Properties of Quasar–Galaxy Associations and Gravitational Mesolensing by Halo Objects

Yu. L. Bukhmastova

St. Petersburg State University, St. Petersburg, Russia

Received November 14, 2000

Abstract—A new catalog of 8382 close quasar–galaxy pairs is presented. The catalog was composed using published catalogs of quasars and active galactic nuclei containing 11 358 objects, as well as the LEDA catalog of galaxies, which contains on the order of 100 thousand objects. Based on these new data, the dependence of the number of pairs on $a = z_G/z_Q$ is analyzed, where z_G and z_Q are the redshifts of the galaxy and quasar, respectively, revealing an excess of pairs with $a < 0.1$ and $a > 0.9$. This means that the galaxies in pairs are preferably located close to either the observer or the quasar and avoid intermediate distances along the line of sight to the quasar. Computer simulations demonstrate that it is not possible to explain this number of pairs with the observed distribution in a as the result of chance positional coincidences with a uniform spatial distribution of galaxies. Data on globular clusters show that the excess of pairs with $a < 0.1$ and $a > 0.9$ is consistent with the hypothesis that we are observing distant compact objects that are strongly gravitationally lensed by transparent lenses with a King mass distribution located in the halos of nearby galaxies. The Hubble diagram for galaxies and quasars is presented. Observational tests of the mesolensing hypothesis are formulated. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The problem of physical associations between distant quasars and nearby galaxies has been discussed in the literature for more than thirty years. Burbidge *et al.* [1] published a catalog of galaxy–quasar associations containing 577 quasars and 500 galaxies. They indicated that quasars show a tendency to be located near the halos of normal galaxies much more often than expected in the case of chance projections, and that this physical relationship requires an explanation. The physical connection between quasars and galaxies was first interpreted as a result of gravitational lensing by halo stars in the galaxies (microlensing) by Canizares [2]. However, in subsequent studies [3–5], it was shown that gravitational microlensing cannot explain the observations, due to the extremely low surface density of weak quasars, which should be amplified by microlensing. In addition, Arp [6] presented several arguments against the gravitational microlensing hypothesis.

(1) Quasars in pairs with nearby galaxies are predominantly separated from the galaxy by several galaxy diameters, so that microlenses should occupy a huge volume of space around the galaxies. The observed number of associations requires anomalously large halo masses.

(2) In groups of galaxies, quasar–galaxy associations are more often encountered for companion galaxies than for dominant galaxies in the group.

(3) Analysis of archival data on the variability of quasars in associations [7] showed a lack of variability

on timescales of several dozen years, in contradiction with expectations of the microlensing hypothesis.

As a result, Arp [6] concluded that the redshifts of quasars could have a non-cosmological nature.

Baryshev and Ezova [8] attempted to explain the appearance of quasar–galaxy pairs by theorizing that distant active galactic nuclei experienced mesolensing by globular clusters in the halos of more nearby galaxies. In this case, the quasar corresponds to a distant active galactic nucleus whose brightness is amplified by several magnitudes. Estimates of the expected number of associations based on calculations of the probability that active galactic nuclei will be lensed lead to values of 10^2–10^4, assuming brightness enhancements of 5^m–8^m. If the mean brightness enhancement is lowered 3^m, the number of expected associations grows to 10^5 (based on the number of active galactic nuclei being $\sim 10^4/\text{deg}^2$ for galaxies with apparent magnitudes to 29^m). In light of these hypotheses, the arguments of Arp [6] lose their force.

The current paper is a logical continuation of [8]. We do not yet consider the problem of the luminosity function of quasars [4, 5], which will be the focus of the next paper in this series. Here, we compose a new, appreciably expanded, catalog of pairs. The number of pairs selected using the new observational material (8382 pairs) is in agreement with the predictions of theoretical computations. The search for pairs was carried out in such a way that the linear distance between the galaxy and projected quasar does not exceed 150 kpc. We used four criteria for the pair search:

(1) there are the spatial coordinates α, δ and z of the quasars and galaxies, so that the quasar magnitude m_Q is known;

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

(3) the quasar should be projected onto the halo of the galaxy;

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

The criteria for selecting new pairs were based on the main assumptions of the gravitational lensing hypothesis. In particular, this theory led to criteria (2) and (3). Criterion (4) was derived in [8].

The structure of the paper is as follows. Section 2 describes the algorithm for selecting the pairs. Section 3 is dedicated to a general analysis of the resulting pairs. In Section 4, we present histograms reflecting an important property of the associations and attempt to analyze the result obtained. Section 5 is concerned with the Hubble diagram for galaxies and quasars, and Section 6 discusses possible tests for the adopted model. We briefly summarize our main conclusions in Section 7.

2. PAIR SELECTION ALGORITHM

Since the quasars are located at greater distances than the galaxies, we will initially project the quasar onto a celestial sphere with a radius corresponding to the distance to the galaxy. We denote the smallest distance between the quasar and galaxy projections $l = GQ$ (Fig. 1). We now solve for the spherical triangle PGQ with sides $PG = 90 - \delta_G$, $PQ = 90 - \delta_Q$, and $GQ = l$. According to the cosine theorem, $\cos l = \cos\delta_G\cos\delta_Q\cos(\alpha_Q - \alpha_G) +$
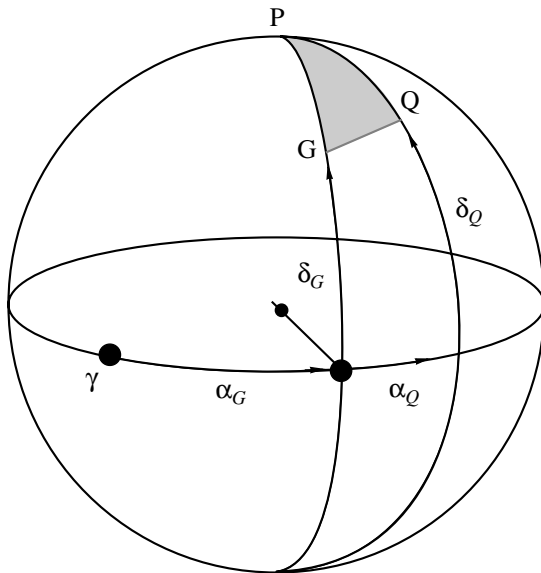
Fig. 1. Celestial sphere of a galaxy. P is the pole, G the galaxy, and Q the projection of the quasar onto the celestial sphere.

$\sin\delta_G\sin\delta_Q$. The linear distance x in kpc is $x = z_G\dfrac{\sqrt{1 - \cos^2 l}}{\cos l} \times 5 \times 10^6$ for $H = 60$ km s^{-1} Mpc^{-1}. Thus, if $x \le 150$ kpc and $z_G > 4 \times 10^{-4}$, the quasar–galaxy pair is specified. The input quasar catalog [9] is available at the web address

ftp://cdsarc.u-strasbg.fr/cats/VII/207.

The galaxy catalog [10] is available at the address

http://www-obs.univ-Lyon1.fr.

Our application of the above simple selection operation resulted in the new catalog of associations, which can be found at the web address

http://www.astro.spbu.ru/staff/baryshev/gl_dm.htm.

3. MAIN PROPERTIES OF THE CLOSE QUASAR–GALAXY PAIRS

The selection criteria indicated above are satisfied by 77843 galaxies and 11358 quasars. Of these, 1054 galaxies and 3164 quasars were included in pairs, making up a total of 8382 quasar–galaxy pairs. Cases when one quasar is near several galaxies are encountered fairly frequently. Examples of such situations are presented in Table 1. One simple explanation for the data can be obtained if the distribution of galaxies along the line of sight has a fractal nature; i.e., if, in accordance with [11], we adopt the following parametric representation for the concentration of objects along the line of sight between the source and observer:

$$n_l(R) = 0.5 n_{ol}\left[\left(\frac{R}{R_0}\right)^{D_F - 3} + \left(\frac{R_s - R}{R_0}\right)^{D_F - 3}\right],$$

where $n_{ol}(R_0)$ is the concentration of objects at the distance R_0, R is the distance to the lens, R_s is the distance to the source, and D_F is the fractal dimension of the galaxy distribution. Usually, a uniform distribution of lenses is obtained if $D_F = 3$, in which case $n_l(R) = $ const. The observed value for the fractal dimension is close to $D_F \sim 2$ [12, 13]. The character of this relation is such that, if some quasar (according to the Barnotti–Tyson hypothesis, the nucleus of an active galaxy) is projected onto the halo of a nearby galaxy, the probability that the quasar is also projected onto other galaxies is enhanced, since the galaxy is part of a fractal structure. The inverse is also true, since two objects of a single class along the line of sight in a fractal structure have equal validity relative to the path of a light ray. This means that cases when there are several quasars near a single galaxy should be rare. Examples of such cases are presented in Table 2. The galaxies PGC 0003290, PGC 0003589, PGC 0001014, PGC 0003238, PGC 0003721, PGC 0002789, and a number of others are surrounded by tens or hundreds of quasars.

Table 1. Examples from the catalog of quasar–galaxy associations showing several galaxies near a single quasar*

1	2	3	4	5	6	7	8	9
7627	TEX 1322+479	2.26	20.5	PGO 0045939	0.0005	14.207	112.6	0.0002
7628	TEX 1322+479	2.26	20.5	PGO 0046039	0.0006	11.697	140.7	0.0003
7629	TEX 1322+479	2.26	20.5	PGO 0046127	0.0009	13.477	115.6	0.0004
7630	TEX 1322+479	2.26	20.5	PGO 0047270	0.0015	13.865	139.1	0.0007
7631	TEX 1322+479	2.26	20.5	PGO 0047404	0.0015	8.298	141.3	0.0007
7632	TEX 1322+479	2.26	20.5	PGO 0047413	0.0016	9.938	141.9	0.0007
7880	SBS 1400+541	0.646	17.5	PGO 0049448	0.0005	12.37	46.7	0.0007
7881	SBS 1400+541	0.646	17.5	LEDA 0140246	0.0006		0.7	0.0009
7882	SBS 1400+541	0.646	17.5	PGC 0050063	0.0008	7.919	31.1	0.0012
7883	SBS 1400+541	0.646	17.5	PGC 0050216	0.0009	10.883	37.5	0.0014
7884	SBS 1400+541	0.646	17.5	PGC 0050262	0.0010	13.902	63.2	0.0016
7885	SBS 1400+541	0.646	17.5	LEDA 0165626	0.0005		15.5	0.0008
7886	SBS 1400+541	0.646	17.5	LEDA 0165627	0.0010		38.0	0.0015
7887	SBS 1400+541	0.646	17.5	LEDA 0165629	0.0010		48.7	0.0015
8324	Q 2315–4230	2.83	20.0	LEDA 0123614	0.0538	16.081	71.8	0.019
8325	Q 2315–4230	2.83	20.0	PGC 0071001	0.0052	10.644	64.9	0.0019
8326	Q 2315–4230	2.83	20.0	PGC 0071031	0.0053	11.145	57.2	0.0019
8327	Q 2315–4230	2.83	20.0	PGC 0071066	0.0056	10.908	99.8	0.0020

* The columns contain: (1) pair number in the original catalog, (2) quasar name, (3) quasar redshift, (4) quasar apparent magnitude, (5) galaxy name, (6) galaxy redshift, (7) galaxy apparent magnitude, (8) distance between the galaxy and projection of the quasar onto the plane of the galaxy in kpc, (9) $a = z_G / z_Q$. The table presents data for three quasars and eighteen galaxies with which they are associated.

Why should a distant quasar be projected onto the halo of a more nearby galaxy? There exist at least three possible explanations.

(1) If the galaxy is very close to the observer, it covers a relatively large fraction of sky. There is some probability for the chance projection of quasars onto the halo of the nearby galaxy.

(2) If the spatial distribution of galaxies is fractal and at least some fraction of quasars are the nuclei of active galaxies (i.e., they belong to the same general class of objects–galaxies), there should be a significant number of pairs due to the properties of the fractal distribution, as noted above.

(3) Due to gravitational mesolensing by halo objects with intermediate mass, such as globular clusters, whose number in galactic halos can reach several thousand [4], some fraction of objects viewed through these lenses will appear to have enhanced luminosities and will be interpreted as being quasars [3]. As a result, the clustering of the mass in the halos of nearby galaxies "magnifies" the light from a distant galaxy that would otherwise simply not be visible. The result is the detection of a close quasar–galaxy pair. The appearance of numerous quasars around a single galaxy can be explained by the large number of globular clusters acting as lenses in the halo.

We will analyze the distribution of the galaxies along the line of sight from the observer to the quasar in order to elucidate the origin of the observed associations.

4. MUTUAL LOCATIONS OF THE GALAXIES AND QUASARS IN ASSOCIATIONS

The distance to the quasar (which we take to be indicated by its redshift) is denoted z_Q. The analogous distance to the galaxy is z_G. We introduce the quantity $a = z_G/z_Q$, which is the normalized distance from the observer to the galaxy in a system in which the distance to the quasar is equal to unity. We determined a for all the quasar–galaxy pairs. The result is shown as a histogram of x as a function of the number of pairs with that a value (not to scale). We obtain four relations for various halo sizes (50 and 150 kpc) and quasar redshifts (Fig. 2).

Even without a more careful analysis, we can see that galaxies in associations are preferably located either near the observer, $a < 0.1$, or near the quasar, $a > 0.9$, avoiding intermediate distances between the observer and quasar. The appearance of the first "tail" at $a < 0.1$ is not surprising. Most galaxies are located nearby, so that the substantial contribution to the number of pairs from such galaxies could be an observational selection effect. What about the second tail at $a > 0.9$, which is stable to variation of the halo size and quasar redshift? Similar dependences were obtained for the catalog [1]
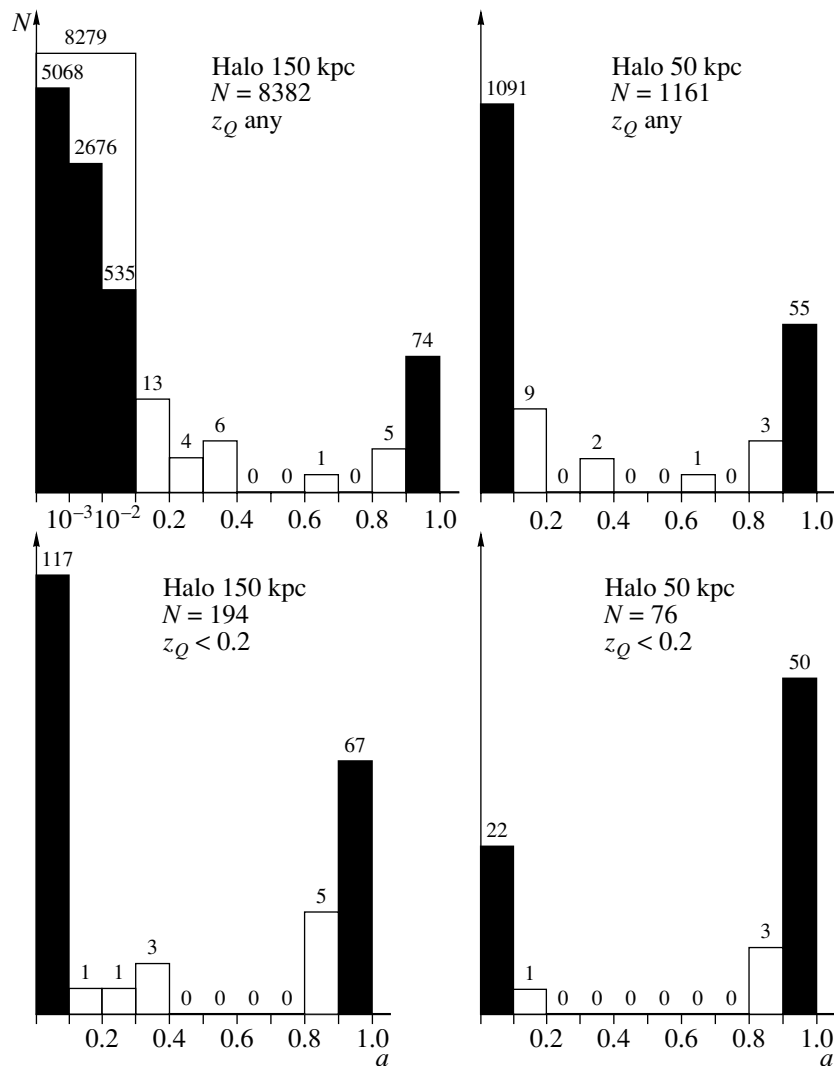
Fig. 2. Dependence of the number of quasar–galaxy associations found by comparing the quasar catalog [9] and the galaxy catalog [10] on the quantity $a = z_G/z_Q$ for various halo sizes and quasar redshifts. The study includes 8382 pairs.

in [8], based on 241 pairs. The data have now been expanded appreciably to 8382 pairs, but the dependences have remained essentially unchanged. Let us try to elucidate their intrinsic properties. We assume that the following effects are responsible for these dependences to some degree:

(1) the fractal distribution of galaxies along the line of sight, with a fractal dimension close to two;

(2) gravitational lensing by globular clusters in the galactic halos (or by other halo objects displaying clustering with a King mass distribution);

(3) chance projected pairs.

We studied this question using computer simulations. We fixed the galaxy positions and scattered the quasars randomly, with z_Q from 0.1 to 3. In this way, we attempted to remove the effect of possible lensing. There remain two effects that could lead to the formation of pairs: random projections and the nature of the distribution of the galaxies along the line of sight. Fig-

ure 3 shows the result of identifying pairs for the case of a random distribution of background quasars. The results have changed both qualitatively and quantitatively. The number of pairs has dropped dramatically, to 2000 (compared to 8382), and the second tail in the distributions has disappeared; i.e., its presence in the observational data is not due to random effects.

As a second step in our computer simulations, we randomly specified the positions of 77483 galaxies, with z_G from 0 to 0.25, and of 11358 quasars, with z_Q from 0.1 to 3. Figure 4 shows the result of identifying pairs for this case of randomly distributed galaxies and quasars. An analysis of this histogram indicates that about 1200 pairs could be obtained due to random positional coincidences. However, another qualitative result is more reliable and interesting: in the case of chance coincidences, the quasars in pairs are not projected onto nearby galaxies whose distances are close to those of the quasars; i.e., again, the tail at $a > 0.9$ disappears.
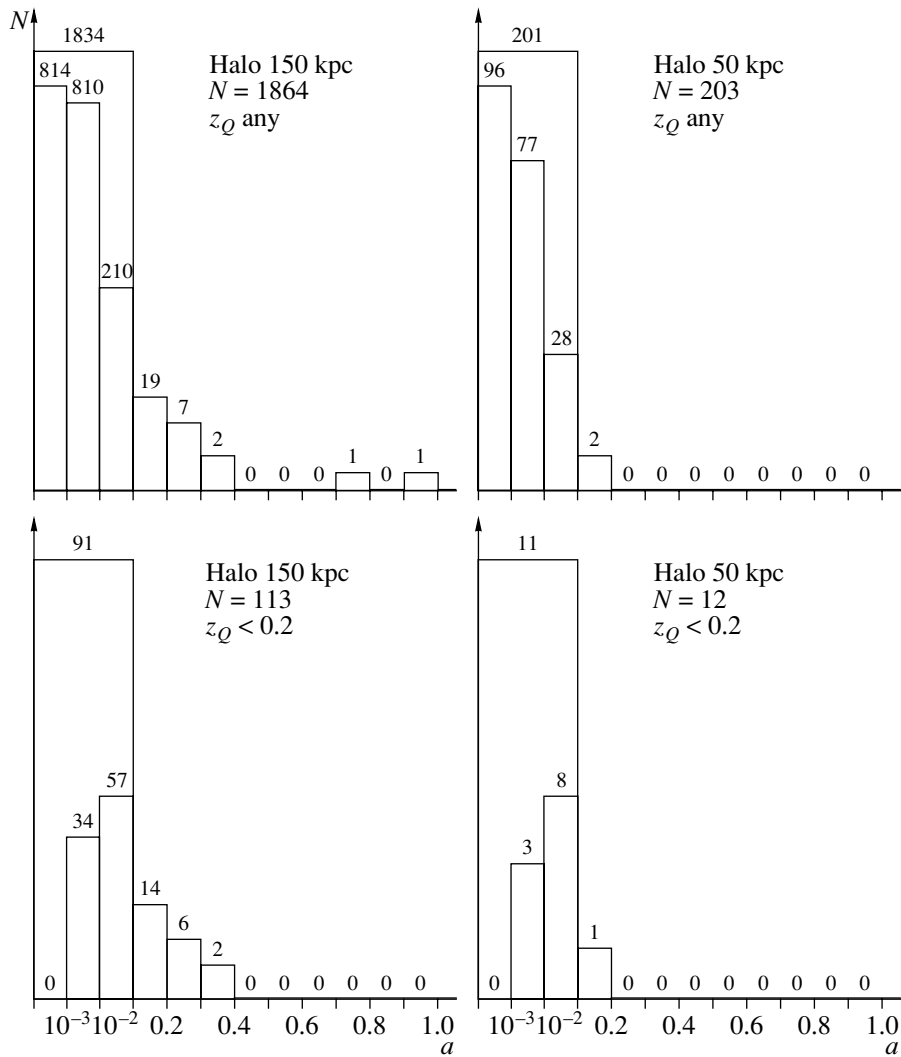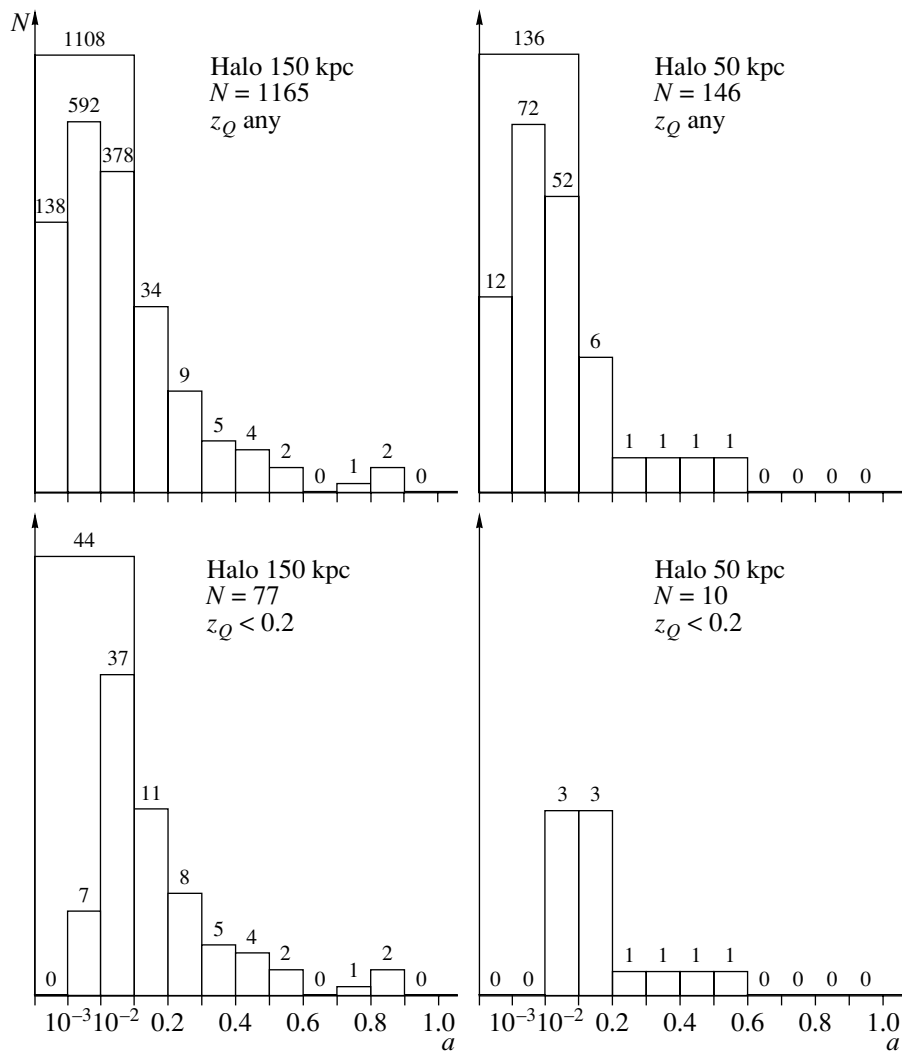
Fig. 3. Dependence of the number of quasar–galaxy associations found by comparing the galaxy catalog [10] and a catalog of 11358 quasars with randomly specified positions (with z_Q from 0.1 to three) on $a = z_G/z_Q$ for various halo sizes and quasar redshifts.

Thus, crude estimates of the model relations suggest that about 15% of the pairs formed due to random projected positional coincidences. In the framework of our adopted assumptions, the remaining pairs form as a result of the fractal distribution of galaxies along the line of sight and due to lensing by halo objects with a King mass distribution, with the effects of lensing and of the fractal galaxy distribution complementing each other [8].

Let us now consider the following selection of 8382 quasar–galaxy pairs. If some quasar is projected onto the halos of several galaxies, we select the pair with the minimum distance between the quasar and galaxy projected onto the plane of the galaxy. If several quasars are projected onto a single galaxy, we consider all these pairs, since several hundred, or even several thousand, globular clusters can be located in the galaxy's halo, each of which can act as a lens, resulting in several amplified active nuclei of distant galaxies, inter-

preted as quasars. We now construct the same histogram for such a sample of 3164 quasars and 1054 galaxies, shown in Fig. 5. Qualitatively, the results are as before: there are excesses of pairs with $a < 0.1$ and $a > 0.9$. Note the numerical values for $a > 0.9$ (the tail with $a < 0.1$ has a high fraction of chance pairs, and is not informative in a numerical sense). It turns out that the postulated lenses are located primarily in halos less than 50 kpc in size. The histogram shows that about 83% of the lenses are located in such halos, whose size does not exceed 50 kpc while the remaining 17% of the lenses are located at distances of 50 to 150 kpc. If these lenses are globular clusters, they should be located at precisely such distances from the centers of their galaxies, in accordance with the data of [14].

As an example, we present data for 146 globular clusters in the Milky Way (Fig. 6). We can see that 140 of the 146 globular clusters are located in the halo out to distances of 50 kpc, in agreement with the histo-

Fig. 4. Dependence of the number of quasar–galaxy associations found by comparing a catalog of 77 483 galaxies with randomly specified positions (with z_G from 0 to 0.25) and a catalog of 11 358 quasars with randomly specified positions (with z_Q from 0.1 to three) on $a = z_G/z_Q$ for various halo sizes and quasar redshifts.

grams in Figs. 2 and 5. The same dependences of the number of globular clusters with distance from the galactic center are characteristic, for example, of the Abell galaxies A754, A1644, A2124, A2147, A2151, and A2152 [15], as well as for NGC 4874, NGC 4889, NGC 4472, and NGC 4486 [16].

5. THE HUBBLE DIAGRAM $z = f(m)$ FOR GALAXIES AND QUASARS

Figure 7 presents a Hubble diagram $z = f(m)$, for the 77 483 LEDA galaxies and 11 358 quasars. All the quasars are shifted to the left from the straight line corresponding to the 0.2 m law.

In the context of the gravitational lensing hypothesis, this shift for the quasars can be explained as the effect of brightness amplification during lensing. If the quasars are the amplified nuclei of active galaxies, they

can be shifted to the strip corresponding to the brightest galaxies by "removing" several magnitudes from their brightnesses. The presence of a distinct lower boundary to the region occupied by the quasars can be explained by the fact that there are objects brighter than -23^m in the initial catalog of quasars.

6. OBSERVATIONAL TESTS

Since the observational data for the quasar–galaxy associations are consistent with the gravitational lensing model, this raises the question of further observational tests of the model. We propose the following observational tests for the properties of the quasars that are members of associations.

(1) The expected angular separation of multiple images of quasars due to lensing by objects such as
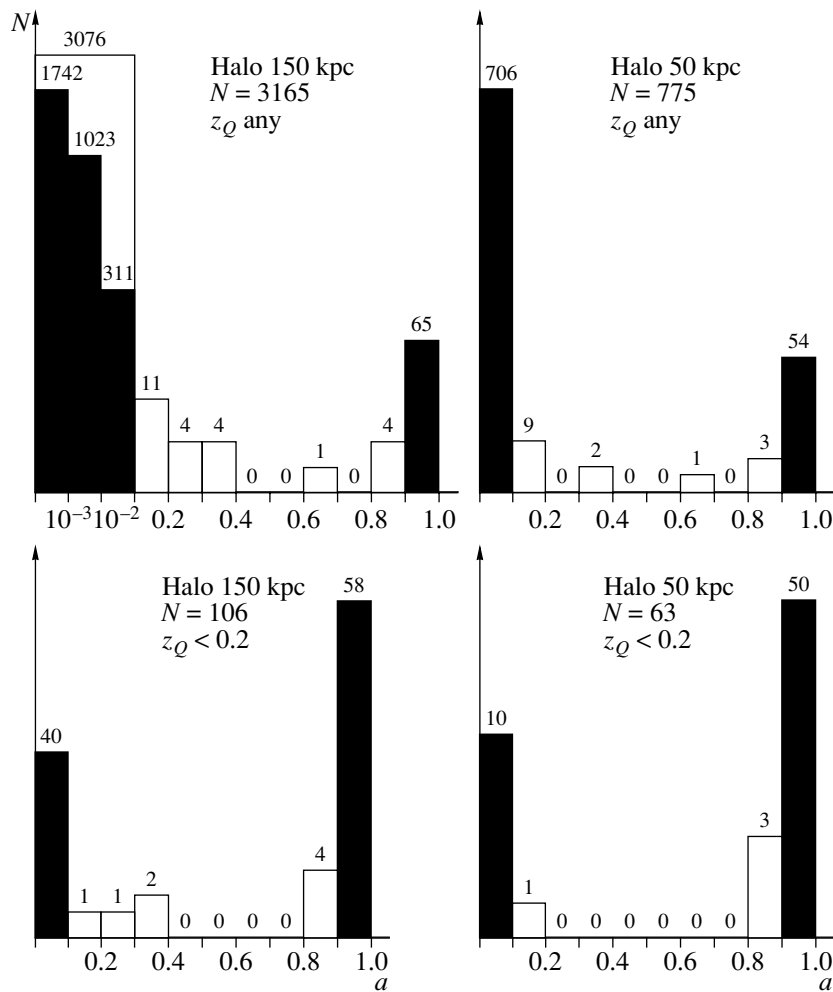
Fig. 5. Dependence of the number of quasar–galaxy associations found by comparing the quasar catalog [9] and the galaxy catalog [10] on $a = z_G/z_Q$ for various halo sizes and quasar redshifts. The study includes 3164 pairs obtained after imposing additional restrictions (see text).

globular clusters is several milliarcseconds. Therefore, it is important to study the structure of the compact radio components of quasars in associations that have sufficient radio fluxes for VLBI observations. Of course, the theory does not exclude the possibility of a single image being formed. In addition, the image splitting could be unobserved due to insufficient dynamical range for the observations.

(2) The expected variability of the quasar's brightness due to the motion of the globular clusters takes place on time scales of more than a thousand years. Variability on timescales of less than a year is also possible, as a consequence of microlensing by individual globular cluster stars.

