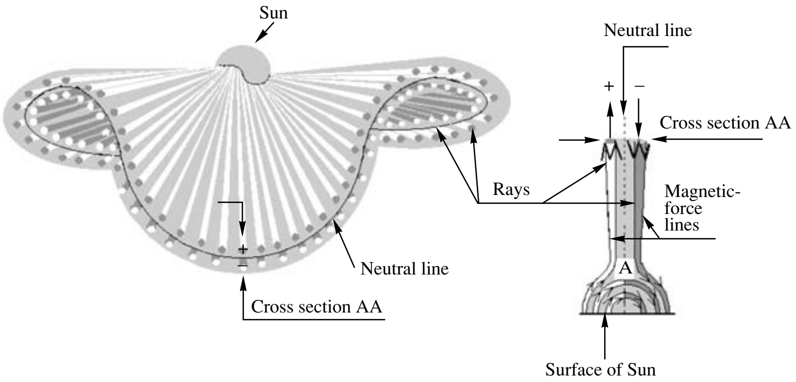


Fig. 6. Schematics of the spatial ray pattern of the coronal streamer belt (left) and the cross section (AA) of the streamer belt (right). The magnetic field is directed away from the Sun (+) in the light rays of the upper row of the streamer belt and toward the Sun (−) in the dark rays of the lower row. The apex of the helmet is labeled A in the right-hand schematic.

case, the belt is formed of a set of pairs of rays of enhanced brightness (or two closely spaced rows of rays). The distance between the rays in each pair is comparable to the ray size d. In general, the rays in any pair can have different brightnesses, and will have opposite magnetic polarities. Therefore, a neutral line of the radial component of the solar magnetic field probably runs along the belt between each pair of rays.

ACKNOWLEDGMENTS

The SOHO/LASCO data used here were obtained by a consortium including the Naval Research Laboratory (US), the Max-Planck-Institut für Aeronomie (Germany), the Laboratoire d'Astronomie (France), and the University of Birmingham (UK). This work was supported by the Program "Leading Scientific Schools of Russia" and the Federal Science and Technology Program in Astronomy.

REFERENCES

1. S. K. Vsekhsvyatskii, G. M. Nikol'skii, V. I. Ivanchuk, *et al.*, *The Solar Corona and Particle Radiation in Interplanetary Space* (Kiev. Univ., Kiev, 1965) [in Russian].
2. G. Newkirk, Annu. Rev. Astron. Astrophys. **5**, 213 (1967).
3. J. D. Bohlin, Solar Phys. **12**, 240 (1970).
4. R. A. Howard, M. A. Koomen, D. J. Michels, *et al.*, World Data Center for Solar-Terrestrial Physics. Report UAG-48. Synoptic Observations of the Solar Corona during Carrington Rotations 1580–1596 (July, 1975).
5. L. Svalgaard, J. M. Wilcox, and T. L. Duvall, Solar Phys. **37**, 157 (1974).
6. N. P. Korzhov, Solar Phys. **55**, 505 (1977).
7. R. A. Gulyaev, Usp. Fiz. Nauk **162**, 161 (1992) [Sov. Phys. Usp. **162**, 135 (1992)].
8. V. G. Eselevich and M. V. Eselevich, Solar Phys. **188**, 299 (1999).
9. V. G. Eselevich, V. G. Fainshtein, and M. V. Eselevich, Solar Phys. **200**, 259 (2000).
10. N. R. Sheeley, Jr., Y.-M. Wang, S. H. Hawley, *et al.*, Astrophys. J. **485**, 472 (1997).
11. Y.-M. Wang, N. R. Sheeley, Jr., R. A. Howard, *et al.*, Geophys. Res. Lett. **26**, 1203 (1999).
12. Y.-M. Wang, N. R. Sheeley, Jr., R. A. Howard, *et al.*, Geophys. Res. Lett. **26**, 1349 (1999).
13. V. G. Eselevich and M. V. Eselevich, Solar Phys. **203**, 165 (1999).
14. R. M. Illing and A. J. Hundhausen, J. Geophys. Res. **91**, 10951 (1986).
15. V. G. Eselevich and M. V. Eselevich, Solar Phys. **197**, 101 (2000).
16. A. J. Hundhausen, J. Geophys. Res. **98**, 13177 (1993).
17. M. V. Eselevich and V. G. Eselevich, Astron. Zh. **81**, 757 (2004) [Astron. Rep. **48**, 688 (2004)].
18. V. G. Eselevich and M. V. Eselevich, in *Abstrats of Papers to the All-Russian Conference "Magnetic Fields and the Three-Dimensional Structure of the Solar Atmosphere", Irkutsk, August 25–29, 2003*, p. 30.
19. Y.-M. Wang, N. R. Sheeley, Jr., and N. B. Rich, Geophys. Res. Lett. **27**, 149 (2002).
20. Y.-M. Wang, N. R. Sheeley, Jr., D. G. Socker, *et al.*, J. Geophys. Res. **105**, 25133 (2000).
21. K. M. Thieme, E. Marsch, and R. Schwenn, *Proceedings of Sixth International Solar Wind Conference*, Ed. by V. J. Pizzo, T. Holzer, and D. G. Sime (NCAR, Colorodo, 1987), p. 317.
22. C. E. DcForest, J. T. Oeksema, J. B. Gurman, *et al.*, Solar Phys. **175**, 393 (1997).
23. V. M. Gubchenko, V. V. Zaitsev, Kh. K. Birnat, *et al.*, in *Proceedings of Conference of the CIS and Baltic States on Current Problems in the Physics of Solar and Stellar Activity* (Nizhni Novgorod, 2003).

Translated by A. Getling