(3) Since globular clusters are located in the halos of galaxies, absorption lines with z_{abs} corresponding to the redshift of the galaxy may arise in the quasar spectra. It would be interesting to continue the work begun in [17], comparing the properties of quasars with absorption lines and quasars in associations. In particular, the
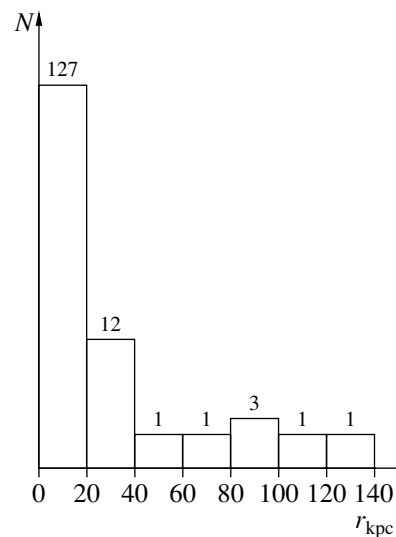

Fig. 6. Dependence of the number of globular clusters on distance from the center of their galaxy for 146 globular clusters in the Milky Way [14].

Table 2. Examples from the catalog of quasar–galaxy associations showing a number of quasars near a single galaxy*

1	2	3	4	5	6	7	8	9
895	Q 0054+0200	1.872	18.4	PGC 0003844	0.0008	9.598	137.6	0.0004
934	Q 0054+0236	1.654	17.6	PGC 0003844	0.0008	9.598	139.1	0.0005
948	Q 0055+0141	2.232	18.6	PGC 0003844	0.0008	9.598	123.1	0.0003
983	Q 0055+0225	0.373	18.6	PGC 0003844	0.0008	9.598	119.6	0.0021
995	UM294	1.914	17.1	PGC 0003844	0.0008	9.598	146.4	0.0004
1059	PC0056+0125	3.154	18.6	PGC 0003844	0.0008	9.598	98.8	0.0002
1060	Q 0056+0009	0.613	18.0	PGC 0003844	0.0008	9.598	148.9	0.0013
1066	Q 0056+0118	1.1	18.1	PGC 0003844	0.0008	9.598	98.7	0.0007
1095	Q 0057+0000	0.776	17.2	PGC 0003844	0.0008	9.598	149.4	0.0010
1106	Q 0057+0230	0.716	18.3	PGC 0003844	0.0008	9.598	90.5	0.0011
1147	PHL938	1.959	17.16	PGC 0003844	0.0008	9.598	67.4	0.0004
1149	Q 0058+0215	2.868	18.91	PGC 0003844	0.0008	9.598	71.2	0.0003
1171	Q 0058+0218	0.929	17.7	PGC 0003844	0.0008	9.598	67.0	0.0008
1182	Q 0058+0121	1.432	17.6	PGC 0003844	0.0008	9.598	66.8	0.0005
1193	Q 0058+0205	0.6	18.5	PGC 0003844	0.0008	9.598	58.3	0.0013
1195	Q 0059+0147	1.143	18.0	PGC 0003844	0.0008	9.598	55.7	0.0007
1255	Q 0059+0035	2.545	18.5	PGC 0003844	0.0008	9.598	95.1	0.0003
1311	Q 0100+0228	1.543	18.4	PGC 0003844	0.0008	9.598	50.8	0.0005
1314	Q 0100+0146	1.909	18.5	PGC 0003844	0.0008	9.598	30.3	0.0004
1315	UM301	0.393	16.39	PGC 0003844	0.0008	9.598	32.0	0.0020
1330	Q 0.100+0106	1.405	18.5	PGC 0003844	0.0008	9.598	57.3	0.0005
1335	Q 0101+0024	1.436	18.9	PGC 0003844	0.0008	9.598	100.6	0.0005
1344	Q 0101+0009	0.394	17.4	PGC 0003844	0.0008	9.598	116.6	0.0020
1370	Q 0101+0118	1.133	18.61	PGC 0003844	0.0008	9.598	38.3	0.0007
1403	Q 0102+0036	0.649	18.4	PGC 0003844	0.0008	9.598	84.3	0.0012
1408	Q 0102+0241	1.509	18.6	PGC 0003844	0.0008	9.598	55.3	0.0005
1414	LMA 15	2.7	0.0	PGC 0003844	0.0008	9.598	6.7	0.0003
1444	Q 0103+0234	1.7	18.0	PGC 0003844	0.0008	9.598	49.8	0.0005
1448	Q 0103+0024	1.075	17.4	PGC 0003844	0.0008	9.598	99.7	0.0007
1454	PC 0103+0123	3.066	19.7	PGC 0003844	0.0008	9.598	36.4	0.0003
1467	Q 0103+0123	0.782	18.8	PGC 0003844	0.0008	9.598	37.9	0.0010
1475	BRI 0103+0032	4.437	18.6	PGC 0003844	0.0008	9.598	92.5	0.0002
1486	Q 0103-0014	1.629	18.5	PGC 0003844	0.0008	9.598	144.7	0.0005
1487	Q 0104+0001	0.91	18.3	PGC 0003844	0.0008	9.598	127.6	0.0008
1493	Q 0104+0030	1.874	18.5	PGC 0003844	0.0008	9.598	96.9	0.0004
1499	PC 0104+0215	4.171	19.73	PGC 0003844	0.0008	9.598	41.4	0.0002
1578	PKS 0160+01	2.107	18.39	PGC 0003844	0.0008	9.598	73.0	0.0004
1606	Q 0107+0051	0.966	19.0	PGC 0003844	0.0008	9.598	106.4	0.0008
1613	Q 0107+0022	1.968	18.3	PGC 0003844	0.0008	9.598	131.3	0.0004
1640	MS 01080+0139	0.713	17.49	PGC 0003844	0.0008	9.598	96.3	0.0011
1643	Q 0108+0028	2.005	18.25	PGC 0003844	0.0008	9.598	134.2	0.0004
1651	Q 0108+0030	0.428	19.0	PGC 0003844	0.0008	9.598	135.8	0.0018
1678	MS 01094+0242	0.262	18.1	PGC 0003844	0.0008	9.598	132.0	0.0029
1682	UM 87	2.343	17.3	PGC 0003844	0.0008	9.598	126.5	0.0003
1705	PB 6325	0.774	17.8	PGC 0003844	0.0008	9.598	146.0	0.0010
1708	PB 6327	1.509	18.1	PGC 0003844	0.0008	9.598	149.5	0.0005

* The columns contain: (1) pair number in the original catalog, (2) quasar name, (3) quasar redshift, (4) quasar apparent magnitude, (5) galaxy name, (6) galaxy redshift, (7) galaxy apparent magnitude, (8) distance between the galaxy and projection of the quasar onto the plane of the galaxy in kpc, (9) $a = z_G/z_Q$. The table presents data for the galaxy PGO 0003844 and 47 quasars with which it is associated.
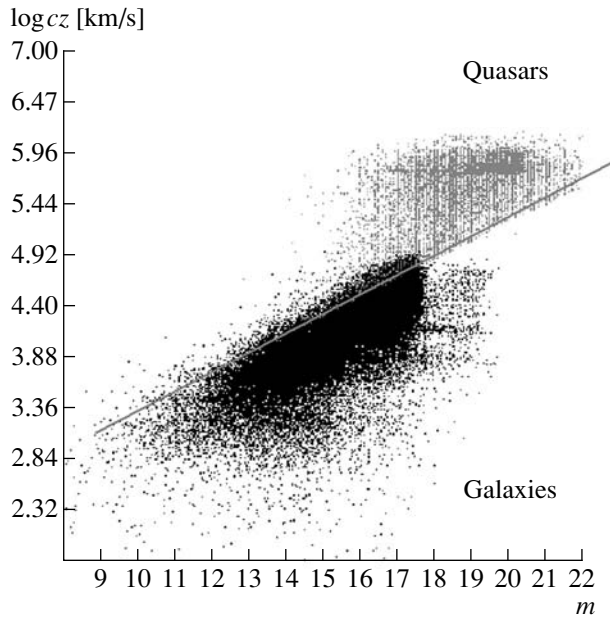
Fig. 7. Hubble diagram $z = f(m)$ for 77 483 LEDA galaxies [10] and 11358 quasars [9]. The quasars, marked in grey, are shifted to the left from the line corresponding to the Hubble law. The galaxies are shown in black.

well known increase in the number of absorption lines with approach toward the quasar could be associated with the fractal nature of the large-scale distribution of galaxies along the line of sight. In this case, we should also expect an increase in the number of absorption lines with approach toward the observer.

(4) Comparisons of the spectral properties of quasars in associations with those of various types of active galactic nuclei could serve as a probe of the structure of the emission-line formation region, which could be affected differently during lensing by King objects.

(5) Analyzing the properties of quasar–galaxy associations, we can estimate the fraction of galaxies of various morphological types in the total number of associations and the distribution of global clusters in each type of galaxy.

(6) In the case of quasars with jets, the curvature in the jet trajectories that is not associated with real curved motion of the material ejected from the core, but is related to the refraction of rays in the presumed lens is possible.

(7) In the hypothesis of gravitational microlensing of quasars, the Hubble diagram presented above can be used to estimate the amplification coefficient for quasars, by moving each quasar to the strip of brightest galaxies. Knowing the dependence of the amplification coefficient on the position of the lens between the observer and source, it is possible to estimate the distance to the postulated lens.

(8) It is possible to devise a computer simulations of the large-scale fractal distribution of galaxies and qua-

sars, in order to organize searches for quasar–galaxy pairs and compare the resulting data with the catalog of observed associations.

7. CONCLUSION

We have derived a new catalog of close quasar–galaxy pairs including 1054 galaxies and 3164 quasars comprising 8382 pairs. All the galaxies in the associations show a strong tendency to be located close to either the observer or the quasar, avoiding intermediate positions along the line of sight to the quasar. This property of the pairs and their considerable number can easily be explained if we adopt the following assumptions.

(1) Galaxies on scales ~150 mpc have a fractal distribution with a fractal dimension close to two. A uniform galaxy distribution makes it difficult to explain the enhanced number of pairs with $a > 0.9$.

(2) The quasars in associations could be gravitationally amplified active nuclei of distant galaxies. There are sufficient numbers of active galactic nuclei to provide the observed number of quasars if there is a probability of 10^{-4} for a brightness amplification by 3^m.

(3) The role of the gravitational lenses can be played by objects such as globular clusters or clusters of dark matter characterized by a King density distribution, located primarily in galactic halos at distances of up to 50 kpc from the galactic center. The importance of the King distribution is that it has a conical caustic that can explain the enhanced probability for the galaxies harboring the lenses to be near either the observer or the quasar. The point and isothermal-sphere lens models that are often considered in the literature give an enhanced probability for the galaxy to be located in a central position between the observer and the source, in contradiction with the observed properties of the associations. The derived histograms indicate that quasars in associations are most often projected onto the galactic halo at distances out to 50 kpc, in spite of the fact that the pair selection criteria allowed projections out to 150 kpc. This provides an additional argument that the relationship between the distant quasars and nearby galaxies is associated with globular clusters in the galaxy halos.

(4) Quasars that were not included in the catalog of associations can also be gravitationally amplified galactic nuclei, since it is probable that a galaxy surrounded by lensing objects is present near the projection of the quasar and is simply not detected. This possibility follows from the above histograms.

The hypothesis that we are dealing with gravitational lensing of the distant nuclei of galaxies can provide a simple physical interpretation of the Arp effect; i.e., the observed frequency with which quasar–galaxy associations are encountered. The quasar redshifts are cosmological in nature, and do not require any new physics.

ACKNOWLEDGMENTS

REFERENCES

1. G. Burbidge, A. Hewitt, J. V. Narlikar, and P. Das Gupta, Astrophys. J., Suppl. Ser. **74**, 675 (1990).

2. C. R. Canizares, Nature **291**, 620 (1981).

3. M. Vietri and J. P. Ostriker, Astrophys. J. **267**, 488 (1983).

4. E. V. Linder and P. Schneider, Astron. Astrophys. **204**, L8 (1988).

5. P. Schneider, J. Ehlers, and E. E. Falko, *Gravitational Lenses* (Springer-Verlag, New York, 1992).

6. H. Arp, *Quasars, Red-shifts, and Controversies* (Cambridge Univ. Press, Cambridge, 1988).

7. W. C. Keel, Astrophys. J. **259**, L1 (1982).

8. Yu. V. Baryshev and Yu. L. Ezova, Astron. Zh. **74**, 497 (1997) [Astron. Rep. **41**, 436 (1997)].

9. M. P. Veron-Cetty and P. Veron, *A Catalogue of Quasars and Active Galactic Nuclei* (European Southern Observatory, Garching, 1998), ESO Scientific Report Series, Vol. 18.

10. G. Paturel, H. Andernach, L. Bottinelli, *et al.*, Astron. Astrophys., Suppl. Ser. **124**, 109 (1997).

11. Yu. V. Baryshev, A. A. Raikov, and A. A. Tron, in *Gravitational Lenses in the Universe*, 31st Liege International Astrophysics Colloquium, 1993, p. 365.

12. Yu. V. Baryshev, F. Sylos-Labini, M. Montuori, and L. Pietronero, Vistas Astron. **38**, 419 (1994).

13. F. Sylos-Labini, M. Montuori, and L. Pietronero, Phys. Rep. **293**, 61 (1998).

14. W. E. Harris, Astron. J. **112**, 1487 (1996); http-//wnw. physics.memaster.ca/Globular-html.

15. J. Blakeslee, http://babbage.sissa.it/astro-ph/9906356 (1999).

16. J. P. Blakesslee and J. L. Tonry, Astrophys. J. **442**, 579 (1995).

17. A. F. Dravskikh, Astron. Zh. **73**, 19 (1996) [Astron. Rep. **40**, 13 (1996)].

Translated by D. Gabuzda

Identification of Sources of Ultrahigh Energy Cosmic Rays

A. V. Uryson

Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia

Received February 23, 2000

Abstract—The arrival directions of extensive air showers with energies $4 \times 10^{19} < E \le 3 \times 10^{20}$ eV detected by the AGASA, Yakutsk, Haverah Park, and Fly's Eye arrays are analyzed in order to identify possible sources of cosmic rays with these energies. We searched for active galactic nuclei, radio galaxies, and X-ray pulsars within 3-error boxes around the shower-arrival directions and calculated the probabilities of objects being in the 3 error boxes by chance. These probabilities are small in the case of Seyfert galaxies with redshifts $z < 0.01$ and BL Lac objects, corresponding to $P > 3\sigma$ (σ is the parameter of Gaussian distribution). The Seyfert galaxies are characterized by moderate luminosities ($L < 10^{46}$ erg/s) and weak radio and X-ray emission. We also analyzed gamma-ray emission at energies $E > 10^{14}$ eV recorded by the Bolivian and Tian Shan arrays. The source identifications suggest that the gamma rays could have been produced in interactions of cosmic rays with the microwave background radiation and subsequent electromagnetic cascades in intergalactic space. We estimate the strength of intergalactic magnetic fields outside galaxy clusters to be $B \le 8.7 \times 10^{-10}$ G. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Particles initiating extensive air showers with energies $E > 4 \times 10^{19}$ eV are likely to have an extragalactic origin [1–5]. In this case, the spectrum of extragalactic cosmic rays (CRs) may abruptly steepen at $E \approx 3 \times 10^{19}$ eV, since their interaction with the microwave background radiation reduces the particle flux at energies $E \approx 6 \times 10^{19}$ eV to half the level implied by extrapolation of the power law [6, 7]. No such blackbody cutoff should be observed, however, if CR sources are primarily relatively nearby objects: the mean free path of particles with energies $E < 10^{20}$ eV in the background radiation field is about ~40–50 Mpc, and particles with energies up to $E \approx 10^{21}$ eV should traverse distances of about 10–15 Mpc [8] essentially unattenuated.

Currently, more than ten showers with energies $E \ge 10^{20}$ eV have been recorded. The very detection of such energetic particles rules out the presence of a blackbody cutoff in the spectrum at energies up to $E \approx 10^{21}$ eV. However, the spectrum has a rather complex shape at energies 10^{17}–10^{20} eV, which could be interpreted as a cutoff [3].

The following distances have been estimated for potential sources of CRs with $E > 4 \times 10^{19}$ eV. Analysis of the compositions and spectra of CRs obtained in various studies yielded the distance estimate ~50 Mpc [9]; about 30% of detected particles come from the plane of the Local Supercluster, from distances ~15–30 Mpc [5]; Fe nuclei can arrive from distances of ~100 Mpc [10].

CR sources considered in the literature can be divided into three categories. The first includes astrophysical objects, such as pulsars, active galactic nuclei [1], the hot spots and cocoons of powerful radio galaxies and quasars [11], and interacting galaxies [12]. The second category of proposed CR sources is topological defects [13, 14], and the third is the decay of supermassive metastable particles of cold dark matter that have accumulated in galactic halos [15, 16]. Direct identification of astrophysical objects is possible only in the first case. In the second case, any objects falling within 3 error boxes centered on particle arrival directions should be chance coincidences. In the third case, most of the particle flux should arrive from the Galactic halo and, perhaps, partly from the Virgo cluster of galaxies [16].

We adopted the first model (CR sources are astrophysical objects) as our working hypothesis and supposed that possible sources could be X-ray pulsars (i.e., the most powerful ones), Seyfert galaxies, BL Lac objects, or radio galaxies. We further assumed that particle trajectories in intergalactic space are almost rectilinear and ignored particle deflection by magnetic fields in the Galaxy. We searched for possible sources within 3-error boxes centered on the arrival directions of showers, and calculated the probabilities of objects being in the error boxes by chance.

This work continues a series of previous papers [17–21], in which we searched for CR sources using the catalog [22]. We analyzed 17 AGASA air showers with $E > 3.2 \times 10^{19}$ eV [23, 24], 12 Yakutsk showers with $E > 4 \times 10^{19}$ eV [25], two Haverah Park showers with $E \ge 10^{20}$ eV [26], and one shower with $E \approx 3 \times 10^{20}$ eV detected by the Fly's Eye array [27]. In the current paper, we analyze a more extensive sample of AGASA showers, including a total of 48 events with

$E > 4 \times 10^{19}$ eV [28], and use the catalog [29] to search for the sources. As before, we find a small probability that the Seyfert galaxies with $z < 0.01$ lying in the 3-error boxes are there by chance, corresponding to $P \approx 1.40 \times 10^{-3} \approx 3.20\sigma$, where σ is the Gaussian parameter. The radio galaxies and X-ray pulsars found in the 3-error boxes appear to be chance coincidences. We obtained a new result for BL Lac objects: the probability that these objects are in the 3-error boxes by chance is $P \approx 1.7 \times 10^{-3} \approx 3.10\sigma$, corresponding to 3.10σ, in contrast to our previous estimate $P < 3\sigma$ for these objects [21]. This discrepancy can be attributed to the fact that the new catalog [29] contains three times more BL Lac objects with declinations $\delta > -10°$ than the old catalog [22] (159 as opposed to 55).

In addition to showers with $E > 4 \times 10^{19}$ eV, we consider high-energy gamma-ray emission detected by the Bolivian and Tian Shan high-altitude arrays.

We showed in [20] based on an identification of CR sources that the gamma-ray photons that initiated the showers could have originated in electromagnetic cascades in intergalactic space, resulting from interactions of extragalactic protons with the cosmic background radiation. The possible development of such cascades with microwave background photons was demonstrated theoretically in 1966–1970 [30, 31]. Here, we give further arguments in favor of an extragalactic origin for gamma rays with $E > 10^{14}$ eV. We also estimate the intensity of the intergalactic magnetic fields (outside clusters of galaxies) implied by the reported source identification.

2. SOURCE IDENTIFICATION

2.1. Experimental Data

The arrival directions of showers with energies $4 \times 10^{19} < E \leq 3 \times 10^{20}$ eV and their errors are reported in [25–28]. For the purposes of source identification, we used only showers for which the error in their arrival direction did not exceed $(\Delta\alpha, \Delta\delta) \lesssim 3°$ in equatorial coordinates. Our data set includes a total of 53 showers: 48 showers recorded by the AGASA array [28], one detected by the Haverah Park array [26] (its error was computed in [32]), and four Yakutsk showers [25] (errors computed in [21]).

We compared the arrival directions of these showers with the coordinates of active galactic nuclei, radio galaxies, and X-ray pulsars. The optical coordinates of the galaxies and pulsars are accurate to several arcseconds, so that the fields of search were determined solely by the errors in the shower coordinates. Statistically [33], the probability that the proton initiating the shower is within the one- and three-mean-square error boxes is 66 and 99.8%, respectively. We therefore searched for sources within $(3\Delta\alpha, 3\Delta\delta) \leq 9°$ fields around the shower-arrival directions.

2.2. Search for Seyfert Galaxies

We searched for Seyfert galaxies using the catalog [29]. As a rule, more than one galaxy falls in the 3-error box of each shower. In each 3-error box, we looked for galaxies with redshifts $z \leq 0.0092$, i.e., for those located within 40 Mpc of the Sun ($H = 75$ km s^{-1} Mpc^{-1}). We subdivided the showers into several groups, depending on Galactic latitude b and arrival direction. The number of shower groups K and the number of showers N with at least one $z \leq 0.0092$ galaxy falling in the 3-error box were $K = 48$ and $N = 27$ for $|b| > 11°$; $K = 33$ and $N = 22$ for $|b| > 21°$; $K = 26$ and $N = 22$ for $|b| > 31°$; $K = 23$ and $N = 21$ for $|b| > 41°$; and $K = 53$ and $N = 27$ with no prior selection in Galactic latitude and without constraining b. We subdivided the showers in Galactic latitude b in order to exclude events clearly lying in the galaxy "avoidance zone." We calculated the probabilities that at least one galaxy with the given z would fall in the fields of search of N of the total of K showers by chance. We analyzed groups of simulated showers arriving from areas of sky with $|b| > 11°$, $21°$, $31°$, and $41°$, and from all Galactic latitudes with randomly distributed arrival coordinates, with the numbers K of showers in these groups equal to those observed. The coordinates of the simulated showers in each group were determined by a random-number generator [34] within a survey band $\alpha = 0–24$ hr and $\delta = -10°–90°$. We then counted in each simulated group the number of showers $N_{sim} = N$ having at least one galaxy with $z \leq 0.025$, 0.017, and 0.0092 located in the $(3\Delta\alpha, 3\Delta\delta)$ error box (N_{sim} can take values in the interval $0 \leq N_{sim} \leq K$). In a group of K showers, the probability P of a chance occurrence of galaxies in the field of search of a given number of showers N_{sim} was determined as $P = \sum_{i=1}^{M} (N_{sim})_i / M$, where $M = 10^5$ is the number of trials performed for each group.

The resulting probabilities P for $z \leq 0.0092$ were $P(|b| > 11°) \approx 2.84 \times 10^{-3}$, $P(|b| > 21°) \approx 1.85 \times 10^{-3}$, $P(|b| > 31°) \approx 1.40 \times 10^{-3}$, $P(|b| > 41°) \approx 5.98 \times 10^{-3}$, and $P(|b| > 0) \approx 2.25 \times 10^{-3}$. These probabilities increase with increasing z; for example, the probabilities for $|b| > 31°$ and $z \leq 0.025$ and 0.017 are 1.27×10^{-1} and 3.15×10^{-2}, respectively (see [17] for $P(z)$ relations for showers arriving from sky areas located at arbitrary b). These results indicate that, as the Galactic latitude $|b|$ increases, chance occurrences of galaxies with $z \leq 0.0092$ in the error boxes of the given number of showers become very unlikely. The large probability for a chance occurrence of galaxies in the error boxes of showers with $|b| > 41°$ may be due to an insufficient data set.

2.3. Search for BL Lac Objects

Kardashev [35] was the first to suggest that BL Lac objects could be sources of ultrahight-energy CRs. A search for BL Lac objects within the 3-error boxes of

the shower-arrival directions yielded the following results: $|b| > 11°$ $K = 48$, $N = 39$, and $P = 2.56 \times 10^{-2}$; for $|b| > 21°$ $K = 33$, $N = 28$, and $P = 2.93 \times 10^{-2}$; for $|b| > 31°$ $K = N = 26$, and $P = 1.71 \times 10^{-3}$; for $|b| > 41°$ $K = N = 23$, and $P = 1.56 \times 10^{-2}$; and for b and $K = 53$, $N = 41$, and $P = 4.28 \times 10^{-3}$ with no selection in b. The law of probability indicates that the BL Lac objects in the fields of search at $|b| > 31°$ are not chance coincidences. Along with the Seyfert galaxies, these should be considered possible CR sources.

2.4. Search for X-Ray Pulsars

We searched for X-ray pulsars using the catalog [36]. The probability of the pulsars in the shower 3-error boxes falling there by chance is 0.17, consistent with these pulsars being chance coincidences.

2.5. Search for Radio Galaxies

Powerful radio galaxies are considered to be possible CR sources due to their high energies. In some models [11], particles in radio galaxies can be accelerated to energies of $5 \times 10^{19}–10^{20}$ eV. However, radio galaxies are also considered to be the most likely sources for particles with energies 2×10^{20} and 3×10^{20} eV. Stanev *et al.* [4] and Stecker [10] report attempts to identify these galaxies under various assumptions about the intensity of the intergalactic magnetic field and the atomic numbers and propagation distances of the particles. Rachen *et al.* [9] also analyzed the chemical composition and spectrum of detected particles assuming that their sources were radio galaxies.

We searched for radio-galaxies using the catalogs [37, 38]. Our identification list included all galaxies from [37] and galaxies with $z \leq 0.1$ from [38] that were absent from [37]. The probabilities of the radio galaxies with any z and $z < 0.1$ falling in the error boxes of showers by chance are $P \approx 0.01–0.1$, for showers arriving from fields of sky with various b values. Thus, the radio galaxies in the cosmic-ray fields of search could represent chance coincidences.

2.6. Analysis of Showers with $E \geq 10^{20}$ eV

We analyzed a total of 11 showers with energies $E \geq 10^{20}$ eV, including seven showers from [28], two from [26] (see [32] for their error boxes), one from [25], and one from [27]. Probability estimates for these showers are unsuitable for our purposes due to the small data set and large errors in the arrival directions of some of the particles. We therefore checked whether moderate Seyfert galaxies located within 100 Mpc were found in the 3-error boxes of the arrival directions of these particles. No such galaxies were found in the error boxes for two events—one Yakutsk and one AGASA shower—however, these were located in the galaxy avoidance zone, with arrival directions at Galactic latitudes $b = 3 \pm 2°$

and $-4.8°$, respectively. We found Seyfert galaxies in the 3-error boxes of each of the remaining nine showers: galaxies with $z \leq 0.0092$ for seven showers and with $z = 0.016$ and $z = 0.018$ (i.e. located within 72 Mpc) for two showers. This result is consistent with the hypothesis that showers could be initiated by galactic nuclei (the chemical composition of CRs with energies $E \geq 10^{20}$ eV remains unknown). Of these nine showers, seven had BL Lac objects located in their fields of search, most with redshifts $z > 0.1$ or without published redshifts. However, this result does not contradict the CR propagation pattern for the following reasons. If CRs in BL Lac objects can be accelerated to energies $E \sim 10^{27}Z$ eV [35], the particles could arrive in the form of fragments of atomic nuclei with $E \geq 10^{20}$ eV from distances much greater than 100 Mpc. The lack of redshift determinations could be due to a lack of absorbing material along the line of sight, which seems more likely for objects located within ~100 Mpc than for more distant galaxies. Therefore, BL Lac objects with unknown z are likely to be located at distances of ~100 Mpc [39].

2.7. Errors in the Shower-Arrival Directions

In the previous sections, we searched for sources within circles of radius $(3\Delta\alpha, 3\Delta\delta) = 9°$. These areas of search corresponded to shower-arrival direction errors of $(\Delta\alpha, \Delta\delta) = 3°$. We now analyze the dependence of source identification on the coordinate errors.

For showers arriving from areas of sky with $b > 31°$, we determined the probabilities of Seyfert galaxies with $z \leq 0.0092$ and BL Lac objects falling in $(3\Delta\alpha, 3\Delta\delta) = 12°$, $15°$, and $18°$ 3-error boxes by chance. In this sequence, the numbers of showers with at least one object in the search fields increases. The resulting probabilities for Seyfert galaxies are $P(12°) \approx 3.65 \times 10^{-3}$, $P(15°) \approx 9.80 \times 10^{-2}$, and $P(18°) \approx 2.84 \times 10^{-1}$. The corresponding probabilities for BL Lac objects are $P(12°) \approx 5.52 \times 10^{-2}$, $P(15°) \approx 2.40 \times 10^{-1}$, and $P(18°) \approx 5.14 \times 10^{-1}$. Therefore, in cases where data are available only for several dozen showers, only events with arrival direction errors $(\Delta\alpha, \Delta\delta) < 4°$ should be used for identification. For example, the coordinate errors for the shower with energy $E \approx 3 \times 10^{20}$ eV of [27] ($\alpha = 85.2 \pm 0.4°$, $\delta = (48.0^{+5.2}_{-6.3})°$) make this event unsuitable for inclusion in the sample considered here.

2.8. Intergalactic Magnetic Fields

When identifying CR sources, we assumed that the corresponding particles propagate almost rectilinearly through intergalactic space; the searches for possible sources were carried out in circles of radius $(3\Delta\alpha, 3\Delta\delta) = 9°$ centered on the shower axes. This is equivalent to assuming that the maximum deflection angle in the intergalactic magnetic field is $\alpha_0 = 9°$.
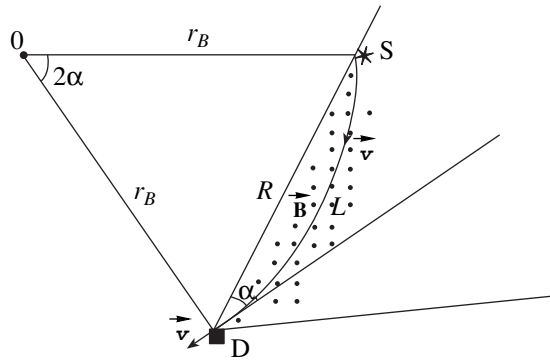
Fig. 1. Motion of a particle in the plane perpendicular to the intergalactic magnetic field B. S is the source, D the detector, O the center of the Larmor orbit, L the arc along which the particle moves, and α the angle between the direction toward the source and the shower axis (the direction of particle motion during detection).

At present we know neither the magnitude nor the character of inhomogeneities in the intergalactic magnetic field. Only theoretical upper limits for the magnetic-field intensity are available [40]: $B < 10^{-9}$ G, based on measurements of the rotation measures of quasars with $z = 2.5$; $B \ll 10^{-9}$ G, if ultra-high-energy protons propagate rectilinearly through intergalactic space; and $B < 10^{-11}$ G for the regular component of the magnetic field. Clusters of galaxies can harbor much stronger fields: $B \sim 10^{-6}$–10^{-7} G at distances of up to ~0.5 Mpc from the cluster center.

We estimated the intensity of the intergalactic magnetic field outside clusters of galaxies for the simple case when particles move in the plane perpendicular to the magnetic field B, assuming that $\alpha \approx 9°$ and that field inhomogeneities can be neglected throughout the propagation path. We also ignored the effect of the magnetic field of our own Galaxy.

We looked for sources located in a given cone around the shower axis (Fig. 1). The half-angle α of the cone is related to the arc $L*$ along which the particle moved from the source to the detection facility: $\alpha = L*/2$. The length of the arc is $L = 2\alpha\, r_B$, where the Larmor radius $r_B = E/(300ZB)$ [1]. Here, r_B is in cm, B in G, and Z is the charge of the particle. Since $\alpha = \alpha_0$, we have $L/(2\,r_B) = \alpha_0$ and $B = (2\alpha_0 E)/(300ZR)$ for angles $\alpha \leq 10°$; $L \approx R$ where R is the distance between the source and detector array. It follows that $B = (2\alpha_0 E)/(300ZL)$ and $\alpha_0 = 90°$. We found above that most CR sources are located at distances $R < 40$ Mpc. Therefore, we have $B \leq 8.7 \times 10^{-10}$ G for protons ($Z = 1$) with energies $E \approx 10^{20}$ eV. If heavy nuclei with such energies arrive from distances of 70 Mpc [10], we have for $Z = 26$, $B \leq 2.7 \times 10^{-11}$ G. These limits for the magnetic-field intensity B are consistent with the estimates of [40].

We now consider the case when a CR source is a member of a cluster. According to [41], the sizes of gal-

axy clusters range from $D \approx 1$ Mpc to $D \approx 5$–7 Mpc. If the distance to the cluster is 40 Mpc, its angular size is $\alpha_{cl} \approx 1.4°$ and $\alpha_{cl} \approx 10°$ for $D \approx 1$ Mpc and $D \approx 7$ Mpc, respectively. In both cases, the cluster containing the source should fall in the 3-error box, since the angular size of the box exceeds that of the cluster, $2\alpha_0 \approx 18° > \alpha_{cl}$.

Thus, the proposed source identification method and the results obtained remain valid if the magnetic fields outside clusters are $B \ll 10^{-9}$ G. In this case, we should search for sources within $(3\Delta\alpha, 3\Delta\delta) \approx 9°$ and areas of showers with arrival direction errors $(\Delta\alpha, \Delta\delta) < 3°$.

2.9. Seyfert Galaxies that are CR Source Candidates

In this section, we discuss the conditions required in order for accelerated particles to escape from their sources, and describe the characteristics of the proposed source galaxies. We adopted the radio and X-ray fluxes and galaxy axial ratios b/a from [42].

The regions within 10–100 pc of the nuclei are believed to be surrounded by geometrically and optically thick dusty tori that emit in the infrared [43]. Infrared radiation from the region inside the torus is absorbed and re-emitted by clouds. Photopion losses by accelerated particles in this region are small if the source luminosity $L < 10^{46}$ erg/s [11]. It is such Seyfert galaxies that are CR sources. Interactions of particles with infrared photons produce electromagnetic cascades and X-ray radiation [44], implying that galactic nuclei in which particles are first accelerated and then lose their energy should be X-ray sources. The galaxies that are possible cosmic-ray sources are also weak X-ray sources.

An accelerated particle does not come into the dusty torus if the angle between its velocity and the perpendicular to the galactic plane is such that $\tan i < l/h$, where l and h are the inner radius and thickness of the torus, respectively, i.e., if the galactic plane is viewed at a relatively small angle (if the torus orientation correlates with the orientation of the galactic plane). The angle i depends on the galactic axial ratio: $\cos i = b/a$ ($i = 55°$ for $b/a = 0.6$) [45]. Therefore, galaxies that are CR sources should have relatively high b/a ratios.

Accelerated particles will interact with galactic gas and dust. According to [41], the density of matter (equivalent to the H density) in Seyfert galaxies is $n \sim 10^8$–10^{11} cm^{-3} in the central $D \sim 0.1$–1 pc; $n \sim 10^4$–10^6 cm^{-3} at $D = 100$ pc; and $n \sim 30$–600 cm^{-3} in the toroidal dusty layer $D \sim 100$ pc thick at a distance of ~100 pc from the central region. The mean density averaged over the entire galaxy is $n \approx 10$ cm^{-3}. Below, we calculate the free paths of accelerated protons. (The free paths of nuclei are shorter by a factor of a few.)

The cross section for proton interactions with H atoms is equivalent to that for pp interactions, which increases logarithmically with the proton energy and at $E \approx 10^{21}$ eV and may be $\sigma_{pp} \approx 200$ mbarn, according to the estimates of

[46]. The proton-interaction cross section is $\lambda = 1/(n\sigma_{pp})$. It follows that $\lambda \sim 10^{-5}$–0.02 pc, $\lambda \sim 1.5$–150 pc, and $\lambda \sim 3$–60 kpc in the central region, at $D = 100$ pc, and in the $D \sim 100$ pc thick toroidal dusty layer, respectively; the mean free path for the entire galaxy is $\lambda \approx 150$ kpc. The proton interaction mean free path is $\lambda \gg D$ throughout the galaxy, except for the central region. An accelerated particle leaves the central region without large energy losses, since the gas there is distributed nonuniformly and is concentrated in clouds with filling factors $\eta \ll 0.1$ [1, 11].

Thus, the main characteristic features of galaxies that are cosmic-ray sources should be relatively small distances ($z \leq 0.01$), weak X-ray and radio emission, and relatively large axial ratios b/a.

The shower 3-error boxes often contain more than one such galaxy. It is indeed more likely to detect CRs toward directions where there are more such galaxies, and this statistical sample provides further support for the hypothesis advanced here.

Our results suggest that the sources of detected CRs are located at distances $R < 40$ Mpc, and that the number of sources at greater distances decreases. This could be due to two effects. First, intergalactic and intercluster voids can have sizes up to ~100 and ~250 Mpc, respectively [47], so that we observe mainly particles arriving from the Local Supercluster, from distances $R \sim 30$ Mpc. Second, it was believed earlier that the number of sources should increase with distance due to evolutionary effects, since the luminosity and density of sources in a comoving volume increases with redshift [1]. However, the source energy seems to be uncorrelated with the efficiency of particle generation. The most powerful galaxies appear not to be identified with CR sources. On the contrary, CR sources are identified with intermediate-luminosity galaxies that are weak radio and X-ray sources: the more luminous the galaxy, the more difficult it is for protons to escape, due to energy losses through interactions of the accelerated protons with infrared photons.

3. ELECTROMAGNETIC CASCADES IN INTERGALACTIC SPACE

One manifestation of proton interactions with the microwave background radiation could be a blackbody cutoff in the CR spectrum. However, if cosmic-ray sources are located within several tens of Mpc of the Earth, their energy spectrum should not exhibit such a cutoff. In our earlier analysis [18, 19] of the spectrum obtained for data from the giant ground-based facilities of Sydney, Yakustk, Akeno and AGASA, Haverah Park, and Fly's Eye [48], we found no apparent cutoff.

In this section, we show that one possible manifestation of the interaction of protons with the microwave background radiation could be the very-high-energy gamma rays detected by ground-based arrays.

### 3.1. Very-High-Energy Gamma Rays and Attempts at Interpretation

Showers initiated by gamma rays have been studied with the Tian Shan and Bolivian high-altitude arrays [49–52]. Such events can be distinguished from the intense background of numerous cosmic-ray showers by their substantially smaller fractions of muons and hadrons. This criterion was used to select out gamma-ray initiated showers observed by the Bolivian facility. However, a methodical analysis performed at the Tian Shan array [49, 53] showed that using this criterion alone could result in the inclusion of an appreciable number of particle-initiated showers. The main criterion used at the Tian Shan facility is that showers not only have low muon and hadron contents, but also that the energy flux in the hadron component be low compared to that in the electron and photon components. Due to the use of different criteria, the showers selected at the different facilities have different mean age parameters s. Note that, at mountain altitudes, a considerable fraction of showers initiated by gamma rays with $E > 10^{14}$ eV have ages $s < 1$ [54, 55]. The Tian Shan showers discussed here have $s < 1$.

The photons detected by the Tian Shan and Bolivian arrays had energies of $(5$–$8.5) \times 10^{14}$ and $(10^{14}$–$10^{17})$ eV, respectively. The shower-arrival directions did not coincide with any Galactic gamma-ray sources. The measured fluxes exceeded theoretical estimates for the diffuse gamma-ray flux due to interactions of CRs with interstellar gas by more than a factor of ten [1, 56]. The calculation results could not be brought into agreement with the measurements by varying the models or by allowing for a possible change in the CR composition at energies $E > 1$ TeV/nucleon [57, 58]. Neither could this radiation be explained as coming from unresolved Galactic sources [1], as suggested in [59, 60].

3.2. Sky Distribution of Showers Initiated by Gamma Rays

The Tian Shan showers were distributed nonuniformly on the sky: the coordinates (α, δ) of five of eight gamma-ray photons coincide within 3-error boxes (Fig. 2). We used the following methods to determine the probability P of a chance coincidence of the coordinates of five of the eight gamma-ray photons. First, the gamma-ray photons came from the band $\alpha = 13$–16^h. The probability of finding by chance five photons of eight in a band with $\Delta\alpha = 3^h$ is rather small: $P = C_8^5(\Delta\alpha/24)^5(1 - \Delta\alpha/24)^3 \approx 1.1 \times 10^{-3}$ ($\approx 3.3\sigma$). Second, the 3-σ arrival directions of the photons occupy $\Delta S \approx 0.19$ of the facility's field of view. The probability of finding five of the eight photons within this area by chance is also fairly low, $P = C_8^5(\Delta S)^5(1 - \Delta S)^3 \approx 7 \times 10^{-3}$ ($\approx 2.7\sigma$).

These five photons arrived from Galactic latitudes $b \geq 30°$. This means that they are unlikely to be of diffuse origin, since the diffuse flux in this direction is
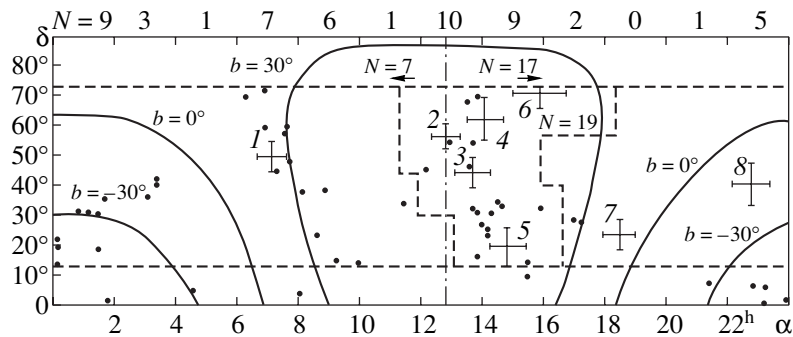
Fig. 2. Sky distributions of showers initiated by gamma rays (numbered bar crossings showing position errors) [49] and of $B < 15^m$ Seyfert galaxies (dots) from [61]. The solid lines show the loci of Galactic latitudes $b = \pm30°$ and the Galactic plane $b = 0°$. The dotted lines show the field of view for the Tian Shan array and the positional 3-error boxes for the arrival directions of showers 2–6. The dashed–dotted line $\alpha = 12^h49^m$ is the symmetry axis of the region $b > 30°$. Also shown is the number of galaxies N in bands $\Delta\alpha = 2^h$ in the region of showers 2–6 and in the region $b > 30°$ to the right and to the left of the symmetry axis.

lower than along the Galactic disk [1], and most diffuse photons are expected to originate at lower latitudes. The Bolivian array recorded an excess of showers (at the 3.8σ level) initiated by gamma rays with energies above 10^{14} eV in the region $\alpha = 180°–210°$, $\delta = 0$ to $(-40°)$. This region is also located at high Galactic latitudes, $b < -25°$ (no sky coordinate distribution was published for these events). The high Galactic latitudes of the arrival directions may indicate an extragalactic origin for the gamma rays, which could have been produced by interactions of extragalactic protons with the microwave background radiation, followed by electromagnetic cascades in intergalactic space. The theoretical possibility of such cascades was analyzed by Hayakawa [30] and Prilutsky and Rozental [31].

3.3. Distribution of Possible Sources of Ultrahigh-Energy Cosmic Rays

If the above conclusions are valid, the distributions of gamma rays and of CR sources in the sky should be correlated. Accordingly, we compared the distributions of gamma rays and of supposed CR sources. The numbers of galaxies can differ from one region of the sky to another due to differing observing conditions. Therefore, we compared the distribution of gamma rays with that of galaxies of the Byurakan survey [61] with $B < 15^m$. The list of such galaxies is relatively complete [62] in the sense that differences between the sky densities of galaxies in different areas are real, rather than the result of observational selection effects.

To search for possible deviations of the galaxy distribution from uniformity, we subdivided the region with $b > 30°$ into two equal areas at the $\alpha = 12^h49^m$ line and counted the number of galaxies N in each of these areas. The line at $\alpha = 12^h49^m$ is the symmetry axis of the region $b > 30°$ in (α, δ) coordinates. In addition, we counted the numbers of galaxies in $\Delta\alpha = 2^h$ bands and in the 3-error boxes for the arrival directions of five showers (nos. 2–6).

Figure 2 presents the distribution of the $B < 15^m$ galaxies from [61]. This figure shows that the region of showers 2–6 coincides with areas of the sky in which the numbers of galaxies are maximum: 17 galaxies fall in the region of showers 2–6, 25 galaxies fall in the array survey band at $b = 30°$, and a total of 52 galaxies fall in the entire survey band. The area occupied by showers 2–6 is ≈0.52, equal to ≈0.19 of the survey-band areas, while the fraction of galaxies falling in this region is 0.64 and 0.34, respectively.[1]

At $b > 30°$, the numbers of galaxies with $z < 0.1$ to the right ($N = 14$) and to the left ($N = 8$) of the $\alpha = 12^h49^m$ line are not the same. However, the number of galaxies in the region of showers 2–6 is equal to the number of galaxies in areas of sky at $b > 30°$ that are free of gamma rays.

We did not analyze the sky distribution of BL Lac objects, since we did not have even a minimally complete list of such objects.

3.4. Gamma-Ray Generation

In [20], we analyzed the development of electromagnetic cascades, and the distances from the sources at which gamma rays with energies $(5–8.5) \times 10^{14}$ eV form in these cascades. The mean free paths of such photons in the field of the microwave background radiation are 10–20 kpc [63], so that, in order to be detected on the Earth, the gamma rays must originate at the edge of our Galaxy. We obtained numerical estimates for electromagnetic cascades initiated by protons with energies $E = 6 \times 10^{19}$, 3×10^{20}, and 10^{22} eV. The main reactions in the cascade are interactions between protons and microwave

---

[1] Here, we consider galaxies at distances ~100 Mpc. The mean redshift and distance to these galaxies are $\langle z \rangle = 0.0283$ and $\langle R \rangle = 113.2$ Mpc, respectively. The nearest galaxy has $z_{min} = 0.0069$ and $R_{min} = 27.6$ Mpc. The most distant galaxy has $z_{max} = 0.0797$ and $R_{max} = 318.8$ Mpc. The galaxies in the region of showers 2–6 have $\langle z \rangle = 0.0347$ and $\langle R \rangle = 139$ Mpc.

background photons $p + \gamma_{rel} \longrightarrow p + \pi^{\circ}$ (I); $p + \gamma_{rel} \longrightarrow p + \pi^{+}$ (II); decay of pions $\pi^0 \longrightarrow 2\gamma$ (III), $\pi^{+} \longrightarrow \mu^{+} + \nu$ (IV) and muons $\mu^{+} \longrightarrow e^{+} + \nu + \nu$ (V); scattering of positrons by microwave background photons and, at energies $\varepsilon_e > 10^{15}$ eV, by radio photons $e + \gamma_b \longrightarrow e' + \gamma'$ (VI); and the production of electron–positron pairs by photons (III) and (VI) $(e^{+} + e^{-})$ pairs in the field of the background photons $\gamma + \gamma_b \longrightarrow e^{+} + e^{-}$ (VII). We adopted the cross sections and kinematic parameters of these processes from [64, 65]. The resulting estimates $\langle R \rangle \approx 200$ Mpc agree with the distances to Seyfert galaxies [61] located within the arrival areas of gamma rays initiated by showers 2–6 ($H = 75$ km s^{-1} Mpc^{-1}). These estimates also indicate that gamma rays with energies $(5–8.5) \times 10^{14}$ eV can be produced in the above reaction chains at distances $\sim(1–2) \times 10^3$ Mpc from the source.

Due to fluctuations, electromagnetic cascades can develop at smaller than average lengths, and photons with energies $(5–8.5) \times 10^{14}$ eV can be produced at distances $\langle R \rangle = 1.5$ Mpc from the sources. See [66] for electromagnetic cascade schemes in which such photons are produced at distances $R \leq 10$ Mpc from the sources.

In order for the intergalactic magnetic fields not to prevent the development of electromagnetic cascades, the synchrotron losses of electrons must be smaller than their Compton losses. The formulas for these losses are given in [67]. Based on this condition, we estimated in [20] $B < 2 \times 10^{-8}$ and $B < 2 \times 10^{-9}$ G for Compton scattering on microwave-background and radio photons, respectively.

In Section 1.8, we found that $B \leq 8.7 \times 10^{-10}$ G, so that the intergalactic magnetic field should not hinder the development of electromagnetic cascades.[2]

3.5. Detection of Cascades with Energies $E < 10^{14}$ eV

During its development, a cascade produces several tracks in the directions of the primary particle and of gamma rays produced in connection with π^0 decays and $e^{+}e^{-}$ pairs, and also a large number of branches along the directions of the motion of electrons and positrons with energies $E_e < 10^{14}$ eV scattered by microwave background photons. The number of branches for $E = 10^{21}$ eV can be roughly estimated as $N = E/E_e = 10^7$. The integrated intensity of ultrahigh-energy CRs [68] yields estimates of their density $\rho_p = 3 \times 10^{-20} (4\rho/c) = 10^{-29}$ cm^{-3}. The density of the branches is then $\rho_{br} = 10^{-22}$ cm^{-3}. As a

result of inverse Compton scattering on microwave background photons, electrons produce gamma rays [66] with a mean energy of $\approx 3.6 \times 10^{-16} E_e^2$, whose mean free path is ~ 1 kpc. In the observer's frame, the photon scattering angle θ_{IC} is determined from the condition $\tan \theta_{IC} = \theta_{IC}^* / (2\Gamma)$, where θ_{IC}^* is the photon scattering angle in a frame comoving with the electron; $\Gamma = E_e/m_e$ is the Lorentz factor of the electron ($m_e \approx 0.5$ MeV is the electron mass). Photons with energies below 10^{14} eV are scattered nearly isotropically in a frame comoving with the electron [69], so that the transverse distance between two photons with energy ~ 10 MeV successively scattered over a length of $l = 1$ kpc is ~ 0.2 kpc. The distance d between the tracks after passage through $l = 1$ kpc is the following. If a π° had an energy of $E_\pi \approx 10^{20}$ eV, then the angle of divergence between two produced photons is $\theta_\pi \sim m_\pi/E_\pi \approx 1.4 \times 10^{-12}$ ($m_\pi \approx 140$ MeV is the proton mass), and the distance between the tracks produced by these photons is $d = l\theta_\pi \sim 4 \times 10^{10}$ cm. The angle divergence between the particles of an $e^{+}e^{-}$ pair with energy $E_e \approx 10^{19}$ eV is $\theta_e \sim m_e/E_e \approx 5 \times 10^{-14}$, and the distance between the two tracks formed is $d = l\theta_e \approx 1.5 \times 10^8$ cm. It follows that the density of cascade photons in the energy interval accessible to satellite instruments is low, preventing the detection of such photons with space facilities. The extensive air showers initiated by gamma rays provide the only tool for investigating intergalactic cascades.

4. CONCLUSIONS

We have analyzed the arrival directions of extensive air showers with energies $4 \times 10^{19} < E \leq 3 \times 10^{20}$ eV and of showers initiated by gamma rays with $E > 10^{14}$ eV, searching for possible CR sources, and assuming that showers with $E > 4 \times 10^{19}$ eV have an extragalactic origin [1–5]. Searches for active galactic nuclei, radio galaxies, and X-ray pulsars were carried out in three-error boxes around the shower-arrival directions. We computed the probabilities of objects of these types being located in the 3-error boxes by chance, and found these probabilities to be small: $P \approx (3.1–3.2)\sigma$ for Seyfert galaxies with redshifts $z < 0.01$ (i.e., located within 40 Mpc of the Earth, with $H = 75$ km s^{-1} Mpc^{-1}) and for BL Lac objects. The Seyfert galaxies have luminosities $L < 10^{46}$ erg/s, weak radio and X-ray emission, and relatively large galactic axial ratio a/b (if the dust-torus orientation correlates with the orientation of the galactic plane).

The identification of sources depends on the errors in the shower arrival directions. We looked for sources in $(3\Delta\alpha, 3\Delta\delta) \approx 9°$ areas centered on the particle arrival directions. If we consider a total of several dozen showers, only events with errors $(\Delta\alpha, \Delta\delta) < 4°$ should be selected for source identification. Based on the assumption that the particle paths deviate from linear trajectories by no

---

[2] In [20], we analyzed the gamma-ray flux from the viewpoint of the development of electromagnetic cascades in intergalactic space. Here, we also point out that, prior to comparing the Bolivian array flux with calculations [1], the observed flux must be averaged over the area of sky $\alpha = 0°–180°$, $\delta = 0°–90°$, since the radiation was detected from the area $\alpha = 180°$ to $210°$, $\delta = 0°$ to $-40°$.

more than $(3\Delta\alpha, 3\Delta\delta) \approx 9°$ as they pass through the intergalactic magnetic fields on their way from the source to the detector, we derived an upper limit for the intensity of the intergalactic magnetic field outside galaxy clusters, $B \leq 8.7 \times 10^{-10}$ G. If this estimate for the field intensity is valid, source searches should be conducted within $(3\Delta\alpha, 3\Delta\delta) \approx 9°$ areas also for showers with arrival direction errors $(\Delta\alpha, \Delta\delta) < 3°$.

Seyfert galaxies located within 70 Mpc of the Earth could be sources of particles with $E \geq 10^{20}$ eV. The cosmic-ray energy spectrum should exhibit no blackbody cutoff if the galaxies that are the CR sources are no farther than 40–70 Mpc. Our analysis of the energy spectrum obtained with giant ground-based arrays in [18, 19] showed no apparent blackbody cutoff.

When analyzing showers initiated by gamma rays, we searched for explanations for why photons arrive mainly from high Galactic latitudes $|b| > 30°$ and why the Tian Shan photons are distributed nonuniformly in the sky. These deviations from uniform distributions are consistent with the nonuniform distribution of $B < 15^m$ Seyfert galaxies from [61]. The list of galaxies with such magnitudes is relatively complete [62], and the average distance to these galaxies is 100 Mpc (radio galaxies do not exhibit such nonuniformity in their sky distribution). Based on these facts, we showed that the gamma rays that initiate extensive air showers with $E > 10^{14}$ eV could have been born from interactions between extragalactic CRs and microwave background photons, with subsequent electromagnetic cascades in intergalactic space. The possible development of such cascades was predicted in [30, 31, 64]. At present, the only way to study intergalactic cascades is via analyses of extensive air showers initiated by cosmic gamma rays.

The model for the generation of cosmic gamma rays suggested in this paper can explain the characteristics of gamma-ray emission with $E > 10^{14}$ eV, and suggests that Seyfert galaxies are cosmic-ray sources.

Should we expect that the distribution of showers initiated by particles with $E > 4 \times 10^{19}$ eV should also be nonuniform in the band $\Delta\alpha \approx 13–16^h$? This question remains unanswered, since available catalogs of Seyfert galaxies [29, 42] are incomplete, and the escape of accelerated particles from the source galaxies is hindered by interaction between the particles and infrared photons. As a result, not all relatively nearby Seyfert galaxies are CR sources, and the distribution of particles on the sky should instead reflect the distribution of Seyfert galaxies with favorable particle-escape conditions.

These results can be verified using new observations obtained with giant ground-based detecting facilities— the Yakutsk, AGASA, Fly's Eye, and Haverah Park arrays—by analyzing archive data of the Sydney and Volcano Ranch arrays, and using observations made with future giant facilities (both planned and under construction) intended for studies of showers with energies

$E \geq 10^{20}$ eV, which will have much higher angular (to 0.2°) and spectral resolution: the EAS-1000 Pierre Auger Project, new observations with the high-altitude Tian Shan and ANI arrays, and satellite-borne instruments.

ACKNOWLEDGMENTS

REFERENCES

1. V. S. Berezinsky, S. V. Bulanov, V. L. Ginzburg, *et al.*, *Astrophysics of Cosmic Rays*, Ed. by V. L. Ginzburg [in Russian] (Nauka, Moscow, 1990).

2. M. N. Dyakonov, T. A. Egorov, N. N. Efimov, *et al.*, *Ultrahigh-Energy Cosmic Radiation* [in Russian] (Nauka, Novosibirsk, 1991).

3. M. Hillas, Nature **395**, 15 (1998).

4. T. Stanev, P. L. Biermann, J. Lloyd-Evans, *et al.*, Phys. Rev. Lett. **75**, 3056 (1995).

5. N. Hayashida, K. Honda, M. Honda, *et al.*, Phys. Rev. Lett. **77**, 1000 (1996).

6. G. T. Zatsepin and V. A. Kuz'min, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 114 (1966) [JETP Lett. **4**, 78 (1966)].

7. K. Greisen, Phys. Rev. Lett. **16**, 748 (1966).

8. F. W. Stecker, Phys. Rev. Lett. **21**, 1016 (1968).

9. J. Rachen, T. Stanev, and P. Biermann, Astron. Astrophys. **273**, 377 (1993).

10. F. W. Stecker, Phys. Rev. Lett. **80**, 1816 (1998).

11. C. A. Norman, D. B. Melrose, and A. Achterberg, Astrophys. J. **454**, 60 (1995).

12. C. Cesarsky and V. Ptuskin, in *Proc. 23rd International Cosmic Ray Conference, Calgary, 1993*, Ed. by D. A. Leahy, R. B. Hicks, and D. Venkatesan, (World Scientific, Singapore, 1993), Vol. 2, p. 341.

13. C. T. Hill, D. N. Schramm, and T. P. Walker, Phys. Rev. D **36**, 1007 (1987).

14. V. Berezinsky and A. Vilenkin, Phys. Rev. Lett. **79**, 5202 (1997).

15. V. A. Kuz'min and V. A. Rubakov, Yad. Fiz. **61**, 1122 (1998) [Phys. At. Nucl. **61**, 1028 (1998)].

16. V. Berezinsky, M. Kachelrié, and A. Vilenkin, Phys. Rev. Lett. **79**, 4302 (1997).

17. A. V. Uryson, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 71 (1996) [JETP Lett. **64**, 77 (1996)].

18. A. V. Uryson, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 729 (1997) [JETP Lett. **65**, 763 (1997)].

19. A. V. Uryson, Zh. Éksp. Teor. Fiz. **113**, 12 (1998) [JETP **86**, 6 (1998)].

20. A. V. Uryson, Zh. Éksp. Teor. Fiz. **113**, 385 (1998) [JETP **86**, 213 (1998)].

21. A. V. Uryson, Zh. Éksp. Teor. Fiz. **116**, 1121 (1999) [JETP **89**, 597 (1999)].

22. M. P. Veron-Cetty and P. Veron, ESO Sci. Rep., No. 13 (1993).

23. N. Hayashida, K. Honda, M. Honda, *et al.*, in *Proc. 22nd International Cosmic Ray Conference, Dublin, 1991*, Vol. 2, p. 117.

24. N. Hayashida, K. Honda, M. Honda, *et al.*, Phys. Rev. Lett. **73**, 3491 (1994).

25. B. N. Afanasiev, M. N. Dyakonov, V. P. Egorova, *et al.*, in *International Symposium on Extremely High Energy Cosmic Rays. Astrophysics and Future Observatories, Tokyo, 1996*, Ed. by M. Nagano, p. 32.

26. A. Watson, *Particle and Nuclear Astrophysics and Cosmology in the Next Millenium*, Ed. by E. W. Kolb and R. D. Peccei (World Scientific, Singapore, 1995), p. 126.

27. D. Bird, S. C. Corbato, H. Y. Dai, *et al.*, Astrophys. J. **441**, 144 (1995).

28. M. Takeda, N. Hayashida, K. Honda, *et al.*, Astrophys. J. **522**, 225 (1999).

29. M. P. Veron-Cetty and P. Veron, http://dbsrv.gsfc.nasa.gov/heasarc_veron98.

30. S. Hayakawa, Prog. Theor. Phys. **37**, 594 (1966).

31. O. Prilutsky and I. L. Rozental, Acta Phys. Hung., Suppl. 1 **29**, 51 (1970).

32. G. R. Farrar and P. L. Biermann, Phys. Rev. Lett. **81**, 3579 (1998).

33. G. L. Squires, *Practical Physics* (McGraw-Hill, New York, 1968; Mir, Moscow, 1971).

34. G. E. Forsythe *et al., Computer Methods for Mathematical Computations* (Prentice–Hall, Englewood Cliffs, 1977).

35. N. S. Kardashev, Mon. Not. R. Astron. Soc. **276**, 515 (1995).

36. S. B. Popov, http://xrai.sai.msu.su/~polar/ -2000.

37. H. Spinrad, S. Djorgovsky, J. Marr, *et al.*, Publ. Astron. Soc. Pac. **97**, 932 (1985).

38. H. Kuhr, A. Witzel, and I. I. K. Pauliny-Toth, Astron. Astrophys., Suppl. Ser. **45**, 367 (1981).

39. N. S. Kardashev, private communication (2000).

40. P. P. Kronberg, Rep. Prog. Phys. **57**, 325 (1994).

41. A. V. Zasov, private communication (1998).

42. V. A. Lipovetskiĭ, S. N. Neizvestnyĭ, and O. M. Neizvestnaya, Soobshch. Spets. Astron. Obs. **55** (1987).

43. E. A. Pier and J. H. Krolik, Astrophys. J. **418**, 673 (1993).

44. F. W. Stecker, C. Done, M. H. Salamon, *et al.*, Phys. Rev. Lett. **66**, 2697 (1991).

45. R. Simcoe, K. K. McLeod, J. Schachter, *et al.*, Astrophys. J. **489**, 615 (1997).

46. R. M. Barnett, C. D. Carone, D. E. Groom, *et al.*, Phys. Rev. D **54**, 1 (1996).

47. J. Einasto, M. Einasto, and M. Gramann, Mon. Not. R. Astron. Soc. **289**, 801 (1997).

48. M. Teshima, in *Proc. 23rd International Cosmic Ray Conference, Calgary, 1993*, Ed. by D. A. Leakhy, R. B. Hicks,

and D. Venkatesan (World Scientific, Singapore, 1993), p. 257.

49. S. I. Nikol'sky, ĭ. N. Stamenov, and S. Z. Ushev, Zh. Éksp. Teor. Fiz. **87**, 18 (1984) [Sov. Phys. JETP **60**, 10 (1984)].

50. S. I. Nikolsky, J. N. Stamenov, and S. Z. Ushev, J. Phys. G **13**, 883 (1987).

51. F. Kakimoto, T. Kaneko, H. Yoshii, *et al.*, in *Proc. 20th International Cosmic Ray Conference, Moscow, 1987*, Vol. 1, p. 307.

52. K. Suga, Y. Toyoda, K. Kamata, *et al.*, in *Proc. 20th International Cosmic Ray Conference, Moscow, 1987*, Vol. 1, p. 310.

53. S. Z. Ushev, Candidate's Dissertation (Fiz. Inst. Akad. Nauk SSSR, Moscow, 1985).

54. S. Hayakawa, *Cosmic Ray Physics: Nuclear and Astrophysical Aspects* (Wiley, New York, 1969; Mir, Moscow, 1973), Part 1, p. 582.

55. A. V. Pljasheshnikov, A. A. Lagutin, and V. V. Uchaikin, in *Proc. 16th International Cosmic Ray Conference, Kyoto, 1979*, Vol. 7, p. 1.

56. A. V. Uryson, Kratk. Soobshch. Fiz., No. 10, 56 (1988).

57. A. V. Uryson, in *Proc. 24th International Cosmic Ray Conference, Rome, 1995*, Vol. 1, p. 111.

58. A. V. Uryson, in *Proc. 24th International Cosmic Ray Conference, Rome, 1995*, Vol. 2, p. 238.

59. S. Karakula, J. N. Stamenov, and W. Tkazhyk, in *Proc. 24th International Cosmic Ray Conference, La Jolla, 1985*, Vol. 1, p. 268.

60. F. A. Agaronyan, E. A. Mamidzhanyan, S. I. Nikol'sky, and E. I. Tukish, Astrofizika **23**, 55 (1985).

61. B. E. Markaryan, V. A. Lipovetskiĭ, and D. A. Stepanyan, Astrofizika **21**, 35 (1984).

62. D. A. Stepanyan, private communication (1998).

63. J. V. Jelley, Phys. Rev. Lett. **16**, 749 (1966).

64. L. M. Ozernoĭ, O. F. Prilutskiĭ, and I. L. Rozental, *High-Energy Astrophysics* [in Russian] (Atomizdat, Moscow, 1973).

65. V. I. Gol'danskiĭ, Yu. P. Nikitin, and I. L. Rozental, *Kinematic Methods in High Energy Physics* [in Russian] (Nauka, Moscow, 1987).

66. A. V. Uryson, Kratk. Soobshch. Fiz., Nos. 3–4, 30 (1996).

67. V. L. Ginzburg, *Theoretical Physics and Astrophysics* (Nauka, Moscow, 1987; Pergamon, Oxford, 1979).

68. G. Brooke, G. Cunningham, P. J. V. Eames, *et al.*, in *Proc. 19th International Cosmic Ray Conference, La Jolla, 1985*, Vol. 2, p. 150.

69. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics* (Nauka, Moscow, 1985; Pergamon, Oxford, 1982).

Translated by A. Dambis

Variability of the H₂O Maser Emission in S255

M. I. Pashchenko[1], E. E. Lekht[2], and I. I. Berulis[3]

[1]*Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899 Russia*

[2]*Instituto Nacional de Astrofísica, Óptica y Electrónica, Luís Enrique Erro No. 1,*
Apdo Postal 51 y 216, 72840 Tonantzintla, Puebla, México

[3]*Pushchino Radio Astronomy Observatory, Astro Space Center of the Lebedev Institute of Physics,*
Pushchino, Moscow oblast, 142290 Russia

Received October 26, 2000

Abstract—The paper reports the results of observations of the H₂O maser in S255 carried out in 1982–1985 and 1990–2000 on the 22-meter telescope of the Pushchino Radio Astronomy Observatory. The H₂O maser emission extends from –2 to 14 km/s and is mainly concentrated in three radial-velocity intervals. The velocity of the central group of emission features coincides with that of the molecular cloud, while the two lateral groups (blueshifted and redshifted) are positioned in the spectrum more or less symmetrically relative to the central feature. During the monitoring of S255, two phenomena were observed. First, the integrated flux of the H₂O maser emission varied in a cyclic manner with a period of two to four years; this may be connected with activity of the protostar. Second, the fluxes of emission features (or groups of features) were anticorrelated. The emission of the three groups of features noted above dominated in succession. In some time intervals, a triplet spectral structure with anticorrelation between the fluxes of the lateral components and of the central and lateral components was observed. The flux anticorrelation between groups of features and individual features could be due to competition between spatial emission pumping modes in a nonuniform Keplerian disk. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Among H₂O masers associated with star-forming regions, sources with triplet spectra present special interest. The works of Elmegreen and Morris [1], Grinin and Grigor'ev [2], Cesaroni [3], and Lekht *et al.* [4] are devoted to studies of such maser sources. The triplet spectra could be formed in an expanding envelope or in a rotating torus-shaped disk (Keplerian disk) observed edge on. In the latter case, the condition $M_\star \gg M_{\mathrm{Disk}}$ should be satisfied (where $M_\star$ is the mass of the central star and M_{Disk} is the mass of the disk).

One such source whose maser emission arises in a Keplerian disk is the H₂O maser S255, which was discovered by Lo and Burke in 1973 [5] and remains fairly well known. The maser is located in the S255 IR complex, which lies between the two extended radio continuum sources S255 and S257.

S255 IR is a compact region of active star formation 2.4 kpc from the Sun [6], in which eight HH objects, two high-velocity molecular outflows, and numerous infrared sources have been observed. The ultracompact H II region G192.58–0.04 [8], which is nearest to S255 IR [8] and is associated with one of the molecular out-

flows [9], is located at a projected distance of 64″ to the north of S255 IR.

According to CS line observations [7], the velocity of the star-forming molecular cloud is 8.3 km/s. At the same time, the velocity of the central peak in the H₂O spectrum is approximately 7 km/s. Thus, the line-of-sight component of the peculiar velocity of the H₂O maser S255 is small, only about –1 km/s. The estimated mass of the central star is $M_\star \simeq 18 M_\odot$ [3].

Finally, it may be important that two class I and II methanol masers were detected in this direction by Haschick *et al.* [11] and Menten [12], respectively. The point here is that CH₃OH masers of both classes are found in star-forming regions. However, class II methanol masers coincide spatially with OH and H₂O masers and ultracompact H II regions, whereas class I masers are not associated with any known objects, though they are located in the vicinities of star-forming regions, near the boundaries of dense gas–dust clouds.

Class II methanol masers are frequently arranged in straight lines or arcs within high-velocity outflows. Norris *et al.* [13] suggested that they arise in protoplanetary disks. The S255 H₂O maser probably coincides with the class II methanol maser, while the class I meth-
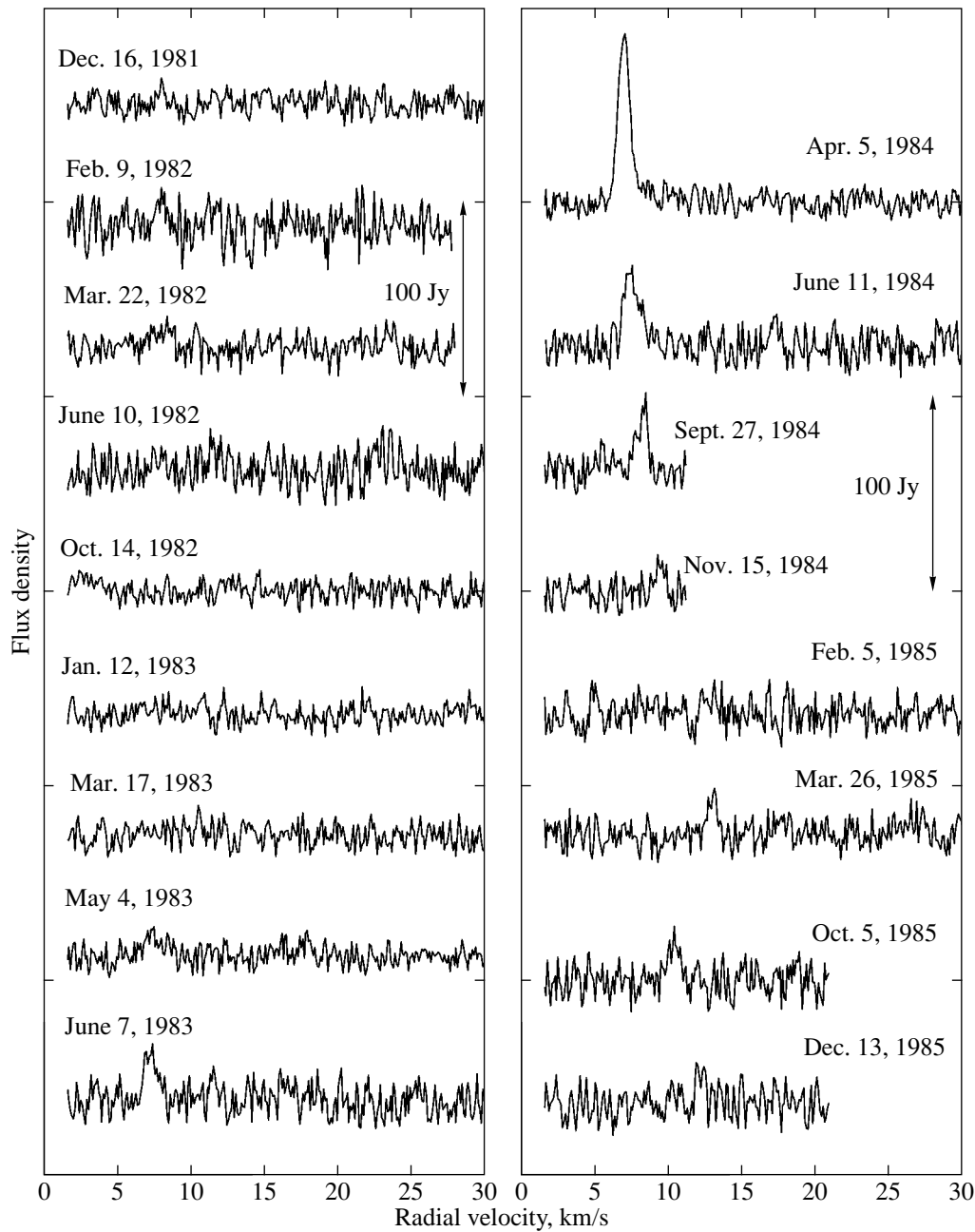
Fig. 1. Spectra of the H$_2$O maser emission of S255 observed in 1982–1985 and 1990–2000. The spectrum of January 16, 2001 was added in proof.

anol maser is slightly shifted to the north of the H$_2$O maser. However, no interferometric measurements or maps in methanol radio lines are available for this region, the exact positions of these masers are not known, and definite conclusions about their connection with the water-vapor maser cannot be drawn.

The spectra of the H$_2$O maser display regular behavior over some period of time. In 1987–1988, Cesaroni [3] observed emission in three narrow, almost equally-spaced radial-velocity intervals. At other times, only

two components were observed. The disappearance of a component in the H$_2$O spectrum could be a consequence of the anticorrelation between the component intensities. This anticorrelation can be explained in models with competing spatial emission modes in a partially saturated disk maser with anisotropic pumping [3, 4].

In the current paper, we study the evolution of the H$_2$O maser emission in S255 on the basis of long-term monitoring data.
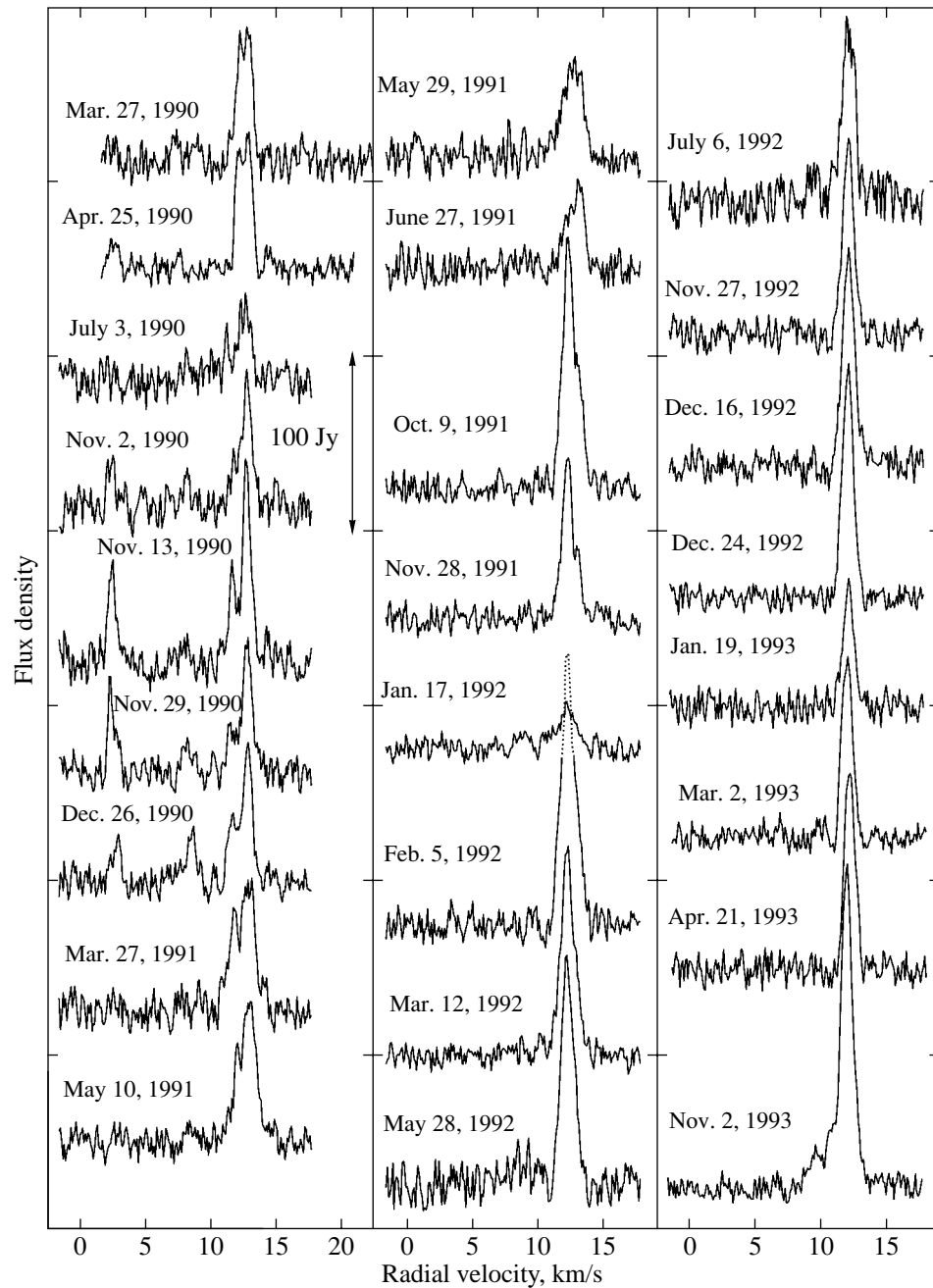
Fig. 1. (Contd.)

2. OBSERVATIONS AND DATA

The observations of the maser emission in the 1.35-cm H_2O line toward S255 IR ($\alpha_{1950} = 06^h09^m58.2^s$, $\delta_{1950} = 18°00'17''$) were carried out on the 22-meter radio telescope in Pushchino from December 1981 to December 1985 and from March 1990 to September 2000. The average interval between our observations was about two months. The noise temperature of the system with a cooled transistor amplifier at the front end was 150–200 K. The spectral analysis of the signal was performed using a 96-channel (starting in July 1997, 128-channel) filter-bank spectrum analyzer with a resolution of 7.5 kHz (0.101 km/s in radial velocity in the 1.35-cm line). We did not observe the S255 maser during the four years from 1986 to 1989. Therefore, the results of similar observations carried out by Cesaroni in 1987–1989 [3] are of considerable interest here.
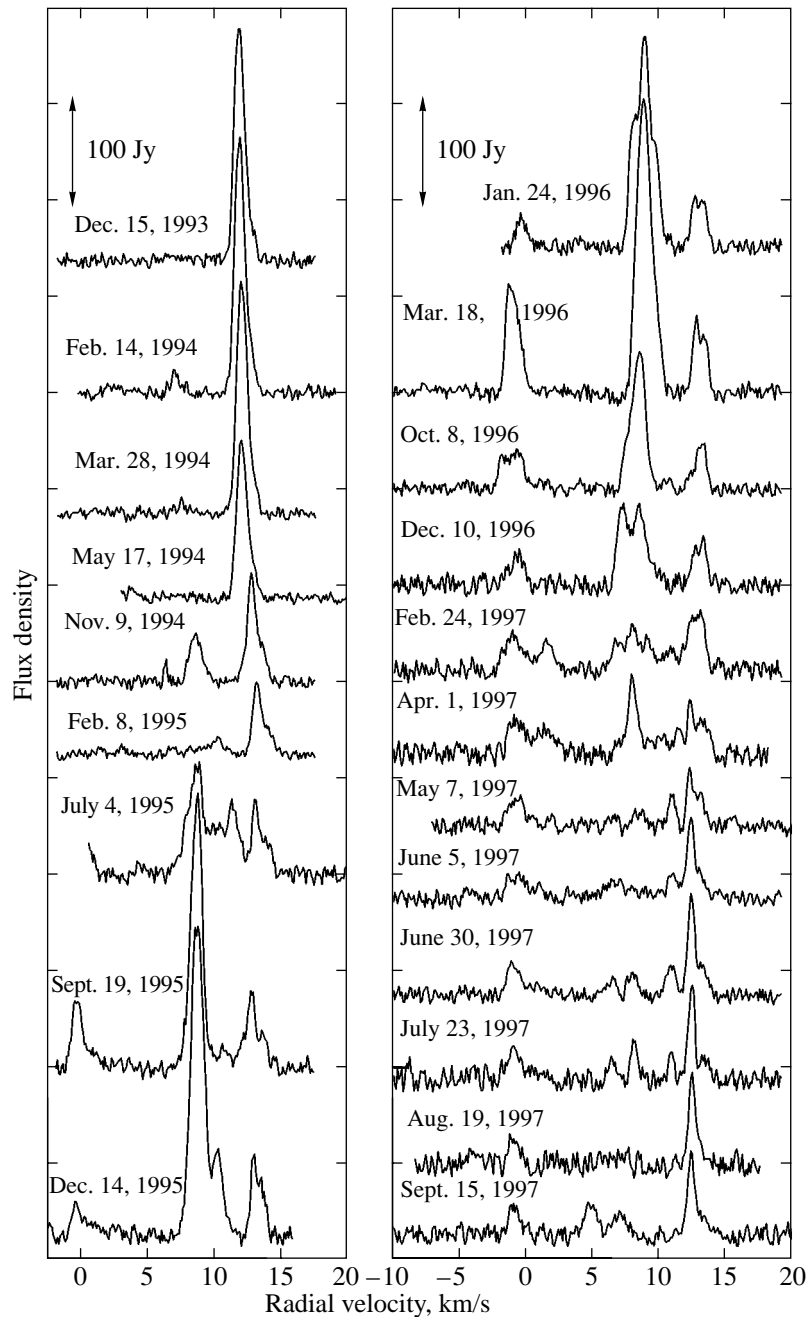
Fig. 1. (Contd.)

Figure 1 presents the results of our observations of the H$_2$O maser in S255. The horizontal axis is the radial velocity in km/s with respect to the Local Standard of Rest, and the vertical axis is the flux density in Jy (the scale in Jy is shown by a vertical arrow). For convenience, the horizontal axes for all the spectra are given on the same scale.

Figure 2 shows the variability of the integrated H$_2$O flux. For completeness, we have added Cesaroni's data for 1987–1989 [3]. Cyclic activity of the maser emission with a characteristic interval between maxima of two to four years is visible. Accordingly, we have divided the 1982–2000 interval into eight sections, denoted "a" through "f." Section "g" is located between the two main maxima. Within each section, we have fitted Gaussians to the spectra and determined their peak values. In Fig. 2, the positions of the maxima are shown with vertical arrows. For section "a," we give an upper limit for its maximum flux.
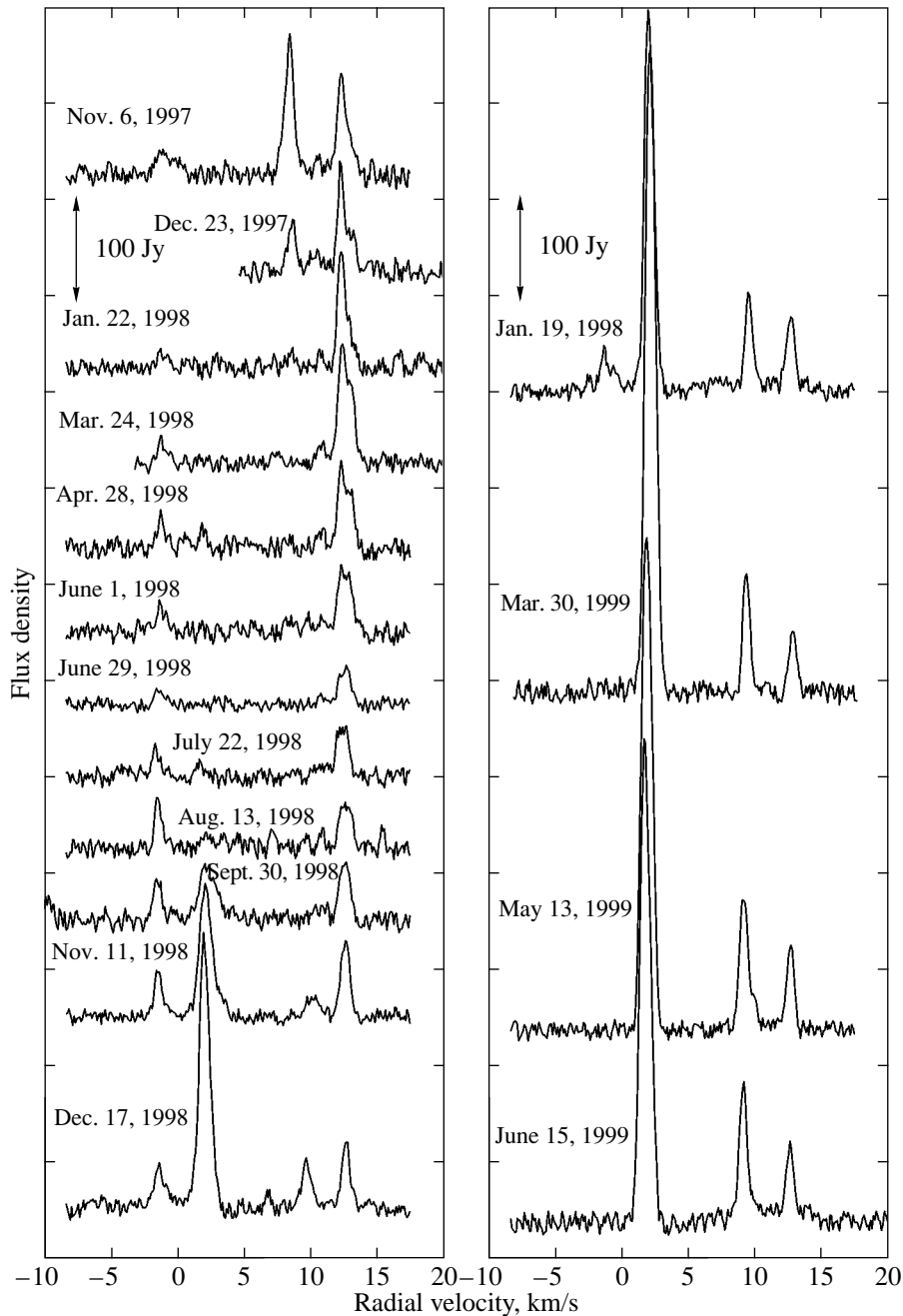
Fig. 1. (Contd.)

Figures 3 and 4 plot the superposition of the spectra and the average spectrum for each time interval. The boundaries between the selected time intervals are shown, with corresponding labels given in parentheses. The time interval 1987–1989, when we did not observe S255, is omitted. To improve the sensitivity (of course, at the expense of the spectral resolution), we have averaged pairs of adjacent spectral channels. We can clearly see that the emission from one group of features—blueshifted, cen-

tral, or redshifted—dominates in each interval. In addition, faint emission appeared for a short time at intermediate radial velocities. A similar pattern was observed in S140 [4], which is identified with a Keplerian disk.

Thus, during our monitoring in various time intervals, we observed emission at nearly all radial velocities from –2 to +14 km/s (except for a narrow interval from 4 to 6 km/s). However, the emission in three sections of the spectrum—blueshifted, central, and red-
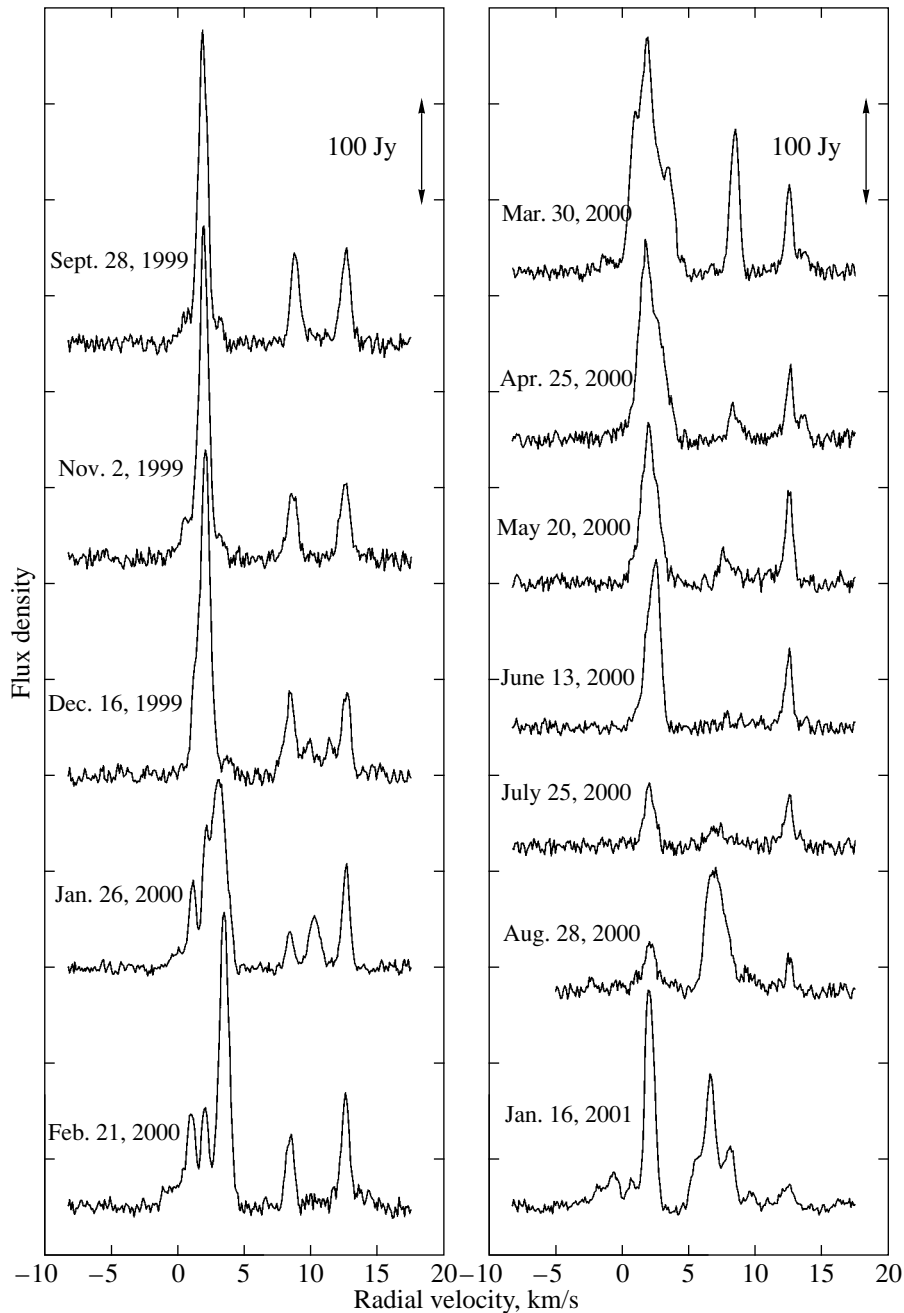
Fig. 1. (Contd.)

shifted—was dominant. Figure 5 shows variations of the integrated emission for these spectral regions. The arrows mark maxima of the total flux. We slightly expanded the integration interval when plotting Fig. 5, since, in our monitoring, emission was observed over a broader range of velocities than for more or less stable main-component emission. We adopted integration intervals from –2 to 4, 6 to 10, and 10 to 14.5 km/s. We detected no emission at velocities from 4 to 6 km/s.

The spectra obtained between July 1998 and January 1999 had a more complicated structure. In comparison with the period observed by Cesaroni [3] (1987–1989), there was no central component, and the blue-shifted and redshifted components had a doublet structure. In turn, each component of the doublets consisted of several emission features with closely spaced radial velocities. The velocities of the main doublet features did not vary with time, except for feature 3 (doublet 2).
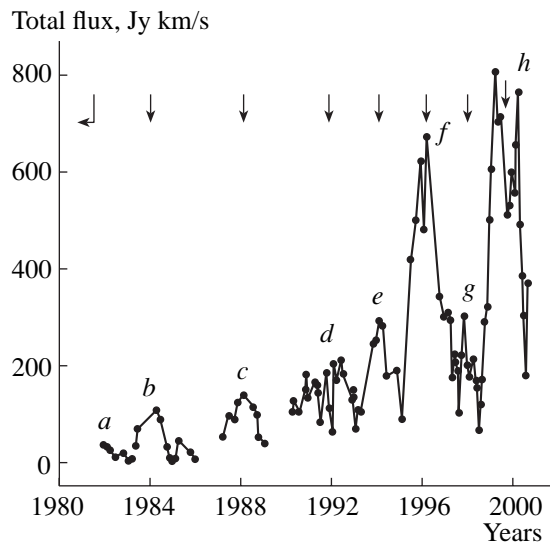
Fig. 2. Variations of the integrated H_2O maser emission flux of S255. The data for 1987–1989 are from [3]. Arrows show the total flux maxima derived from Gaussian fitting.

Figure 6 shows the variations of the velocities of the main features.

After the disappearance of the emission feature at -3 km/s, starting in March 1999, the spectrum again acquired a triplet structure; however, the spectrum differed strongly from those of 1987–1989 [3]. Each component of the triplet consisted of several features with closely spaced radial velocities. The blueshifted component, whose features in 2000 extended from -1.5 to 5 km/s, had the most complicated structure. A peculiarity of this period is a drift of feature 3 from 10.8 to 8.5 km/s. The radial velocity of this emission feature is close to that of the central component observed by Cesaroni [3]. Since about May 2000, the appearance of the H_2O spectrum has been the same as in 1987–1989.

3. DISCUSSION

We now analyze our long-term monitoring data for S255 and compare our and Cesaroni's results [3]. First, we note the differences between our and Cesaroni's observational data.

(1) We monitored S255 for a much longer time.

(2) Our observations had higher spectral resolution.

(3) During our monitoring, the emission displayed considerably larger flux excursions ($F_{max} \simeq 670$ Jy) than in Cesaroni's observations [3] (no more than 70 Jy).

(4) The H_2O spectra of S255 were very complex and contained a large number of emission features. Most of

the time, they did not resemble the spectra of Keplerian-disk sources.

3.1. Integrated Flux Variations

Together with Cesaroni's data [3], the available regular observations of the H_2O maser emission in S255 cover 19 years (1981–2000). Such long-term monitoring has enabled us to reveal important regularities in the integrated-flux variations. We have found cyclic variability with a timescale of two to four years. The variations are similar to those of supergiant circumstellar masers, which are sources of a different type and nature. Such variability could be connected with a nonstationary character for the formation of the central star. However, in addition to the cyclic activity of the H_2O maser, the emission in each of the spectral intervals dominates in turn (Fig. 5). This indicates a more complex evolution of the total H_2O maser emission in S255 than in the case of purely cyclic activity of the central star.

3.2. Structure of the H_2O Spectra

During our observations of the S255 H_2O maser in 1990–2000, we sometimes observed triplet structure in the spectrum. Flux anticorrelations for single features were also rarely observed. In 1986–1989, Cesaroni [3] observed an anticorrelation of the fluxes of the lateral components, however, in 2000, there was a flux anticorrelation for the central and blueshifted components, as in S140 [4]. In this connection, we note that the anticorrelation in the maser emission of S255 in 1986–1989 was observed for fluxes not exceeding 70 Jy. In our observations, the S255 maser was usually considerably more intense, with the spectra consisting of a large number of features. However, the emission of one of the three main groups—blueshifted, central, or redshifted—always dominated (Fig. 5). The phase when the emission of any single group was dominant lasted about two years. The time intervals between the emission maxima were also about two years.

The emission of any lateral or central component can dominate owing to the flux anticorrelation between the lateral and central components or between the lateral components. This takes place in the case of an even partially saturated maser. The degree of maser saturation can be estimated, for example, if there is a functional dependence between intensity and linewidth variations. However, due to the lack of isolated features in the spectra, we could not trace the linewidth variations for individual features with sufficient accuracy to detect such correlations.

Note that, the state of activity of the central star and its effect on the maser activity and spectral structure have not been considered for masers in Keplerian disks. A forming star can have activity maxima and minima. In S255, during periods of higher activity of the central
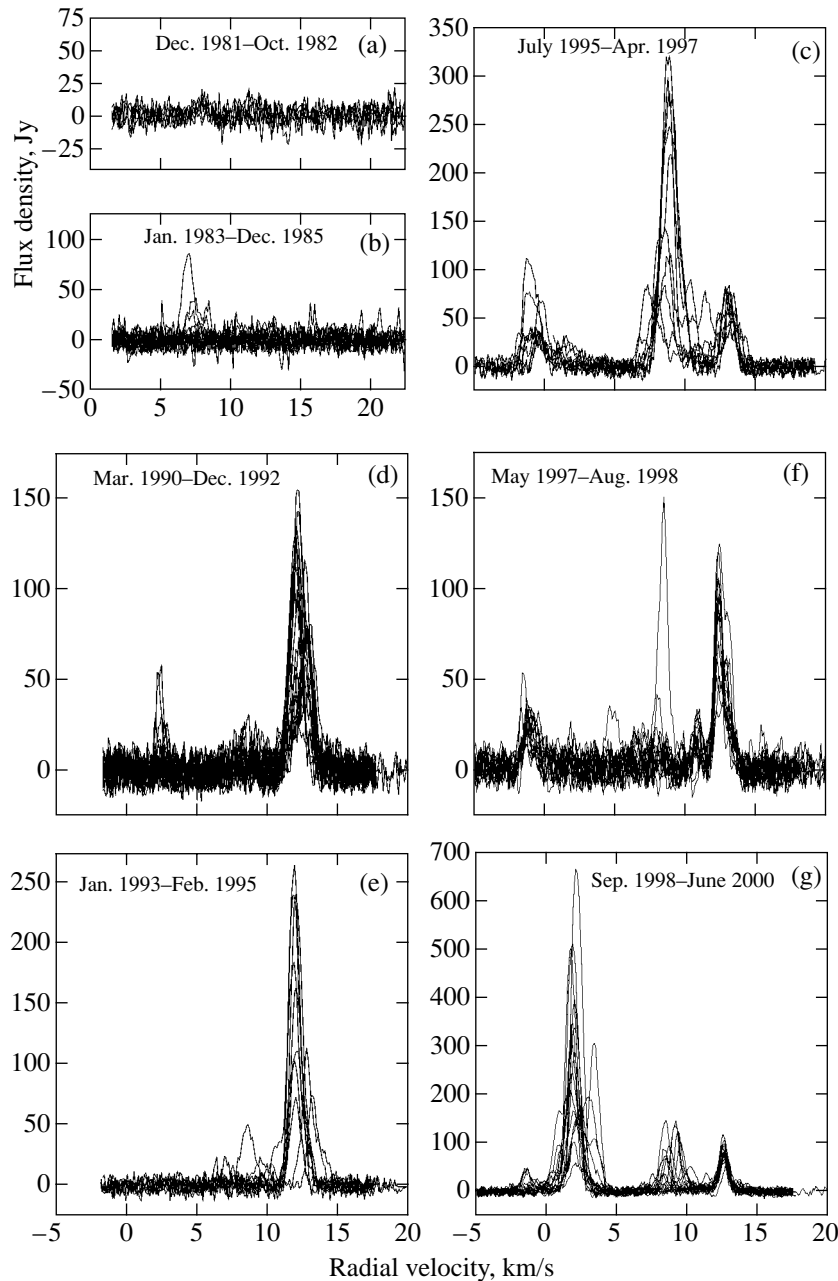
Fig. 3. Superposition of the H$_2$O maser emission spectra of S255 for various time intervals (the time intervals are designated as in Fig. 2).

star, the maser emission is more intense and the spectra contain a larger number of features than in 1987–1988 [3]. In such periods, competition of spatial emission pumping modes can occur between groups of features with closely spaced radial velocities, but not between individual features. As a result, the emission of the central and one lateral component, or of both lateral components, is suppressed. The observed flux anticorrelation supports the model of a Keplerian disk in S255.

3.3. Radial Velocity Variations of Individual Features

In 1998–1999, as in 1987–1989, one of the features in the H$_2$O spectra of S255 drifted in radial velocity. In 1987–1989, the redshifted component drifted [3] at a rate of 0.25 km/s per month. In 1998–1999, drifting of the central component at a radial velocity of about 10 km/s was observed; the drift rate was 0.14 km/s per month. The drift direction was the same in both cases. For comparison, in [3], the central component had a
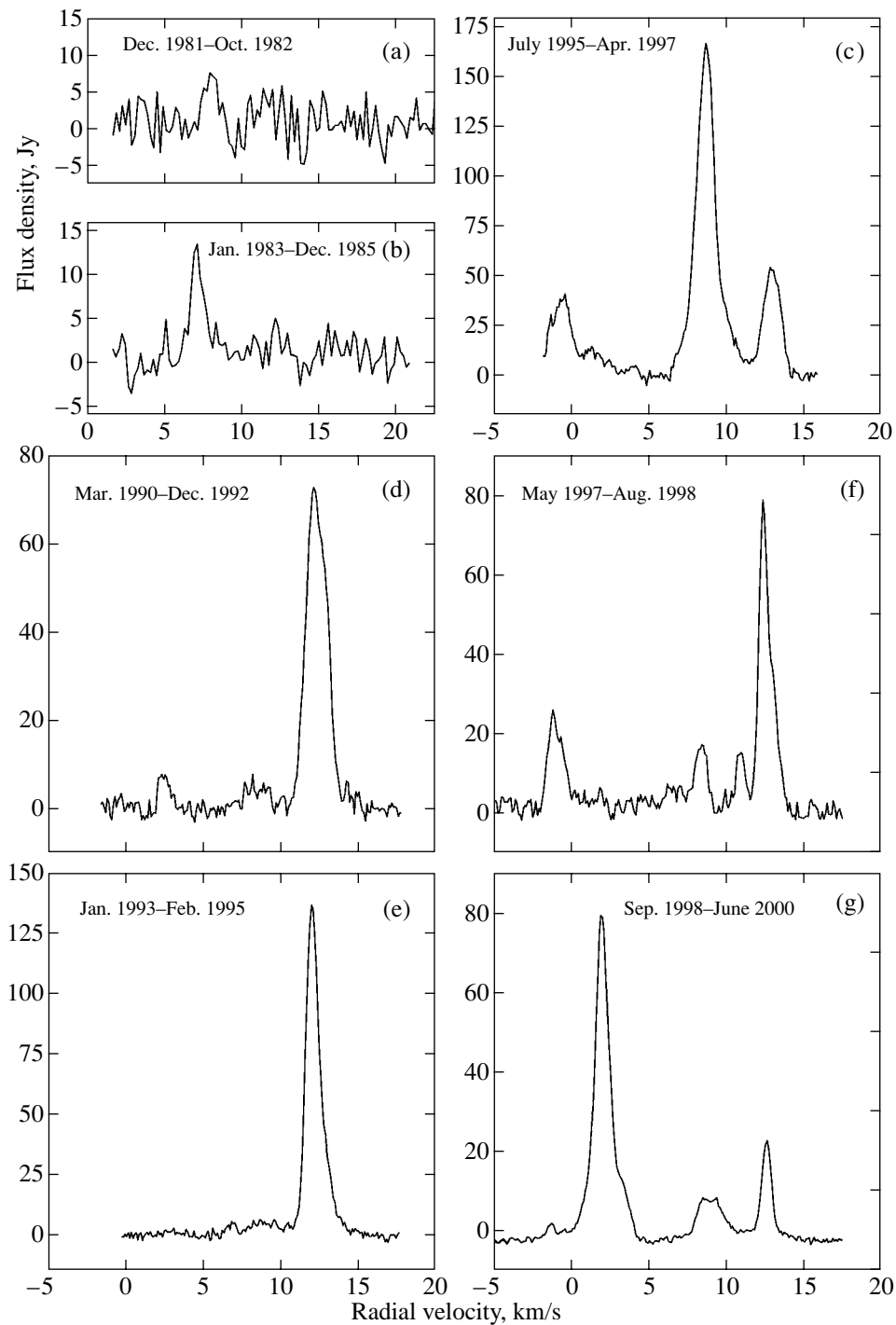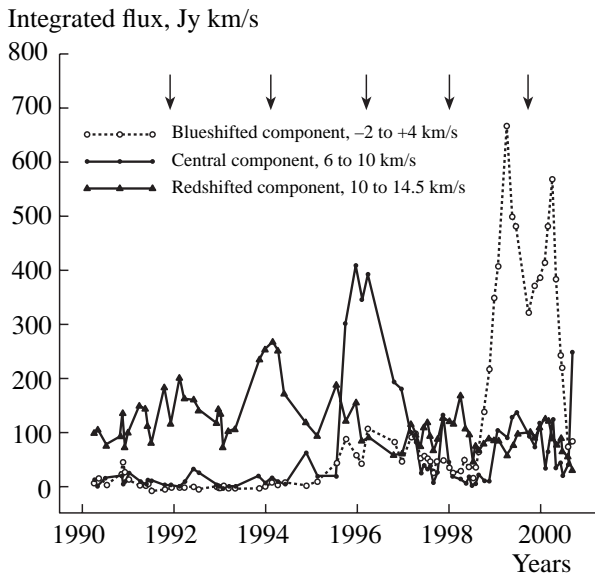
Fig. 4. Averaged spectra of the H$_2$O maser S255 for the same time intervals as in Fig. 3.

radial velocity of 7 km/s. This means that, in our case, the radial velocity was intermediate between the velocities of the central and redshifted components.
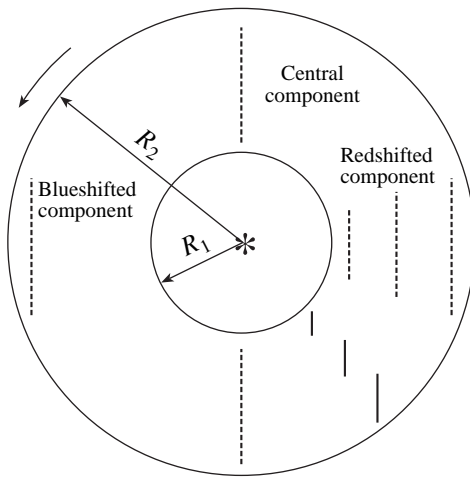
It follows from [3] that, in 1987–1989, the emission of the redshifted component was received from different parts of the Keplerian disk at different observational epochs (from the inner to the outer parts); i.e., different regions of the disk were excited in succession. This explains the decrease of the radial velocity of the redshifted component with time.
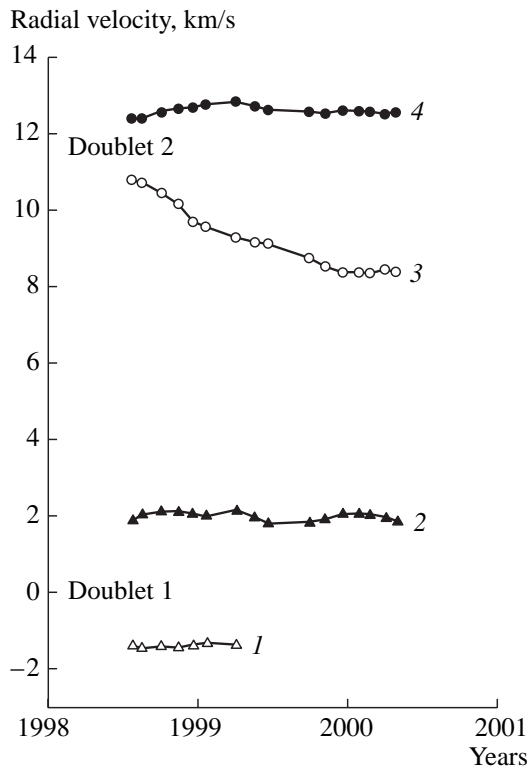
We suggest that the drift mechanism in 1987–1989 and 1998–1999 was the same. The difference consists only in the location of the emitting regions within the disk. Figure 7 shows a possible arrangement of the main maser-

Integrated flux, Jy km/s

Blueshifted component, –2 to +4 km/s
Central component, 6 to 10 km/s
Redshifted component, 10 to 14.5 km/s

Fig. 5. Variations of the integrated flux for three main groups of emission features: blueshifted, central, and redshifted. Arrows show the total flux maxima derived from Gaussian fitting.

Radial velocity, km/s

Doublet 2

4

3

2

Doublet 1

1

Fig. 6. Variations of the radial velocity of the main emission features for 1998–2000.

ing regions in S255. The formation of the maser emission regions in a Keplerian disk was considered in detail in [3, 4].

Central component

Redshifted component

Blueshifted component

Fig. 7. Schematic representation of the main regions of H$_2$O maser emission in S255. The arrow at the left shows the direction of rotation of the Keplerian disk; R_1 and R_2 are the inner and outer radii of the disk. The dotted lines show corridors of maximum amplification for the blueshifted, central, and redshifted components, and the solid lines show the same for the group of features at 8–11 km/s (1999–2000).

4. CONCLUSION

We have reported the results of monitoring the H$_2$O maser S255 from 1982–2000. The data for 1987–1989 were taken from [3]. We have obtained the following results.

(1) The H$_2$O spectra of S255 consist of a large number of features, located throughout the radial-velocity interval from –2 to 14 km/s, except for the interval from 4–6 km/s. However, features were clustered in three main groups: blueshifted, central, and redshifted.

(2) There was cyclic variability of the integrated H$_2$O flux with an interval between maxima of two to four years. This variability is most likely due to the non-stationarity of the star-formation process.

(3) During our observations, the emission of different groups dominated in succession. The variability was more or less periodic, with an interval between maxima of about two years. Anticorrelation of the fluxes of individual features (central and blueshifted) was observed only once, during 2000.

(4) One of the central-group features drifted in radial velocity. This was probably not associated with an actual physical displacement of masing material but simply with a shift of the maser-generation region from the inner to the outer parts of the disk.

(5) The model of a Keplerian disk in S255 is supported by these new observations; the disk must be very nonuniform.

Thus, we have observed two phenomena in S255. The first is connected with cyclic activity of the central star during its formation. The second is an anticorrela-

tion between the fluxes of the lateral components or between the central and lateral components.

ACKNOWLEDGMENTS

REFERENCES

1. B. G. Elmegreen and M. Morris, Astrophys. J. **229**, 593 (1979).
2. V. P. Grinin and S. A. Grigor'ev, Astron. Zh. **60**, 512 (1983) [Sov. Astron. **27**, 298 (1983)].
3. R. Cezaroni, Astron. Astrophys. **223**, 513 (1990).
4. E. E. Lekht, S. F. Likhachev, R. L. Sorochenko, and V. S. Strel'nitskiĭ, Astron. Zh. **70**, 731 (1993) [Astron. Rep. **37**, 367 (1993)].
5. K. Y. Lo and B. F. Burke, Astron. Astrophys. **26**, 487 (1973).
6. D. T. Jaffe, J. A. Davidson, M. Dragovan, and R. H. Heldebrand, Astrophys. J. **284**, 637 (1984).
7. M. Morris, P. Palmer, B. E. Turner, and B. Zuckerman, Astrophys. J. **191**, 111 (1974).
8. S. Kurtz, E. Churchwell, and D. O. S. Wood, Astrophys. J., Suppl. Ser. **91**, 659 (1994).
9. M. P. Miralles, L. Salas, I. Cruz-González, and S. Kurtz, Astrophys. J. **488**, 749 (1997).
10. B. F. Turner, Astrophys. J. Lett. **8**, 73 (1971).
11. A. D. Haschik, K. M. Menten, and W. A. Baan, Astrophys. J. **354**, 556 (1990).
12. K. M. Menten, Astrophys. J. Lett. **380**, L75 (1991).
13. R. P. Norris, J. B. Whiteoak, J. L. Caswell, *et al.*, Astrophys. J. **412**, 222 (1993).

Translated by G. Rudnitskiĭ

The Structure of Matter Flows in Semi-Detached Binaries after the Termination of Mass Transfer

D. V. Bisikalo[1], A. A. Boyarchuk[1], A. A. Kil'pio[1], O. A. Kuznetsov[2], and V. M. Chechetkin[2]

[1]*Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia*
[2]*Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya pl. 4, Moscow, 125047 Russia*
Received January 11, 2001

Abstract—We present the results of three-dimensional gas-dynamical simulations of matter flows in semi-detached binaries after termination of the mass transfer between the components of the system. The structure of the residual accretion disk is studied. When the mass transfer has ended, the quasi-elliptical disk becomes circular and its structure changes: tidal interactions result in the formation of a second arm in the spiral shock wave. In addition, a condensation (blob) moving through the disk with variable velocity is formed. The blob is maintained by interactions with the arms of the spiral shock and exists essentially over the entire lifetime of the disk. We also show that, for a viscosity corresponding to $\alpha \sim 0.01$ (typical for observed accretion disks), the lifetime of the residual disk is about 50 orbital periods. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Semidetached binaries are one type of interacting star in which one of the components fills its Roche lobe, resulting in mass transfer between the components through the vicinity of the inner Lagrange point L_1. We have developed a three-dimensional (3D) gas-dynamical model for this process [1–3] and applied it to binaries with steady mass-transfer rates [1–5], making it possible to determine the basic parameters of the flow pattern. In particular, we found that the flow–disk interaction is shockless in the steady-flow regime. The "hot line" (rather than the traditional "hot spot") model proposed for systems with steady-state mass-transfer rates can successfully describe the light curves of cataclysmic binaries [6, 7], indicating its applicability for these systems.

However, observations indicate that, in a number of semi-detached binaries, along with steady-state stages of mass transfer, there also exist stages in which the rate of mass transfer between the components varies appreciably [8–10]. In addition, studies of the evolution of binary systems indicate that the radius of the donor star decreases at certain stages, so that the star may be entirely contained within its Roche lobe, resulting in a complete termination of the mass transfer in the system.

Here, we study the structure of the matter flows in a semi-detached binary after the termination of mass transfer. As a first step, we carried out 3D gas-dynamical simulations of the mass transfer in a system with a specified transfer rate, up until the establishment of the steady-state regime. Further, we supposed that the outflow in the system is terminated and considered the structure of the residual accretion disk. It is evident that the evolution of the accretion disk is determined by physical processes leading to the redistribution of angular momentum in the disk. To study the effect of viscosity on the flow structure after the termination of mass transfer, we carried out three computations with different viscosities, corresponding to $\alpha \sim 0.08–0.1$, 0.04–0.06, and 0.01–0.02 in α-disk terminology.

2. FORMULATION OF THE PROBLEM

Let us consider a semi-detached binary system with the mass of the accretor M_1, mass of the donor star M_2, distance between the centers of the components A, and angular velocity of the orbital rotation Ω. We will describe the matter flows in the system using three-dimensional Euler equations. We close the system of equations with the equation of state for an ideal gas $P = (\gamma - 1)\rho e$, where P is the pressure, ρ the density, e the specific internal energy, and γ the adiabatic index. To imitate a system with radiative losses, we adopt an adiabatic index close to unity ($\gamma = 1.01$), which corresponds to a case that is close to isothermal [1, 11].

To obtain a numerical solution for the system of equations, we used a high-order Roe–Osher TVD scheme [12, 13] as modified by Einfeldt [14]. In our Cartesian coordinate system XYZ, the X axis was directed from the center of the donor star toward the accretor, the Y axis along the orbital motion of the donor star, and the Z axis perpendicular to the orbital plane. For the sake of convenience, we wrote the initial system of equations in dimensionless form: all spatial variables were normalized to the distance between the components A, times to the inverse angular velocity Ω^{-1}, and densities to the density at the inner Lagrange point L_1. The computations were carried out within the parallelepiped $[1/2A \ldots 3/2A] \times [-1/2A \ldots 1/2A] \times [0 \ldots 1/4A]$ and only in the upper half-space relative to the equatorial plane, since we assumed

that the system was symmetrical about this plane. The accretor was represented by a sphere with a radius of 1/100 and excluded from the computation region. We specified the following boundary conditions:

(1) free outflow of gas onto the accreting star, as well as through the outer boundary of the computation region;

(2) in the first stage, matter with $\rho = \rho(L_1)$, $V_x = c(L_1)$, $V_y = V_z = 0$, where $\rho(L_1)$, $c(L_1)$ are the density and sound speed of the gas at L_1, was injected into the system through the grid point corresponding to L_1;

(3) in the second stage, after the steady-state flow regime was reached, at time $t = t_0$, the density of the matter injected into the system through L_1 was decreased by five orders of magnitude.

As initial conditions, we assumed that the system was filled with a rarified intercomponent "background" gas with $\rho_{back} = 10^{-5}\rho(L_1)$, $P_{back} = 10^{-4}\rho(L_1)c^2(L_1)/\gamma$, and $V_{back} = 0$. The assumption of the presence of a background gas with $\rho_{back} = 10^{-5}\rho(L_1)$ in the system is justified for interacting binaries from an astronomical point of view. In addition, the presence of the background gas simplifies the solution of the gas-dynamical equations, due to the finite density gradients in such a system. Calculations with different background gas densities showed that the gas-dynamical solution is independent of the initial conditions when $\rho_{back} < 10^4\rho(L_1)$.

It becomes clear from analysis of the problem that a gas-dynamical solution for a semi-detached system is specified by three dimensionless parameters [5, 6, 15]: the mass ratio $q = M_2/M_1$, the Lubow–Shu parameter $\varepsilon = c(L_1)/A\Omega$ [15], and the adiabatic index γ. As noted above, in our computations, we assumed that the adiabatic index was close to unity ($\gamma = 1.01$). The computations presented in [2, 4] indicate that the basic gas-dynamical characteristics of 3D flow structures are qualitatively similar over a broad range of q and ε. In our computations, we took the mass ratio of the components to be $q = 1$ and the sound speed to be a factor of ten smaller than the orbital velocity; i.e., $\varepsilon = 1/10$.

To estimate the influence of the viscosity on the solution, we carried out computations with various spatial resolutions. Since the Euler equations do not contain a physical viscosity, we varied the numerical viscosity by varying the parameters of the computation grid. We chose three values for the grid resolution: $31 \times 31 \times 17$, $61 \times 61 \times 17$, and $91 \times 91 \times 25$ (cases A, B, and C, respectively). The grid was uniform in all cases. In α-disk terminology, the numerical viscosity for cases A, B, and C corresponded roughly to $\alpha \sim 0.08$–0.1, 0.04–0.06, and 0.01–0.02.

3. FLOW PATTERN AFTER MASS-TRANSFER TERMINATION

To study the flow structure after mass-transfer termination, we used the previously obtained solutions for steady-state mass transfer as initial conditions. We then assumed that the mass transfer terminates at time $t = t_0$, corresponding to the establishment of the steady-flow regime. Beginning at this time, the rate of mass transfer was decreased by five orders of magnitude, which computationally corresponds to its termination.

We continued the computations in cases A and B until the density in the computation region became equal to the background density $\rho_{back} = 10^{-5}\rho(L_1)$, corresponding to the complete removal of matter from the system, due to either accretion or outflow through the outer boundary. The computation times from t_0 were 5 and $12P_{orb}$ for cases A and B, respectively. For case C, with better grid spatial resolution (minimum viscosity), we continued the computations only to $\sim10P_{orb}$, due to the infeasibly large amount of computational time that would be required to continue to $\sim50P_{orb}$, when the density in the system becomes comparable to the background. We extrapolated the case-C computations for $t > 10P_{orb}$.

Prior to proceeding to analysis of the flow structure, let us estimate the lifetime of the residual disk after mass-transfer termination for the various viscosities. We will use the data for the variations in the mass of gas in the system, presented in Fig. 1. This figure displays the variations in the mass of gas in the computation region (upper panel) $[0.9A…1.1A] \times [–0.1A…0.2A] \times [0…0.05A]$ and in the accretion disk (lower panel) for cases A, B, and C. As noted above, we set the background density as a lower limit for the density in the system; therefore, the mass for $\rho = \rho_{back}$ contained in the system or the accretion disk is a lower limit for the end of the computations. These limits are plotted by the horizontal dashed lines in the upper and lower panels of Fig. 1.

Figure 1 indicates that, as expected, the lifetime of the disk (τ_{disk}) increases as the viscosity decreases: for case A ($\alpha \approx 0.1$) $\tau_{disk} \sim 5P_{orb}$, and for case B ($\alpha \approx 0.05$), $\tau_{disk} \sim 12P_{orb}$. For case C ($\alpha \approx 0.01$), the extrapolated computations suggest that the lifetime of the disk will exceed 50 orbital periods. Note that $\alpha \approx 0.01$ can be considered a typical value for observed accretion disks (see, for example, [16–19]). The characteristic lifetime of the accretion disk in a binary system after mass-transfer termination is therefore about 50 orbital periods, i.e., approximately several days for a typical dwarf nova.

Let us now consider the structural variations of the residual accretion disk as a function of time. Qualitatively, the solutions for cases A, B, and C do not differ, and we will consider in detail the solution for case B. This choice is dictated by two facts. First, in this case, the spatial resolution is sufficiently high to clearly display the structure of the accretion disk in detail. Second, the lifetime of the residual accretion disk is fairly short ($\tau_{disk} \sim 12P_{orb}$), so that we can study the variations
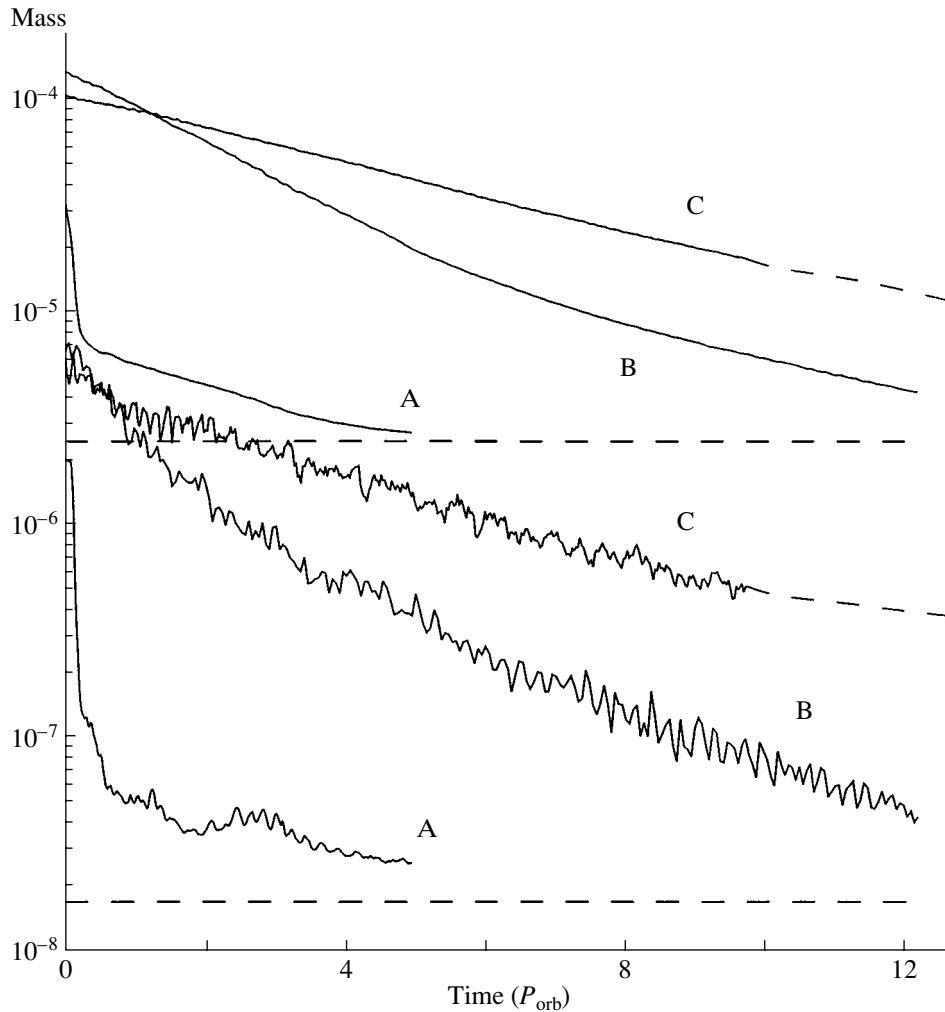
Fig. 1. Variations of the gas mass in the computation region (above) and in the accretion disk (below) for computation cases A, B, and C. The horizontal dashed lines correspond to the mass of gas for the background density $\rho = \rho_{back} = 10^{-5}\rho(L_1)$ contained within the computation region (above) and in the disk (below).

of the disk structure right up to its complete disappearance.

Figure 2 presents density contours (left) and 3D density maps (right) in the equatorial plane of the accretion disk for various times. The upper plots correspond to time $t = 0.08P_{orb}$, the middle plots to $t = 0.16P_{orb}$, and the lower plots to $t = 0.9P_{orb}$ after the time t_0 for the termination of the matter outflow through the inner Lagrange point. This figure indicates that, shortly after the initial time ($t = 0.08P_{orb}$), the structure of the accretion disk does not differ dramatically from the steady-state case (see, for example, [1–5]). A stream flowing from the inner Lagrange point still dominates. We can clearly see the single morphology of the stream–disk system, which is manifest, in particular, in the quasi-elliptical shape of the disk and the absence of a hot spot at the point of contact of the stream and disk. At the same time, the interaction of the binary envelope with the stream results in the formation of an extended shock

wave along the boundary of the stream (a hot line). Figure 2 also indicates that tidal interactions with the donor star result in the formation of a spiral shock wave in the disk, located in the first and second quadrants of the coordinate plane. Double-arm spiral shocks in the disk were first discovered in the studies [20–22]. In our computations, near the initial time, we see only a single-arm spiral shock (upper two plots in Fig. 2). In the place where the second arm should be, the flow is determined by the stream from L_1, which apparently inhibits the formation of the second arm of the shock.

We can see from the middle two plots in Fig. 2 that, after a time $\sim 0.16P_{orb}$, the stream from L_1 ceases to affect the structure of the residual accretion disk, and its initial quasi-elliptical shape becomes essentially circular. The shock wave along the edge of the stream from L_1 (hot line) also disappears from the computed structure of the accretion disk. At the same time, the second arm of the spiral shock forms in the disk, in the third
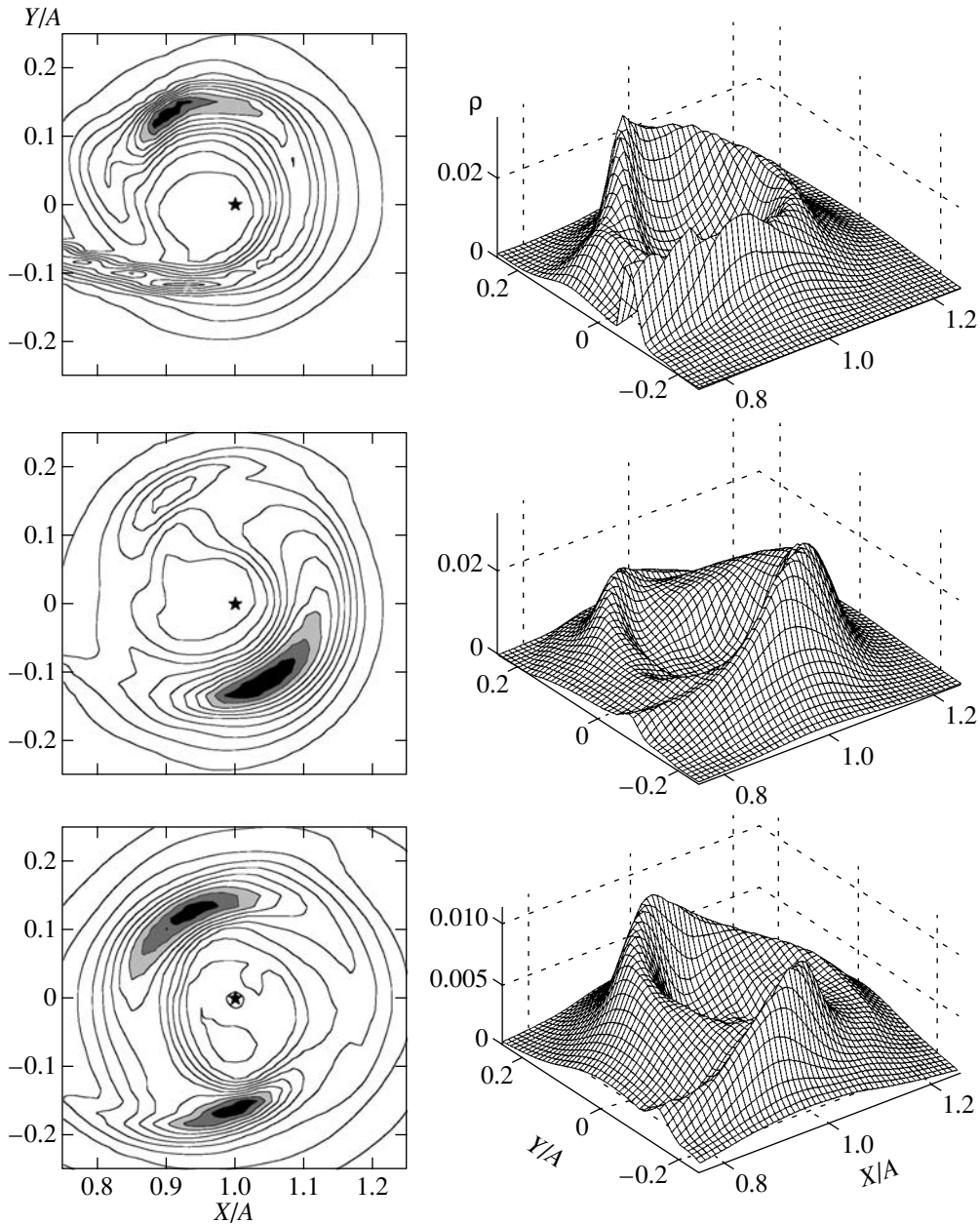
Fig. 2. Density contours (left) and 3D density maps (right) in the equatorial plane of the system for case B. The three pairs of plots from top to bottom correspond to three times after the mass-transfer termination: $t = 0.08$, 0.16, and $0.9P_{orb}$. The maximum densities are 0.0354, 0.0345, and 0.0119 in the upper middle, and lower rows, respectively. The maximum densities for isolines in the left panels were defined as $\rho_{max}/11$.

and fourth quadrants of the coordinate plane. This disk structure, with a double-arm spiral wave, remains unchanged over a long time. This is illustrated by the results of our computations corresponding to $t = 0.9P_{orb}$, displayed in the lower two plots in Fig. 2. Note that both the intensity of the shocks and the mass of the accretion disk decrease with time, although the flow pattern does not change qualitatively. The disk structure changes when the tidal interaction no longer leads to the formation of spiral shocks. This corresponds to the time $t \sim (10–11)P_{orb}$ for case B. We can thus conclude that

spiral shock waves exist in the residual accretion disk essentially over its entire lifetime (recall that this is approximately 12 orbital periods for case B).

The computations indicate that, after the disappearance of the stream, i.e., on a timescale of $\sim 0.15P_{orb}$, a condensation (blob) with variable rotational velocity forms in the residual disk. Figure 3 presents the time variations of the average density of matter passing through the *XZ* half-plane ($Y = 0$, $X > A$) splitting the disk. Against the background of the overall density decrease, we can see periodic oscillations due to the
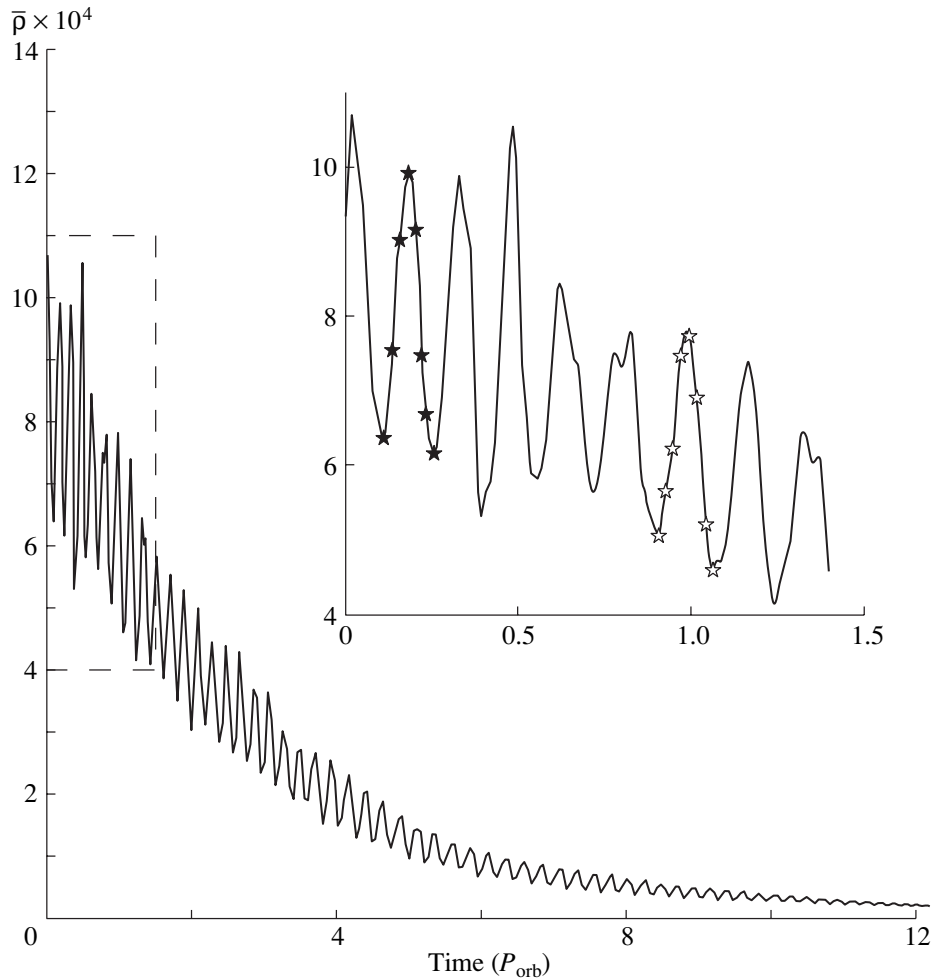
Fig. 3. Average density of matter passing through the XZ half-plane ($Y = 0$, $X > A$) splitting the disk as a function of time for case B. The inset displays a close-up view of the density variations in the time interval from mass-transfer termination to $t \sim 1.4P_{orb}$. Two sets of eight points each are plotted, corresponding to eight times covering one period of the density variations.

passage of the condensation. The blob does not spread under the action of dissipative processes; the periodicity of its rotation ($\sim 0.18P_{orb}$) is maintained essentially right up to the time when the disk disappears. A close–up view of the density variations from the time of mass-transfer termination up to $\sim 1.4P_{orb}$ is presented in the inset in Fig. 3. Here, two sets of eight points are plotted, corresponding to eight times, with each set entirely covering one period of the density variations. The first set corresponds to the time interval $0.11–0.26P_{orb}$ after t_0, and the second, to times $0.91–1.06P_{orb}$ after t_0.

Figure 4 presents the density distribution in the equatorial plane of the system for these two sets of times marked in the inset in Fig. 3. We can conclude from this figure that the condensation is maintained by interactions with the arms of the spiral shock. Let us consider the mechanism that brings about the existence of the blob. After it originates as a remnant of the stream, the condensation tends to spread over the disk uniformly under the action of dissipative processes and the differen-

tial rotation of the disk material. However, when the condensation passes through the spiral shock, the matter behind the wave decelerates, resulting in the formation and maintainance of a compact blob. Further, as the contrast between the densities of the blob and disk—and accordingly also the pressure gradient—increases, there comes a time when the blob begins to move through the disk. However, as the second spiral arm is being reached, matter again decelerates behind shock wave, and the formation and maintainance of the condensation recurs.

Figure 5 provides support for this mechanism. In Fig. 5a, the solid curve indicates the time variation of the velocity of the maximum density of the disk in the equatorial plane (i.e., the velocity of the blob), while the dashed curve indicates the time variation of the Keplerian velocity ($V_{Kep} = GM_1/r)^{1/2}$) for the point corresponding to the maximum density of the blob. The velocity variations are given for the time interval $0.91–1.11P_{orb}$, which includes the interval for the data in
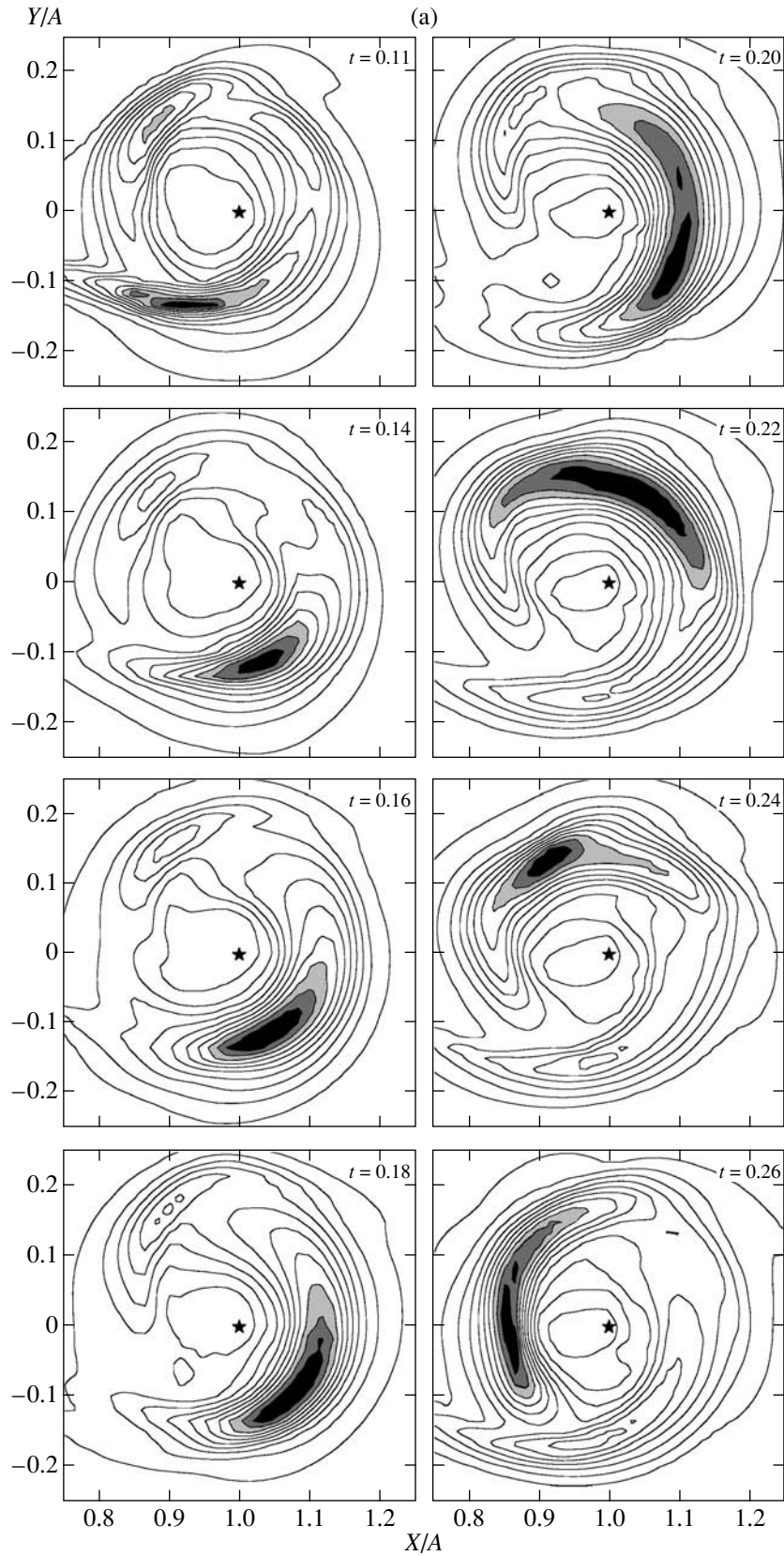
Fig. 4. Density distributions in the equatorial plane of the system for two sets of times marked with asterisks in the inset in Fig. 3: (a) data for the first set, corresponding to time ~$0.2P_{orb}$ after t_0, (b) data for the second set, corresponding to time ~$1P_{orbd}$ after t_0. The maximum densities for (a) and (b) are 0.0345 and 0.015, respectively. The maximum densities for isolines were defined as $\rho_{max}/11$.

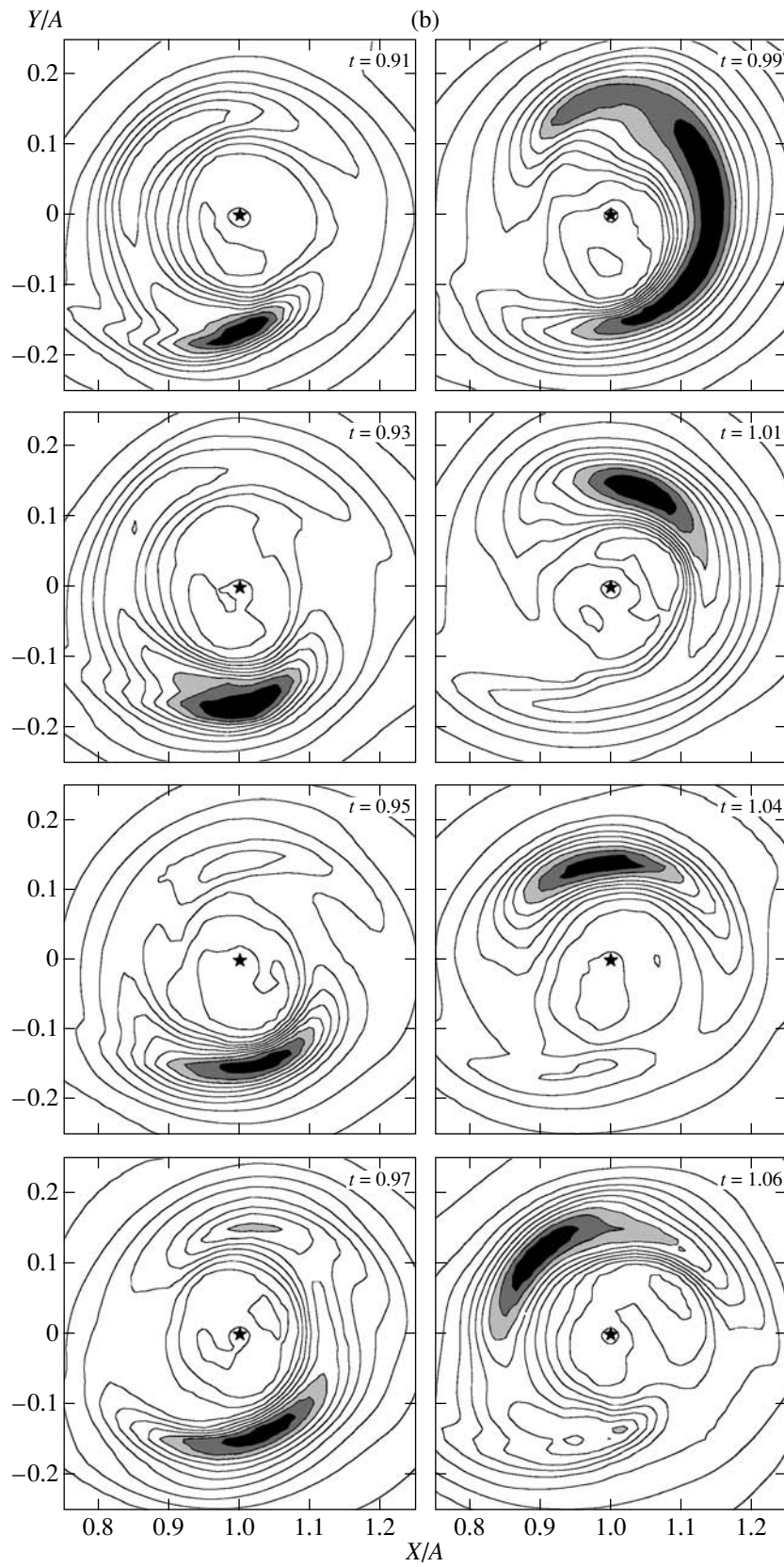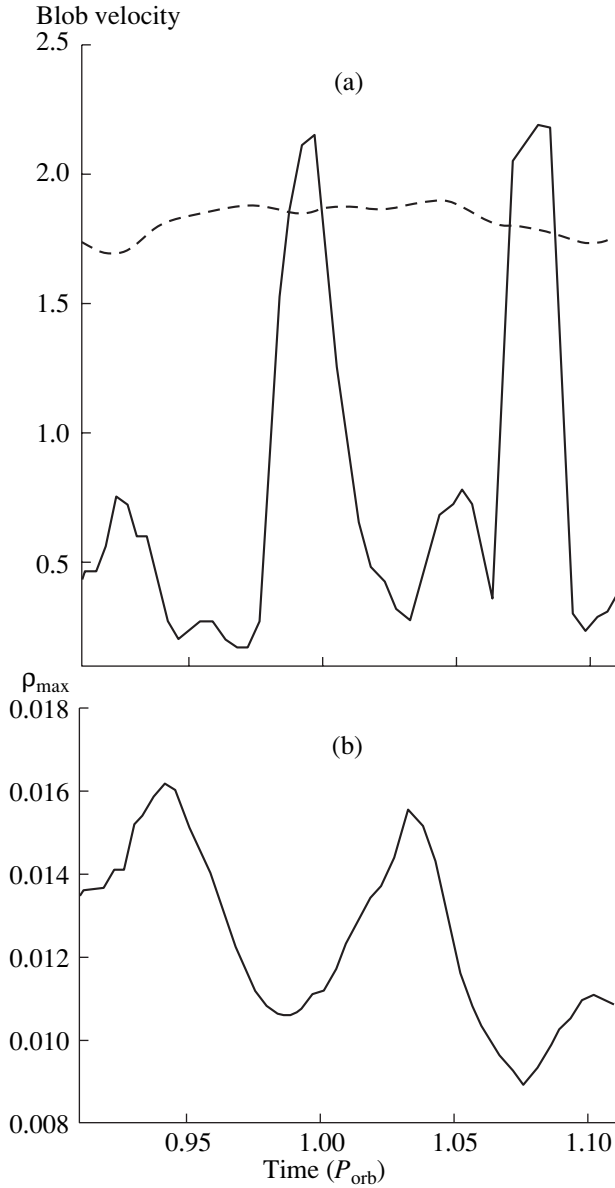**Fig. 4.** (Contd.)

Blob velocity

ρmax

Fig. 5. Time variations in the interval 0.91–$1.11 P_{\mathrm{orb}}$ for case B for: (a) the velocity of the blob (solid) and the Keplerian velocity of the point corresponding to the maximum density of the blob (dashed), (b) the maximum density of the blob.

Fig. 4b and covers the period of rotation of the blob. Figure 5b presents the time variation of the maximum density of the disk in the equatorial plane for the same interval. We can see from Fig. 5 that the blob condenses behind the shock wave where matter decelerates; the density of the blob $\rho_{\max}$ increases by a factor of ~1.5. Later, when a substantial fraction of the disk material has been accumulated into the blob, the latter breaks away from the shock under the action of the pressure; its velocity increases sharply and its maximum density decreases. After passing through the second spiral shock, the blob again decelerates and its density increases. Note that the density contrast between the blob and disk is main-

$\bar{\rho}_{\max}/\bar{\rho}_{\min}$
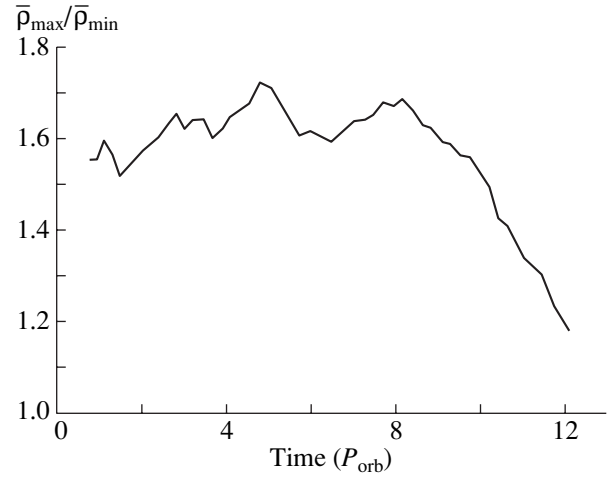
Fig. 6. Blob/disk density contrast as a function of time (case B).

tained right up to the disappearance of the condensation; i.e., until the time when the disk radius becomes so small that no waves remain in it. Further, the blob/disk density contrast begins to decrease: the blob spreads under the action of dissipative processes. However, it does not totally disappear, right up to the very disappearance of the disk. This is confirmed by Fig. 6, which presents the time dependence of the blob/disk density contrast.

4. CONCLUSIONS

We present the results of 3D simulations of the flow structures in semi-detached binaries after mass-transfer termination. We used previously obtained solutions for the steady-state mass-transfer rate as initial conditions; further, starting with the time $t = t_0$, the rate of mass transfer was decreased by five orders of magnitude, which computationally corresponds to its termination. To study the effect of viscosity on the solutions, we carried out three computations with viscosities (in α-disk terminology) corresponding roughly to $\alpha \sim 0.08$–0.1, 0.04–0.06, and 0.01–0.02.

Our study of the residual accretion disk indicates that its lifetime substantially increases as the viscosity decreases. Already when $\alpha \sim 0.05$, the lifetime of the disk exceeds $12 P_{\mathrm{orb}}$, and the lifetime of the residual disk is approximately $50 P_{\mathrm{orb}}$ for typical observed values of the viscosity ($\alpha \sim 0.01$).

Our study provides evidence that the flow pattern is already appreciably altered on time scales time $\sim 0.2 P_{\mathrm{orb}}$ after the termination of mass transfer. The initial stream from L_1 ceases to affect the solution; the shape of the accretion disk changes from quasi-elliptical to essentially circular. A second spiral shock wave arises in the disk structure, due to the gravitational action of the donor star; in systems with constant mass

transfer, this arm is suppressed by the stream from L_1. Despite the decrease of the disk mass with time, the spiral shocks exist in the disk essentially through its lifetime.

The numerical simulations considered here indicate that a condensation (blob) forms in the residual accretion disk, then moves through the disk with variable velocity. The blob is maintained by its interaction with the arms of the spiral shock, right up to their disappearance. The density contrast between the blob and disk is substantial (~1.6), and does not begin to decrease until the disappearance of the spiral shocks at ~$10P_{orb}$. A blob with density contrast relative to the disk material ~1.2 is maintained right up to the disappearance of the disk. This substantial blob/disk density contrast, as well the blob's variable velocity of rotation and approximately constant orbital period (~$0.18P_{orb}$), make this object extremely interesting from an observational point of view.

ACKNOWLEDGMENTS

REFERENCES

1. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **74**, 880 (1997) [Astron. Rep. **41**, 786 (1997)].

2. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **75**, 706 (1998) [Astron. Rep. **42**, 621 (1998)].

3. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).

4. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Astron. Zh. **76**, 905 (1999) [Astron. Rep. **43**, 797 (1999)].

5. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **77**, 31 (2000) [Astron. Rep. **44**, 26 (2000)].

6. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **75**, 40 (1998) [Astron. Rep. **42**, 33 (1998)].

7. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **78** (6), 2001 (in press) [Astron. Rep. **45** (6), 2001 (in press)].

8. J. R. Murray, B. Warner, and D. T. Wickramasinghe, Mon. Not. R. Astron. Soc. **315**, 707 (2000).

9. M. R. Schreiber, B. T. Gänsicke, and F. V. Hessman, Astron. Astrophys. **358**, 221 (2000).

10. H. Ritter, Astron. Astrophys. **202**, 93 (1988).

11. D. Molteni, G. Belvedere, and G. Lanzafame, Mon. Not. R. Astron. Soc. **249**, 748 (1991).

12. P. L. Roe, Ann. Rev. Fluid Mech. **18**, 337 (1986).

13. S. Chakravarthy and S. Osher, AIAA Pap., No. 85-0363 (1985).

14. B. Einfeldt, SIAM J. Numer. Anal. **25**, 294 (1988).

15. S. H. Lubow and F. H. Shu, Astrophys. J. **198**, 383 (1975).

16. D. Lynden-Bell and G. E. Pringle, Mon. Not. R. Astron. Soc. **168**, 603 (1974).

17. E. Meyer-Hofmeister and H. Ritter, in *The Realm of Interacting Binary Stars*, Ed. by J. Sahade, G. E. McCluskey, Jr., and Y. Condo (Kluwer, Dordrecht, 1993), p. 143.

18. C. Tout, in *Cataclysmic Variables and Related Objects*, Ed. by A. Evans and J. H. Wood (Kluwer, Dordrecht, 1996), p. 97.

19. P. J. Armitage and M. Livio, Astrophys. J. **470**, 1024 (1996).

20. K. Sawada, T. Matsuda, I. Inoue, and I. Hachisu, Mon. Not. R. Astron. Soc. **224**, 307 (1987).

21. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **219**, 75 (1986).

22. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **221**, 679 (1986).

Translated by K. Maslennikov

Formation of Low-Mass X-Ray Novae
with Black Holes from Triple Systems

A. G. Kuranov[1], K. A. Postnov[2], and M. E. Prokhorov[2]
[1]Physics Department, Moscow State University, Moscow, 119899 Russia
[2]Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899, Russia
Received January 24, 2000

Abstract—We apply a population synthesis technique to study the formation and evolution of low-mass X-ray binaries with black holes, observed as X-ray novae, from hierarchical triple systems. A scenario is suggested in which an inner close binary system evolves into an X-ray system with a large mass ratio. The high rate of accretion onto the neutron star leads to a common envelope stage, which may result in the formation of a Thorne–Zytkow (TZ) object. During its evolution, the envelope of the TZ object expands, encompassing the third star. The recurrent common-envelope stage decreases the size of the orbit of the third star, leading to the formation of a low-mass X-ray nova with a black hole. The dynamical stability of triple systems automatically ensures that only low-mass X-ray novae form. We also consider the possible formation of an X-ray nova from a binary in the case of asymmetrical core collapse during a supernova explosion. © 2001 MAIK "Nauka/Interperiodica".

1. INTRODUCTION

The search for black holes (BHs) is a fundamental problem of contemporary astronomy. In their pioneering studies, Zel'dovich [1] and Zel'dovich and Guseĭnov [2] showed that the most significant observational manifestations of black holes should be expected in close binaries, during accretion onto the gravitating center. Disk accretion onto compact objects was studied quantitatively by Shakura [3] and Shakura and Sunyaev [4]. These concepts have been dramatically confirmed by numerous observations of X-ray sources in binaries, starting in 1972 with the launch of the specialized UHURU satellite.

Despite appreciable progress in the theory of the internal composition and evolution of stars, the processes involved in the formation of stellar-mass black holes remain incompletely understood. Black holes are assumed to originate during the final stages of evolution of single stars with masses >20–$40 M_\odot$ [5]. Details of the process remain poorly studied; it is clear, however, that the nuclear evolution of massive stars should be accompanied by dramatic mass losses. According to the calculations of [6], the mass of a supernova progenitor does not exceed 12–$15 M_\odot$ and depends only weakly on the initial mass of the star on the main sequence. Another parameter of black-hole formation during stellar core collapse is the core mass. The iron cores of stars at the end of their evolution have masses ~ 2–$2.5 M_\odot$ [5], nearly independent of the initial mass of the star on the main sequence. The observed masses for known black-hole candidates are around $8 M_\odot$ [7–9].

The first observations of a specific class of transient X-ray sources, so-called X-ray novae, were made at the end of the 1970s. By 2000, more than 30 such sources had been discovered [9, 10]. These systems are characterized by a large mass function for the invisible component and the absence of short-period outburst activity, similar to X-ray bursters. It is generally accepted that the compact components of X-ray novae are the most promising black-hole candidates. In seven of nine cases, the optical component is a low-mass star of late spectral type [9]. Attempts to explain the origin of such systems in the standard theory of binary evolution face a number of serious problems.

First and foremost, the binary must have an extreme initial mass ratio $q = M_2/M_1 \ll 1$, which is not directly observed (the component mass-ratio distribution for binaries is reviewed, for example, in [11]). Such a binary can lose more than half of its total mass during the supernova explosion, so that the probability of disruption of the system is high. The mechanism for the decrease of the orbital radius of the system right to the point of Roche-lobe filling by the low-mass component also encounters numerous problems. "Fine tuning" of the parameters and initial conditions is needed in order for the binary to evolve into a low-mass X-ray nova. These problems are considered in detail in [12].

At the same time, a substantial fraction of stellar systems display multiplicity no lower than three (according to various estimates, from 5 to 30%; see, for example, the discussion in [13]). Essentially all observed multiple systems are hierarchical [14]: a close binary frequently possesses a distant component, which, in turn, may be a member of another close binary.

A scenario for the formation of a low-mass X-ray nova with a black hole from a triple system consisting of a massive close binary and a distant low-mass com-

ponent—($15M_\odot$ + $10M_\odot$) + $1M_\odot$—was suggested in 1986 in [15]. One advantage of this scenario is that an initially large mass ratio in the binary is no longer required. In this scenario, the inner close binary evolves as a regular binary star. The primary fills its Roche lobe and the first stage of mass transfer begins. Further, the primary explodes as a supernova and a compact remnant is formed—a neutron star or black hole. When, in turn, the secondary star fills its Roche lobe, catastrophic mass transfer occurs due to the large mass ratio of the binary components, possibly resulting in a common-envelope (CE) stage. The neutron star approaches the center of the supergiant along a spiral path due to deceleration in its envelope.

The possibility that a neutron star could be located at the center of an ordinary star was first considered in the pioneering study by Landau [16] (though not in the context of the evolution of binary systems). It was not until the middle of the 1970s, however, that the first quantitative calculations for this configuration appeared [17, 18], after which the term "Thorne–Zytkow (TZ) object" became generally accepted.

The evolution of such objects is not yet known with certainty due to the complexity of the physical processes involved. Hydrodynamical computations (see, for example, [19–21]) yield only very general (and often directly contradictory) concepts about their fate. For example, it is asserted in [21] that a considerable fraction of energy may be taken away by neutrinos during the accretion of the envelope onto the neutron star, so that a TZ object cannot have a stable phase.

If, on the contrary, there is a regime of gradual subsidence of the envelope material onto the neutron star, the duration of the TZ object stage is estimated to be 10^4–10^6 yr [19, 20] and is specified by the mass-loss rate. From an astronomical point of view, a TZ object should differ only slightly from a supergiant of the same mass. It will possess a radius on the order of $1000R_\odot$ and a mass-loss rate of 10^{-5}–$10^{-6}M_\odot$/yr [19]. The final result of the TZ stage could be either a single neutron star or a black hole with a mass of several solar masses, depending on the regime of accretion onto the neutron star. Thus, in this scenario, the evolution of a massive binary could end with the formation of a black hole.

The formation of a disk around the black hole from the remnants of the ejected envelope is not ruled out; gravitational instability in the disk could lead to the formation of low-mass dense objects (planets and even low-mass stars). In [19], an attempt is made to explain the formation of low-mass X-ray novae from binaries passing through a TZ-object stage in this way.

In spite of the above uncertainties, we will assume below a regime of gradual subsidence of matter onto the neutron star, with subsequent collapse into a black hole.

In a triple system, the presence of a distant component does not significantly affect the evolution of the inner binary up until the formation of the TZ object (the role of Kozai cycles in tidal interactions in close binaries is considered in [22]). If, however, the low-mass star is encompassed by the expanding envelope of the TZ object (whose size is around $1000R_\odot$, while the third star may have a large orbital eccentricity, so that it passes through the envelope at periastron), a second common-envelope stage is unavoidable. The orbital momentum of the third star is partially transferred to the common envelope via dynamical friction; however, unlike the previous common-envelope stage, total coalescence is unlikely, due to the possible decrease of deceleration efficiency in the more rarified upper layers of the envelope and the initially large orbit of the third star.

As a consequence, after depletion of the TZ-object envelope, the initial triple system becomes binary, consisting of a black hole and the now closer low-mass component. Gravitational radiation and magnetic stellar wind will continue to decrease the size of the system's orbit. If the orbit of the low-mass component decreases enough during its lifetime on the main sequence for its Roche lobe to be filled, an accretion disk is formed around the black hole, providing an explanation for the observed X-ray novae [23]. We will assume that any stable, semidetached, low-mass binary with a black hole may be manifest as an X-ray nova. The problem of comparing our results with the observed number of X-ray novae in the Galaxy is considered separately in Section 3.1.

This scenario for the formation of X-ray novae is analogous to the standard scenario of Paczynski [24] for the formation of cataclysmic variables, in which the common-envelope stage is experienced by a binary system consisting of a giant and a white dwarf.

Here, we will apply a population-synthesis method to compute the number of X-ray novae—candidate black holes—in the Galaxy, as well as their mass and spatial-velocity distributions as functions of parameters in the evolutionary scenario.

2. METHODS AND INITIAL CONDITIONS

Our approach to studying the evolution of hierarchical triple systems is based on our modification of the "Scenario Machine" code developed in the Department of Relativistic Astrophysics of the Sternberg Astronomical Institute for binary population syntheses. The method is described in detail in [25]; we will discuss only a few important points below.

2.1. Initial Conditions

The initial distribution of the semimajor axes for the binary orbits is consistent with a uniform logarithmic distribution [26, 27]:

$$f(\log a) = \text{const}, \quad \max \begin{cases} 10R_\odot \\ R_{\text{L}}(M_1) \end{cases} < a < 10^7 R_\odot. \quad (1)$$

Here, $R_L(M_1)$ is the Roche-lobe radius for the primary. We used a Salpeter distribution (for the more massive star) to determine the mass of the stars:

$$f(M_1) \propto M_1^{-2.35}, \quad 0.1 M_\odot < M_1 < 120 M_\odot. \quad (2)$$

This distribution predicts the formation of a massive star in a binary ($M_1 > 10 M_\odot$ for the formation of a compact remnant) once every 60 yr. This estimate is consistent with the observed rate of binary-system formation [30].

We used a power-law distribution for the mass ratio of the components:

$$f(q) \propto q^{\alpha_q}, \quad q = M_2/M_1 < 1, \quad (3)$$

where α_q is a parameter of the distribution.

The lack of sufficient statistical data currently prevents determination of the distributions of the mass and orbital parameters for the distant component in the triple system. It is natural to assume that, since the distant component forms relatively independently of the close binary, it should display an initial mass distribution similar to that for single stars. By analogy with the binary, we used the following distributions for the third star:

$$f(\log a_3) = \text{const}, \quad \max\begin{cases} a_3^{\text{stab}} \\ R_L(M_3) \end{cases} < a < 10^7 R_\odot, \quad (4)$$

where a_3^{stab} is the minimum semi-major orbital axis for the third star for which the triple system remains dynamically stable [specified by condition (7) below], and

$$f(M_3) \propto M_3^{-2.35}, \quad 0.1 M_\odot < M_3 < 120 M_\odot. \quad (5)$$

Observations indicate that the distribution of orbital eccentricities can be represented by the function [28]

$$f(e) \propto e. \quad (6)$$

We used this distribution to randomly select the initial orbital eccentricities for the third star.

A fairly simple and general criterion for the stability of a triple system was derived in [29]:

$$\frac{r_{\text{out}}^{\text{per}}}{r_{\text{in}}^{\text{ap}}} > F^{\min}(M_i), \quad (7)$$

$$F^{\min} = 1 + \frac{3.7}{q_{\text{out}}^{1/3}} - \frac{2.2}{1 + q_{\text{out}}^{1/3}} + \frac{1.4}{q_{\text{in}}^{1/3}} \frac{q_{\text{out}}^{1/3} - 1}{q_{\text{out}}^{1/3} + 1}, \quad (8)$$

where $r_{\text{out}}^{\text{per}} = a_3(1 - e_3)$ and $q_{\text{in}}^{\text{ap}} = a(1 + e)$ are the pericenter distance of the outer orbit and the apocenter distance of the inner orbit, respectively; $q_{\text{in}} = M_1/M_2 > 1$, and $q_{\text{out}} = (M_1 + M_2)/M_3$. The impact on the stability of the system by the fact that the outer and inner orbits are not coplanar does not exceed 20%, which is smaller

than the computational errors in [29]. We used this criterion (in a point-mass approximation) in our computations.

Each evolutionary track consists of a finite number of stages, described in detail in [25]. Condition (7) for stability of the triple system is checked at each stage. If this condition is violated, the distant component is assumed to no longer be gravitationally bound to the close binary system.

The evolution of the systems is computed up to some maximum time $T_{\max}$ ($T_{\max} = 10^{10}$ yr). For each stage i, we find the number of systems S_i that have gone through this stage, and calculate the sums $\sum \Delta f_i$ and $\sum \Delta f_i^2$, where

$$\Delta f_i \equiv \int_{t_1}^{t_2} \varphi_t dT \approx \varphi_t(t_1)\Delta t. \quad (9)$$

Here, φ_t is the probability of the birth of a system at time T (the normalized star-formation rate), t_1 and t_2 are the times corresponding to the beginning and the end of the ith stage, and $\Delta t = t_2 - t_1$. Then, the total number of stars in the ith stage is

$$N_i = N_*(\sum \Delta f_i \pm \alpha \sigma(f_i)),$$

$$\sigma(f_i) = \left(\frac{D(f_i)}{S_i - 1}\right)^{1/2}, \quad (10)$$

$$D(f_i) = (\sum \Delta f_i^2 - (\sum \Delta f_i)^2)/S_i.$$

Here, N_* is the number of stars in the Galaxy and α is a coefficient specified by the confidence level.

In the computations, we used a model of the Galaxy in which the total mass of stars was $10^{11} M_\odot$; the fraction of binaries is 50%, and triple systems and stars of higher multiplicity make up 10% of the total number of stellar systems. In general, the birth rate of systems and the distribution of their initial parameters could be time dependent. We consider only time-independent evolutionary scenarios ($\varphi_t = \text{const}$).

2.2. Common-Envelope Stage

A detailed description of the common-envelope stage requires complicated, three-dimensional, hydrodynamical computations (see, for example, [31]). Since the details of this stage are poorly understood, we describe it using an empirical energy relation containing the parameter α_{ce}—the efficiency of transfer of released gravitational binding energy of the binary to the common envelope. We specify this parameter in accordance with [32]:

$$\alpha_{\text{ce}}\left(\frac{GM_a M_c}{2a_f} - \frac{GM_a M_d}{2a_i}\right) = \frac{GM_d(M_d - M_c)}{R_d}, \quad (11)$$

where M_d and R_d are the initial mass and radius of the star that loses the envelope, M_c the mass of its core remaining after the loss of the envelope, M_a the mass of the secondary component, and a_i and a_f the semi-major axes of the system before and after the common-envelope stage, respectively.

2.3. Parameters of Black Holes Formed

Black holes are currently thought to form in the course of stellar evolution, provided the mass of a star on the main sequence exceeds some critical value $M > M_*$ ($M_* \sim 40 M_\odot$). The mass of a forming black hole is an important parameter; we assumed it to be proportional to the mass of the supernova progenitor M_{pr}:

$$M_{BH} = k_{BH} M_{pr}, \quad 0 < k_{BH} < 1. \quad (12)$$

We assume that the proportionality factor is the same for all stars. M_{pr} is essentially the mass function of the star on the main sequence [33].

Another way for black holes to form is the accretion-induced collapse of a neutron star when its mass exceeds the Oppenheimer–Volkov limit M_{OV} as a result of accretion. This can occur in the TZ-object stage. In this case, we assumed that some fraction of the envelope ends up inside the black hole:

$$M_{BH} = M_{OV} + k_{env} M_{env}, \quad 0 < k_{env} < 1, \quad (13)$$

where $M_{env} = M_{TZ} - M_{OV}$ is the mass of the TZ-object envelope and M_{TZ} the mass of the TZ object itself. The results presented below were obtained assuming $M_{OV} = 2.5 M_\odot$.

2.4. Asymmetric Core Collapse

The distribution of the additional momentum gained by the compact remnant (a neutron star or black hole) in the case of asymmetric core collapse of a massive star is an important parameter in contemporary models of stellar evolution. The initial idea that the collapse could be asymmetric belongs to Shklovskiĭ [34]. Currently, the presence of asymmetry in the formation of neutron stars can be considered an established fact [35]. The distribution of the modulus of the velocity w_{NS} acquired by a neutron star during core collapse can be derived from current data for the velocities of radio pulsars [36]:

$$f(x) \propto \frac{x^{0.19}}{\left(1 + x^{6.72}\right)^{1/2}}, \quad (14)$$

where $x = w_{NS}/w_0$ and w_0 is a parameter (characteristic velocity), which yields the best consistency with observations when $w_0 = 400$ km/s [25].

In our computations, we also used a Maxwellian distribution for the velocity anisotropies:

$$f_m(w) \propto w^2 \exp(-w^2/w_0^2), \quad (15)$$

where w_0 is a parameter related to the mean velocity w_m via the expression $w_m = \dfrac{2}{\sqrt{\pi}} w_0$. The computations indicate that the number and parameters of forming X-ray novae are primarily influenced by the mean anisotropic velocity w_0 rather than the form of the velocity-modulus distribution.

Theoretically, the formation of a black hole could also be asymmetric. This occurs, for example, if the core collapse of a supernova progenitor undergoes two stages: it initially becomes a hot neutron star and then, after cooling, forms a black hole. In this case, the black hole will also gain an additional spatial velocity. We will assume that the additional velocities of the newly formed black holes are distributed in accordance with the same law as that for the neutron star, but with a different amplitude:

$$w_{BH} = w_{NS} \frac{1 - k_{BH}}{1 - \dfrac{M_{OV}}{M_{pr}}}. \quad (16)$$

Using this law, we obtain $w_{BH} = 0$ for $k_{BH} = 1$ (i.e., when the entire mass of the progenitor collapses into the black hole) and $w_{BH} = w_{NS}$ in the limit $M_{BH} = M_{OV}$. For example, for a typical progenitor mass $M_{pr} \sim 15 M_\odot$ and $k_{BH} = 0.5$, we obtain $w_{BH} \sim 0.6\, w_{NS}$. It is apparent that there is no additional momentum gain when a black hole forms from a TZ object.

2.5. A Thorne–Zytkow Object

In our computations, we assume that, from an astronomical point of view, a TZ object differs only slightly from a supergiant of the same mass [19]. According to de Jager [37], the luminosity of the supergiant L is

$$\log \frac{L}{L_\odot} = \begin{cases} 10^{-4.462 + 3.8 \log M/M_\odot}, & M < 23.7 M_\odot \\ 10^{-3.362 + 3.8 \log M/M_\odot}, & 48 M_\odot > M > 23.7 M_\odot \\ 10^{-3.362 + 2.7 \log M/M_\odot}, & M > 48 M_\odot. \end{cases} \quad (17)$$

The maximum radius of a TZ object is determined from its effective temperature (T_{eff}) and luminosity. We used the following relations to calculate T_{eff} [38]:

$$\log T_{eff} = \begin{cases} 4.5, & M \geq 10 M_\odot \\ 3.6, & M < 10 M_\odot. \end{cases} \quad (18)$$

This determination of the radius of the TZ object is fairly formal. Note, however, that, for the chosen initial distribution of orbital semi-axes for the third star (4), the maximum radius of the TZ object, which determines the onset of the common envelope stage, can vary over a broad range without considerably influencing the result.

3. CALCULATION OF THE NUMBER OF X-RAY NOVAE FORMED FROM BINARY AND TRIPLE SYSTEMS

3.1. Estimate of Observed Number of X-ray Novae in the Galaxy

Sufficient statistical data have been accumulated to estimate the number of X-ray novae—candidate black holes—in the Galaxy. According to [39], the observed rate of X-ray novae outbursts is $\mathcal{R} \sim 2.6$ events per year; of these, there are $\mathcal{R}_{NS} \sim 1.1$ events per year in systems with neutron stars and $\mathcal{R}_{BH} \sim 1.5$ in systems with black holes. The observed number of X-ray novae—candidate black holes—is $N_{obs}(BH) \sim 10$; therefore, if the average time between X-ray outbursts in a system is $\delta t_x = 50$ yr, the total number of these systems existing simultaneously in the Galaxy is

$$N_G(BH) \sim N_{obs}(BH)\delta t_x \mathcal{R}_{BH} \sim 800. \qquad (19)$$

Note that these estimates depend appreciably on the time interval between consecutive X-ray outbursts in a system. According to current views [10], this interval may be from several to ~50 yr; in other words, the above estimate of the total number of X-ray novae in the Galaxy is correct to an order of magnitude. Tanaka [40] estimates the total number of X-ray novae with black holes in the Galaxy to be 600–3000 systems. Later, based on observational data for the local density of these systems in the vicinity of the Sun, White and van Paradijs [41] concluded that there should be 500 such systems in the Galaxy. Here, we do not consider possible observational selection effects. However, taking into account the brightness of X-ray novae outbursts (on order of the Eddington luminosity), we assume that the impact of observational selection effects is substantially smaller than the uncertainty associated with the inaccurately known time interval between consecutive X-ray outbursts in a system.

The dispersion of the spatial velocities of the centers of mass of candidate black holes in the Galaxy can be derived from observational data for their z distribution. Brandt *et al.* [42] give the value 22 ± 9 km/s. Later [41], a higher value of 42 ± 8 km/s was obtained, which exceeds the velocity dispersion for stars in the Galactic disk. An independent statistical study of X-ray novae based on monitoring of the Galactic center using the SIGMA telescope (GRANAT satellite observatory) [43] also suggests that the spatial distribution of X-ray novae differs from that of the Galactic disk component, providing additional evidence for a relatively high spatial-velocity dispersion for these objects.

3.2. Formation of X-ray Novae from Binaries

In this section, we present the results of our computations for the formation of X-ray novae with black holes from binary systems in the case of asymmetric core collapse, carried out using the Scenario Machine.

The total number of low-mass X-ray novae formed from binary systems depends critically on the distribution of the binary mass ratio α_q [see (16)]. The distribution of binary-component mass ratios has been investigated in numerous studies (see, for example, the review [11] and references therein). Note, however, that the observed distributions have a break at $q = 0.1$, since systems with smaller q values are essentially not observed. Since the observed distributions are subject to numerous selection effects that primarily lead to underestimation of the number of systems with dramatically different component masses, the problem of the component-mass distribution for $q < 0.1$ remains unsolved.

The possible formation of low-mass X-ray novae with black holes from binaries with extreme mass ratios was studied in detail in [12]. In the course of its evolution, a massive star undergoes a supernova explosion. In the case of a symmetrical core collapse, when the orbit of the supernova is circular, the eccentricity and orbital period become [44]

$$e_{postSN} = \frac{\Delta M}{M_{BH} + M_2},$$

$$P_{postSN} = P_i \frac{\mu}{(2\mu - 1)^{2/3}},$$

where $\mu = \dfrac{M_{BH} + M_2}{M_{He} + M_2} = \dfrac{M_{He} + M_2 - \Delta M}{M_{He} + M_2}$, M_{He} is the mass of the helium star immediately before the supernova, M_{BH} is the mass of the black hole formed, $\Delta M = M_{He} - M_{BH} < 0.5(M_{He} + M_2)$, and the index i indicates parameters of the system before the explosion.

Since all observed X-ray novae are short-period binaries ($P \le 7$ days), tidal interactions rapidly circularize the orbits in these systems. Therefore,

$$P_{re\_circ} = P_{postSN}(1 - e_{postSN}^2)^{3/2} = P_i/\mu^2. \qquad (20)$$

Since $\mu < 1$, the period of the binary P_i before the supernova should be fairly small. This means that the binary should necessarily have undergone a common-envelope stage, but avoided total coalescence.

Therefore, certain conditions must be satisfied for the formation of an X-ray nova from a binary system. The most important of these are as follows.

(1) The binary must have an initial mass ratio $q = M_2/M_1 \ll 1$.

(2) The semi-axis of the binary orbit must be sufficiently small that the massive component filled its Roche lobe and the common-envelope stage began. This requirement imposes restriction on the maximum initial semi-axis of the system.

(3) The binary must avoid coalescence in the common-envelope stage. This depends on a number of parameters, such as the common-envelope efficiency, system semi-axis, mass of the ejected envelope, etc.

(4) The binary must remain gravitationally bound after the supernova explosion. In the case of a symmetrical collapse, this requirement imposes restrictions on the mass ejected during the supernova: $\Delta M < 0.5(M_{He} + M_2)$.

(5) After the explosion, the semi-major orbital axis of the binary must be sufficiently small so that the component (dwarf star) fills its Roche lobe over a time shorter than the age of the Universe. Roche-lobe filling can occur due to either a decrease in the binding energy of the binary owing to energy losses via gravitational radiation or magnetic stellar wind, or to an increase in the radius of the star itself in the course of its evolution.

None of these conditions can be satisfied without very accurate fine tuning of parameters and initial conditions in the binary.

In [12], the number of low-mass X-ray novae (mass of the optical component $<1M_\odot$) formed in this way was compared to the number of X-ray novae of moderate mass (mass of the optical component $>1M_\odot$), taking into account evolutionary parameters. When $\alpha_q = 0$ (a flat distribution), the number of massive X-ray novae exceeds the number of low-mass novae for any reasonable parameters for the formation of the X-ray novae. These numbers become comparable only when $\alpha_q < 0$ ($\alpha_q = -2.7$). In other words, to explain the observations in a scenario in which X-ray novae form from binaries (seven of nine X-ray novae—candidate black holes—have optical components with masses $<1M_\odot$), we must assume that most forming binaries posses extreme mass ratios, which is not confirmed by observations and results in an obvious overestimation of the number of X-ray sources (see [45] and below).

A number of other difficulties are also encountered here. The recent study [6] presents new computations of the evolution of massive stars. In particular, in the course of a massive star's evolution on the main sequence, its mass-loss rate becomes so high that it is unable to fill its Roche lobe if $M > 40M_\odot$. Accordingly, the common-envelope stage does not occur. However, when the compact remnant gains additional velocity during the collapse, the system's orbital size can decrease sharply, making it possible for an X-ray nova to form without a substantial increase in the radius of the massive star on the main sequence.

Using the initial distributions for the semi-axes (1) and the masses (2) and (3), we have estimated the total number and the mean spatial velocity of X-ray novae in the Galaxy for various additional velocities and slopes of the initial distributions from the component mass ratio q. We carried out computations for the case of a high rate of mass loss from massive stars in the course of their evolution, in accordance with [6].

The table presents the results for $\alpha_q = 0$ (a flat distribution), $\alpha_q = 1$, and $\alpha_q = 2$ for two types of X-ray novae: low-mass systems (mass of the optical component $M < 1M_\odot$; Roche-lobe filling occurs due to the decrease of the binding energy of the binary owing to energy losses

Number of X-ray novae with black holes in the Galaxy (total mass of stars $10^{11} M_\odot$, fraction of binaries 50%) originating from binaries with large mass-loss ratios during asymmetrical core collapse, for various indices of the initial mass-ratio distribution α_q

w_{NS}, km/s	Total number of X-ray novae in the Galaxy		Spatial velocity of the center of mass
	$M_{opt} < 1M_\odot$	$M_{opt} > 1M_\odot$	V_{syst}, km/s
$\alpha_q = 0$			
0	0	0	–
50	57	4400	35 ± 15
100	110	4700	38 ± 16
200	250	10000	50 ± 19
300	460	8900	61 ± 22
400	310	7300	72 ± 24
500	100	5000	85 ± 28
600	20	3200	98 ± 12
$\alpha_q = 1$			
0	0	0	–
50	~1	1300	40 ± 15
100	~4	5000	36 ± 15
200	~10	3800	46 ± 18
300	~20	2700	60 ± 20
400	~10	2500	70 ± 22
500	~10	1800	78 ± 23
600	~1	1500	98 ± 12
$\alpha_q = 2$			
0	–	0	–
50	–	200	34 ± 16
100	–	300	35 ± 14
200	–	1900	46 ± 18
300	–	1400	60 ± 20
400	–	1200	70 ± 20
500	–	1000	78 ± 20
600	–	950	98 ± 12

to gravitational radiation and magnetic stellar wind) and moderate-mass systems $M > 1M_\odot$ (the optical component fills its Roche lobe due to its expansion in the course of its evolution).

M_1		M_2	A	E	T
40.67		18.41	120	0	5.5
35.09		18.29	130	0	5.9
20.68	WR	18.28	180	0	6.0
	SNIb				
6.99	BH	18.28	160	0.79	6.0
6.99	BH X Nova	5.9	200	0.63	11.0
7.00	BH WR	5.31	75	0.0	21.0
	BH SNIb				
7.00	BH NS "E"	1.4	Destruction		22.0

Fig. 1. Example of an evolutionary track resulting in the formation of a massive ($M > 1 M_\odot$) X-ray nova with a black hole from a binary system. M_i is the mass of the stars in solar masses, A the semi-major axis of the system's orbit in solar radii, e the eccentricity, T the time since the formation of the binary (in units of 10^6 yr), $k_{BH} = 0.5$, $M_{OV} = 2.5 M_\odot$, $w_{NS} = 300$ km/s. "WR" indicates the Wolf–Rayet star and "X Nova" the X-ray nova.
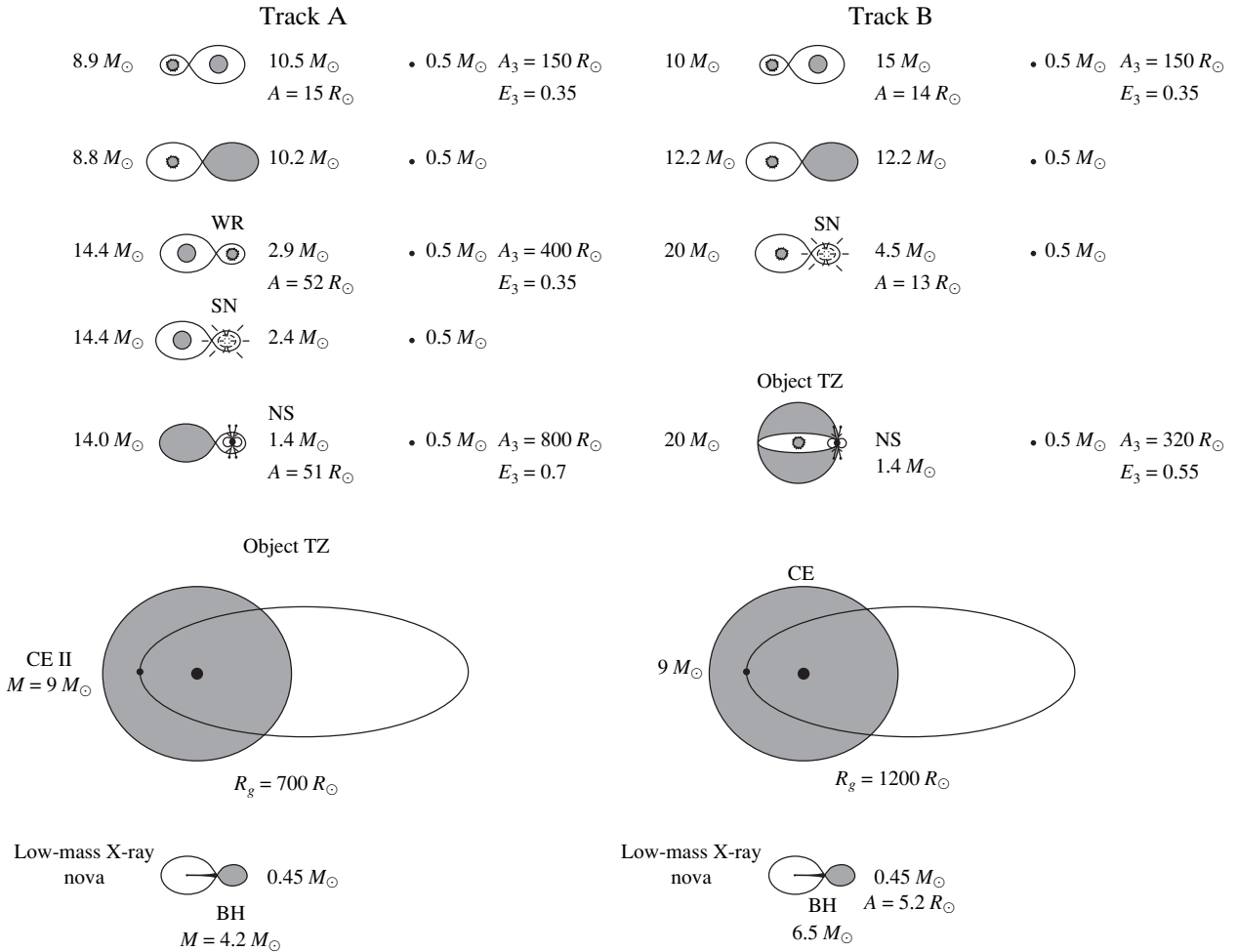
Fig. 2. Examples of evolutionary tracks resulting in the formation of low-mass X-ray novae from hierarchical triple systems. WR denotes the Wolf–Rayet star, $k_{env} = 0.5$, $M_{OV} = 2.5 M_\odot$, $w_{NS} = 300$ km/s, $\alpha_{ce} = 0.5$. Track A corresponds to the scenario of Eggleton and Verbunt [15]; track B is possible only if an additional velocity is gained during the asymmetrical explosion of the supernova (see text).

For any $\alpha_q \geq 0$, the number of X-ray novae of moderate mass formed substantially exceeds the number of low-mass X-ray novae. As α_q increases, the number of X-ray novae of moderate mass decreases, but remains larger than the observational estimate ~200, while the number of low-mass X-ray novae becomes vanishingly small. When $\alpha_q = 2$, low-mass X-ray novae essentially do not form from binary stars.

It was shown in [46] that it is not necessary to invoke additional velocity gained by the compact remnant during the supernova explosion to explain the high spatial velocity of the system GRO J1655–40 (XN Sco = 1994 Cyg), with $V_{syst} = 106 \pm 19$ km/s, assuming the binary underwent a common-envelope stage in the course of its evolution. If no common-envelope stage occurs, the orbital velocities of the stars will be substantially smaller than they would be after the common-envelope stage. In this case, it becomes difficult to explain the high spatial velocity without assuming an anisotropic collapse. In addition, it follows from the Table that X-ray novae do not form when $w_{ns} = 0$.

The formation of massive X-ray novae is illustrated by the track presented in Fig. 1, which is typical of the tracks obtained in our computations.
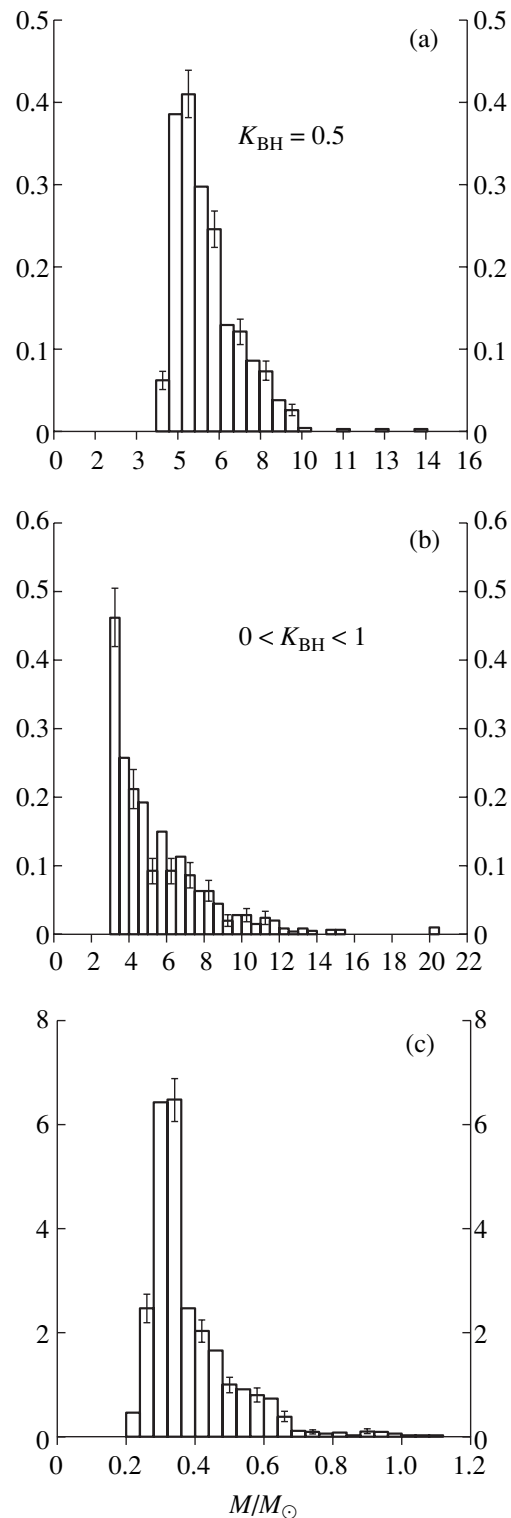
3.3. X-ray Novae from Triple Systems

The computations revealed two basic ways to form X-ray novae from triple systems. In the first (track A in Fig. 2), the evolution follows the above scenario [15]. A TZ object forms due to accretion in a close binary system with a large mass ratio, resulting in catastrophic mass transfer and a common-envelope stage.

The other method (track B in Fig. 2) is possible only when a neutron star gains additional velocity due to asymmetric core collapse of the supernova progenitor. In this case, under appropriate conditions, a direct collision of the neutron star with its companion is possible, resulting in coalescence of the stars and, unlike the scenario [15], the TZ object forms just after the supernova explosion. The number of systems originating in this way is relatively small; it increases from 0% (for $w_{ns} = 0$) to ~5% (for $w_{ns} = 500$ km/s) of the total number of X-ray novae formed. In both cases, the close binary undergoes a stage with a TZ object, whose expanding envelope subsequently encompasses the distant component of the triple system.

Figure 3 presents the mass distribution for the components of the X-ray novae that are formed. The mass of the black hole depends on the Oppenheimer–Volkov limit and the parameter k_{env} (13). Figure 3a is plotted for $M_{OV} = 2.5M_\odot$ and $k_{env} = 0.5$. Figure 3b corresponds to $M_{OV} = 2.5M_\odot$ and a uniform k_{env} distribution in the interval 0–1. In both cases, the average mass of the black holes formed is about $\langle M_{bh} \rangle \sim 5.5M_\odot$.

Figure 3c presents the mass distribution for the black-hole companions. Note that, in spite of the broad
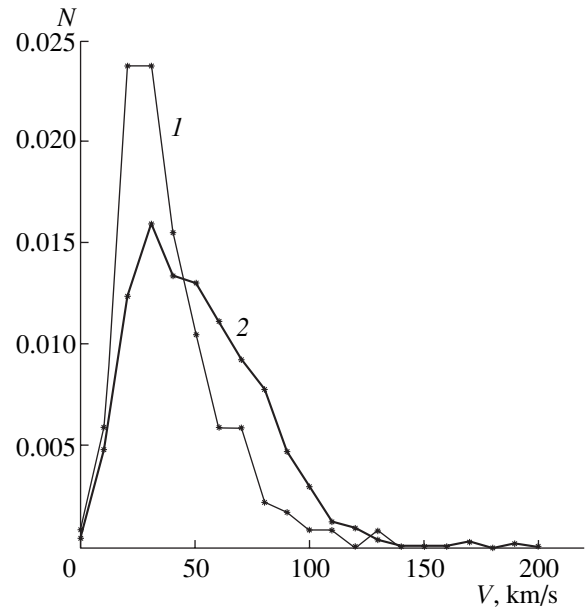

Fig. 3. Component mass distribution for X-ray novae originating from triple systems. (a) Black-hole mass for $M_{OV} = 2.5M_\odot$, $k_{env} = 0.5$; (b) black-hole mass for $M_{OV} = 2.5M_\odot$ and a uniform distribution of k_{env} in the interval from 0 to 1. (In both cases, the average mass of the black holes formed is about $\langle M_{bh} \rangle = 5.5M_\odot$.) (c) Mass distribution for the optical components in X-ray novae originating from triple systems. The error bars reflect the statistical accuracy.
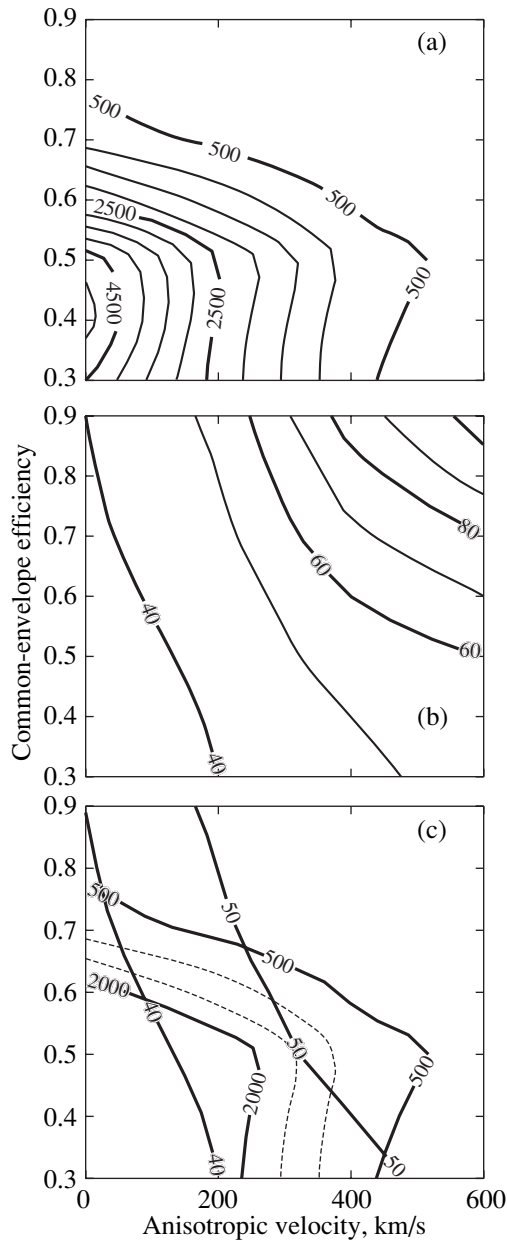
Fig. 4. Area of parameters corresponding to observational data for the number of low-mass X-ray novae (~500) and their velocities (~40 km/s) [41] in the Galaxy (total mass of stars 10^{11} $M_\odot$, fraction of triple systems 10%). (a) Isolines of the numbers of X-ray novae [$k_{env} = 0.5$, $M_{pr} = 3M_\odot$, $w_{BH} \sim 0.6w_{NS}$; see (16)] and (b) isolines of the mean velocity of the X-ray nova's center of mass in km/s as functions of the mean anisotropic velocity w_{NS} and the common-envelope efficiency α_{ce}; (c) individual isolines from the above two graphs.

Fig. 5. Distributions of the modulus of the spatial velocity of the centers of mass of X-ray novae originating from triple systems ($\alpha_{ce} = 0.6$, $w_{NS} = 200$ km/s) for an anisotropic velocity distribution (14) (curve *1*; $\langle V_{syst} \rangle = 54 \pm 25$ km/s) and a Maxwellian distribution (15) (curve *2*; $\langle V_{syst} \rangle = 45 \pm 23$ km/s).

Our computations of the evolution of triple systems revealed certain parameters whose impact on the characteristics and total number of X-ray novae proved to be the most important. These are the common-envelope efficiency α_{ce} and the modulus of the velocity gained by the neutron star during asymmetric core collapse w_{NS}. We computed a large number (~10^6) of evolutionary tracks for differing values of these parameters. Figure 4a presents the number of low-mass X-ray novae in the Galaxy as a function of α_{ce} and w_{NS}. Figures 4b and 5 present the computed spatial-velocity distribution for the X-ray novae as a function of the mean value of the anisotropic velocity w_{NS} and the common-envelope parameter α_{ce}.

It follows from Fig. 4c that the parameter intervals that are consistent with observations of X-ray novae are $\alpha_{ce} \sim 0.3$–0.8 and $w_{NS} \sim 0$–400 km/s. Even given the uncertainty due to the unknown fraction of triple systems, these parameters can be used to derive the expected observed number of low-mass X-ray novae with spatial velocities $V_{syst} \sim 40$ km/s.

4. CONCLUSIONS

Using a population synthesis technique, we have studied the possibility of forming X-ray novae from binary and multiple systems.

Binary systems consisting of a black hole with mass $M_{BH} \sim 5$–$10M_\odot$ and a low-mass main-sequence star filling its Roche lobe (an X-ray nova) can be formed dur-

range for the allowed mass for the third star [see (5)], all systems that evolved to an X-ray nova stage with a black hole contain low-mass components ($M \leq 1M_\odot$). The remaining systems become dynamically unstable and decay into a binary and a single star in the course of their evolution; the stage of interest is not reached.

ing the evolution of hierarchical triple systems. In this case, the assumption of an initially high mass ratio in the binary is not required.

Our computations depend on a number of parameters. The first group is related to the unknown fraction of triple (and larger multiplicity) systems in the Galaxy. Other parameters are associated with the overall scenario for the evolution of binary and multiple stars, and with the process of the formation of a stellar-mass black hole. In the framework of the adopted assumptions, we have determined the parameter values for which the numbers and average spatial velocity of low-mass X-ray novae with black holes originating from triple systems coincide with the observed values to an order of magnitude ($k_{BH} \sim 0.5$, $k_{env} \sim 0.5$, $\alpha_{ce} \sim 0.3$–0.8, $w_{NS} \sim 0$–400 km/s, the intervals for α_{ce} and w_{NS} being interdependent). The mean mass of the black holes formed is $\langle M_{bh} \rangle \sim 5.5 M_{\odot}$, which is close to the observed value.

The preferred explanation for the high spatial velocity in the system GRO J1655–40 = XN Sco = 1994 Cyg, which contains a more massive component, is that this system formed from a binary with a large initial mass ratio, provided the black hole acquired additional velocity during the asymmetrical collapse of the core of the supernova progenitor. Our computations indicate that the number of X-ray novae with moderate masses exceeds the observed value even for $\alpha_q = 2$. However, it is not ruled out that, in reality, the duration of the X-ray nova stage may be shorter for massive systems than for low-mass systems. Due to its large angular size, the massive component intercepts a larger fraction of the X-ray radiation than the low-mass component, which results in additional heating and an increase of its mass-loss rate. It is possible that the duration of the X-ray nova stage is affected by other factors as well. At present, this question remains unclear. In any case, it is difficult to explain the formation of low-mass X-ray novae only from massive binaries with high main-sequence mass-loss rates.

We therefore conclude that there are two different ways to form X-ray novae. A large fraction of the observed candidate black holes with low-mass optical components originate from triple hierarchical systems, whereas moderate-mass systems are formed from binary stars with large initial mass ratios. In the latter case, simultaneous explanation of the number of objects in the Galaxy and their high spatial velocities requires an anisotropic collapse during the formation of the black hole.

ACKNOWLEDGMENTS

REFERENCES

1. Ya. B. Zel'dovich, Dokl. Akad. Nauk SSSR **155**, 67 (1964) [Sov. Phys. Dokl. **9**, 195 (1964)].
2. Ya. B. Zel'dovich and O. Kh. Guseĭnov, Astron. Zh. **43**, 313 (1966) [Sov. Astron. **10**, 251 (1966)].
3. N. I. Shakura, Astron. Zh. **49**, 921 (1972) [Sov. Astron. **16**, 756 (1972)].
4. N. I. Shakura and R. A. Sunyaev, Astron. Astrophys. **24**, 337 (1973).
5. S. E. Woosley, T. A. Weaver, and A. Thomas, Astrophys. J., Suppl. Ser. **101**, 18 (1995).
6. D. Vanbeveren, E. de Donder, J. van Bever, *et al.*, New Astron. **3**, 443 (1998).
7. A. M. Cherepashchuk, Usp. Fiz. Nauk **39**, 759 (1996) [Phys. Usp. **166**, 809 (1996)].
8. A. V. Tutukov and A. M. Cherepashchuk, Astron. Zh. **74**, 407 (1997) [Astron. Rep. **41**, 355 (1997)].
9. A. M. Cherepashchuk **92**, 473 (2000).
10. Y. Tanaka and N. Shibazaki, Astron. Astrophys. **34**, 607 (1996).
11. O. Yu. Malkov, in *Binary Stars. Collection of Scientific Works* [in Russian] (Kosmosinform, Moscow, 1997), p. 162.
12. V. Kalogera, Astrophys. J. **521**, 723 (1999).
13. I. Iben, Jr. and A. V. Tutukov, Astrophys. J. **511**, 32 (1997).
14. D. S. Evans, Q. J. R. Astron. Soc. **9**, 388 (1968).
15. P. P. Eggleton and F. Verbunt, Mon. Not. R. Astron. Soc. **220**, 13P (1986).
16. L. D. Landau, Dokl. Akad. Nauk SSSR **17**, 301 (1937).
17. K. S. Thorne and A. N. Zytkow, Astrophys. J. Lett. **199**, L19 (1975).
18. K. S. Thorne and A. N. Zytkow, Astrophys. J. **212**, 832 (1977).
19. Ph. Podsiadlowski, R. C. Cannon, and M. J. Rees, Mon. Not. R. Astron. Soc. **274**, 485 (1995).
20. M. Barkov, G. S. Bisnovatyĭ-Kogan, and S. A. Lamzin, Astron. Zh. **78**, 269 (2001) [Astron. Rep. **45**, 230 (2001)].
21. C. L. Fryer, W. Benz, and M. Herant, Astrophys. J. **460**, 801 (1996).
22. L. G. Kiseleva, P. P. Eggleton, and S. Mikkola, Mon. Not. R. Astron. Soc. **300**, 292 (1998).
23. G. V. Lipunova and N. I. Shakura, Astron. Astrophys. **356**, 363 (2000).
24. B. Paczynski, in *Structure and Evolution of Close Binary Systems (IAU Symposium 73)*, Ed. by P. Eggleton, S. Mitton, and J. Whelan (Reidel, Dordrecht, 1976), p. 75.
25. V. M. Lipunov, K. A. Postnov, and M. E. Prokhorov, in *The Scenario Machine: Binary Star Population Synthesis*, Ed. by R. A. Sunyaev (Harwood Academic, Amsterdam, 1996); Astrophys. Space Sci. Rev. **9**, 1 (1996).
26. H. Abt and S. G. Levy, Astrophys. J., Suppl. Ser. **36**, 241 (1978).

27. H. Abt, Ann. Rev. Astron. Astrophys. **21**, 343 (1983).

28. D. C. Heggie, Mon. Not. R. Astron. Soc. **173**, 729 (1975).

29. P. P. Eggleton and L. G. Kiseleva, Astrophys. J. **455**, 640 (1995).

30. E. I. Popova, A. V. Tutukov, and L. R. Iungelson, Astrophys. Space Sci. **88**, 55 (1988).

31. M. Livio and N. Soker, Astrophys. J. **329**, 764 (1988).

32. R. F. Webbink, Astrophys. J. **277**, 355 (1984).

33. A. M. Cherepashchuk, Astrophys. J. (2001) (in press).

34. I. S. Shklovsky, Astrophys. J. **159**, L77 (1970).

35. T. M. Tauris and E. P. J. van den Heuvel, astro-ph/0001015 (2000).

36. A. G. Lyne and D. R. Lorimer, Nature **369**, 127 (1994).

37. C. de Jager, *The Brightness Stars* (D. Reidel, Dordrecht, 1980).

38. C. W. Allen, *Astrophysical Quantities* (Athlone Press, London, 1973).

39. W. Chen, C. R. Shrader, and M. Livio, Astrophys. J. **491**, 312 (1997).

40. Y. Tanaka, *Ginga Memo. Symposium, 1992*, Ed. by F. Makino and F. Nagase (Sagamihara ISAS), p. 19.

41. N. E. White and J. van Paradijs, Astrophys. J. Lett. **473**, L25 (1996).

42. N. Brandt, Ph. Podsiadlowski, and S. Sigurdsson, Mon. Not. R. Astron. Soc. **277**, L35 (1995).

43. S. P. Trudolyubov, Candidate's Dissertation (1999).

44. D. Bhattacharya and E. P. J. van den Heuvel, Phys. Rep. **203**, 1 (1991).

45. V. M. Lipunov, K. A. Postnov, and M. E. Prokhorov, Astron. Astrophys. **298**, 677 (1995).

46. G. Nelemans, T. M. Tauris, and E. P. J. van den Heuvel, Astron. Astrophys. **350**, 928 (1999).

Translated by K. Maslennikov

Evolution of the Dust Envelope
of Sakurai's Star (V4334 Sgr) in 1997–1999

V. I. Shenavrin and B. F. Yudin

Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899 Russia
Received March 31, 2000

Abstract—We have studied the structural evolution of the dust envelope of V4334 Sgr, starting with the onset of its condensation in 1997. A model with complete cloud cover, with the optical depth growing until the end of 1999, gives the best fit to the photometric data in the optical and IR. The inner radius of the dust layer remained virtually constant, whereas its thickness increased due to expansion. The deep minimum in the visual light curve of V4334 Sgr in October 1998 is attributed to the arrival at the dust-grain condensation zone of a density discontinuity in the circumstellar envelope. The discontinuity was probably formed early in 1997 due to an increase in the mass-loss rate by a factor of about four, possibly associated with an increase in the luminosity of V4334 Sgr during its transformation into a carbon star. After this luminosity increase, the mass-loss rate was $\dot{M} \approx 2 \times 10^{-6} M_\odot/\mathrm{yr}$. In the summer of 1999, the mass of the dust envelope was $M_{\mathrm{dust}} \approx 2 \times 10^{-7} M_\odot$ ($M_{\mathrm{gas}} \approx 4 \times 10^{-6} M_\odot$). In the complete-cloud-cover model, the envelope consists of graphite grains with $a_{\mathrm{gr}} = 0.05\ \mu\mathrm{m}$, to ~85% per cent in terms of the number of grains. The remaining ~15 per cent of the grains have sizes $a_{\mathrm{gr}} = 0.1$ and $0.25\ \mu\mathrm{m}$. To reproduce the small hump in the spectral energy distribution of V4334 Sgr near $11\ \mu\mathrm{m}$, some silicon carbide grains must be added to the graphite mixture. Their contribution to the V optical depth is $\leq 4\%$. The first deep minimum in the visual light curve could also be reproduced using a model in which the dust cloud has condensed along the line of sight, but a detailed analysis of the resulting characteristics of the cloud and envelope indicates that this model is improbable. © *2001 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

On February 20, 1996, the Japanese amateur astronomer Sakurai [1] discovered a nova-like variable in Sagittarius, later named V4334 Sgr (often called Sakurai's star in the literature). Observations of V4334 Sgr by Duerbeck and Benetti [2] immediately after its discovery showed that the star resembled an F supergiant whose atmosphere was deficient in hydrogen and enriched in carbon and oxygen. They also found an old, low-surface-brightness planetary nebula around the star. These facts led them to conclude that V4334 Sgr had experienced the so-called final helium shell flare when it was already the nucleus of the planetary nebula [3–5]. After this flare, the star again returned to its post-asymptotic-branch (post-AGB) track, initially moving along it backwards, i.e., gradually cooling due to expansion. Another fairly well studied "born-again AGB star" is FG Sge, the central star of the planetary nebula He 1-5.

In March 1997, lines of the C_2 molecule [6], and later, of the CN and CO molecules [7, 8] were found in the spectrum of V4334 Sgr. Thus, sometime during the previous gap in the observations, it became a carbon star. Also in March 1997, an IR excess due to emission from the now condensed dust envelope was revealed in the spectrum of V4334 Sgr [9].

Three minima were recorded in the optical light curve of V4334 Sgr in 1998: in March, June, and Octo-

ber [10–12]. The star's visual brightness decreased by $\sim 7^m$ in the last and deepest of these minima. A still deeper minimum was observed in March 1999 [13]. Thus, V4334 Sgr began to resemble an R CrB star in its atmospheric chemical abundances [14], as well as photometrically (the shape of its light curve).

Tatarnikov *et al.* [15] have presented 1996–1999 IR photometric data for V4334 Sgr. Combined with optical data, these observations made it possible to accurately estimate the bolometric flux from the star and study its photometric characteristics in detail. In the current study, we follow the evolution of the dust envelope since its formation, based on our analysis of the spectral energy distribution of V4334 Sgr from 0.44 to $10\ \mu\mathrm{m}$. Various models for its structure are considered in detail. We discuss the nature of the deep minima in the star's visual light curve.

2. OBSERVATIONAL DATA

Figure 1 presents variations in the V and K brightnesses, $B–V$ and $J–L$ color indices, and bolometric magnitude m_{bol} of V4334 Sgr (cf. [15] and references therein). The parameters of the star's radiation for the dates chosen for the analysis of its spectral energy distribution and estimates of the characteristics of its dust envelope are collected in Table 1. Before calculating m_{bol}, the magnitudes were corrected for interstellar
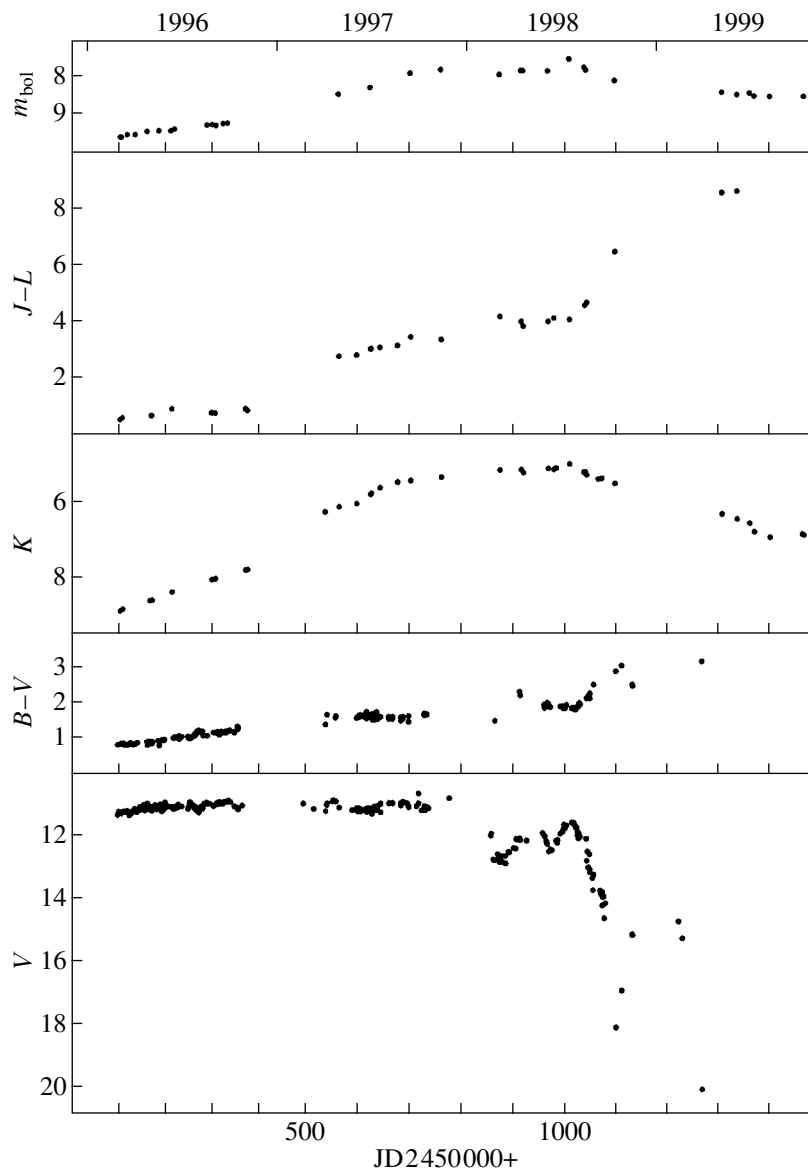
Fig. 1. Changes of the *V* and *K* brightnesses, *B–V* and *J–L* color indices, and bolometric magnitude of V4334 Sgr. The data were taken from [15].

absorption using the color excess $E(B-V) = 0.54$ [16]. This is the difference between the bolometric light curve and the other curves in Fig. 1, which are not corrected for interstellar reddening.

In Spring 1997, an IR excess due to emission from the dust envelope was detected in the spectral energy distribution of V4334 Sgr [9]. The last 1996 *JHKL* brightness estimate for V4334 Sgr was obtained on October 30 (Table 1), and the first IR observations of 1997 were made by the ISO space IR observatory on February 25, in seven filters at 4–16 μm [17]. By these latter measurements, a considerable IR excess was already present at 9–15 μm. Thus, the dust began to condense in January 1997. We will take this time to be Julian date JD 2450470 (January 21, 1997).

The bolometric flux from V4334 Sgr increased significantly in 1997 (Table 1), due solely to the IR flux increase. In other words, fluctuations in the visual brightness remained within the range recorded in 1996. There was no appreciable correlation between the star's *V* brightness and *B–V* in 1997.

In general, *B–V* fluctuated in 1997 within $0^{m}.3$, and its mean value was higher than in late October 1996 by $\sim 0^{m}.4$ (Fig. 1). The reddening of V4334 Sgr in the optical could be associated with the formation of an optically thick dust envelope around the star, or with cooling of the star itself; it is not possible to distinguish between these factors.

Table 1. Magnitude, colors, and bolometric fluxes of V4334 Sgr for selected dates

Date	V	K	$B–V$	$J–K$	$K–L$	$F_{bol}, 1$ $10^{-9} \, \text{erg cm}^{-2} \, \text{s}^{-1}$
Oct. 30, 1996	9.40	7.58	0.78	0.39	0.11	5.5
June 24, 1997	9.57	5.58	1.07	1.52	1.14	12.1
Sept. 9, 1997	9.37	5.23	1.16	1.67	1.41	16.9
Feb. 27, 1998	11.07	4.95	0.98	2.31	1.51	16.4
July 12, 1998	9.96	4.79	1.42	1.97	1.73	23.6
Oct. 7, 1998	16.42	5.31	2.36	4.03	2.08	14.1
Apr. 30, 1999		6.12		5.59	2.62	10.7
July 31, 1999		6.73		6.53	3.15	9.6

Magnitudes and fluxes have been corrected for interstellar reddening with $E(B–V) = 0.54$.

The increased IR flux from V4334 Sgr in the absence of significant changes in its optical brightness implies a gradual increase of the contribution of the dust-envelope emission to the bolometric flux. In other words, dust grains continued to condense and/or grow, and, accordingly, the optical depth of the dust envelope gradually increased during 1997.

The situation changed dramatically in 1998. The overall visual brightness of Sakurai's star decreased, and the corresponding light curve showed first two shallow, but quite significant, minima and then a very deep minimum (Fig. 1). In June 1998 (Table 1), the maximum bolometric flux was recorded during a local maximum of the visual brightness. However, the star was fainter in the visual than in 1997 by $\geq 0\overset{m}{.}5$ (Fig. 1). This means that the contribution of the emission from the dust envelope to the object's total radiation (star + dust envelope) increased in 1998, so that the envelope's optical depth became higher.

In the subsequent deep minimum of October 1998, the bolometric flux decreased by a factor of ~1.7 (Table 1). A still deeper minimum was registered in March 1999 (Fig. 1), after which no optical brightness estimates were obtained. Compared to October 1998, the star became considerably redder in the IR. The bolometric flux also decreased (Table 1). It is fair to say that, since October 1998, only emission from the dust envelope has been observed in the IR, and the envelope's optical depth continued to increase through 1999.

3. CHOICE OF MODEL FOR THE DUST ENVELOPE

Geometrically, the process of condensation of the dust envelope around the star can take one of two forms. In the first, the dust grains condense simultaneously in a spherical layer surrounding the star. In other words, at any time during the dust envelope's evolution, it can be considered to offer complete cloud cover, with the optical depth along the line of sight, τ_{sl}, not significantly different from the optical depth averaged over the whole sphere, τ_{av}. In this case, a spherically symmetrical model for the dust envelope can be used to compute the spectral energy distribution [18]. This form of condensation is observed for the dust envelopes of classical novae, such as DQ Her. For such stars, as for V4334 Sgr, the dust envelope does not significantly redden the emerging stellar radiation. This also applies to FG Sge, the evolutionary analog of V4334 Sgr, which has already displayed six deep minima since 1992.

In the second form of condensation, the dust envelope around the star forms gradually out of individual thick dust clouds and appears as a "patchy dust blanket," with τ_{sl} and τ_{av} possibly differing significantly. If the star is viewed through a "hole" in the blanket ($\tau_{sl} \ll \tau_{av}$), the dust envelope will have no effect in the optical. This is precisely the appearance of the first stage of the dust envelope's condensation around Sakurai's star, which lasted until the first minimum in its optical light curve in February 1998 (Fig. 1), and this how it is interpreted by Arkhipova *et al.* [12] and Duerbeck *et al.* [19].

In turn, both shallow and deep minima in the visual light curve can be attributed to the crossing of an optically thick dust cloud through the line of sight. In each of the two forms of condensation, the observed bolometric flux will differ from that of the central star (in other words, the observed bolometric flux will not correspond to the star's luminosity). In the first case, the observed bolometric flux will be higher and, in the second case, lower than the bolometric flux from the central star.

Let us suppose that, in 1997, Sakurai's star was seen through the hole in a dust blanket. Then, the dip in its visual brightness in February 1998 (the first minimum) should have been accompanied by a significant decrease in the bolometric flux, which was not observed (Fig. 1). The bolometric flux in February 1998 was the same as in September 1997 (Table 1), only ~10% lower than in November 1997. This relative flux deficiency was already absent in the second shallow minimum of June 1998, and a maximum of the bolometric flux was
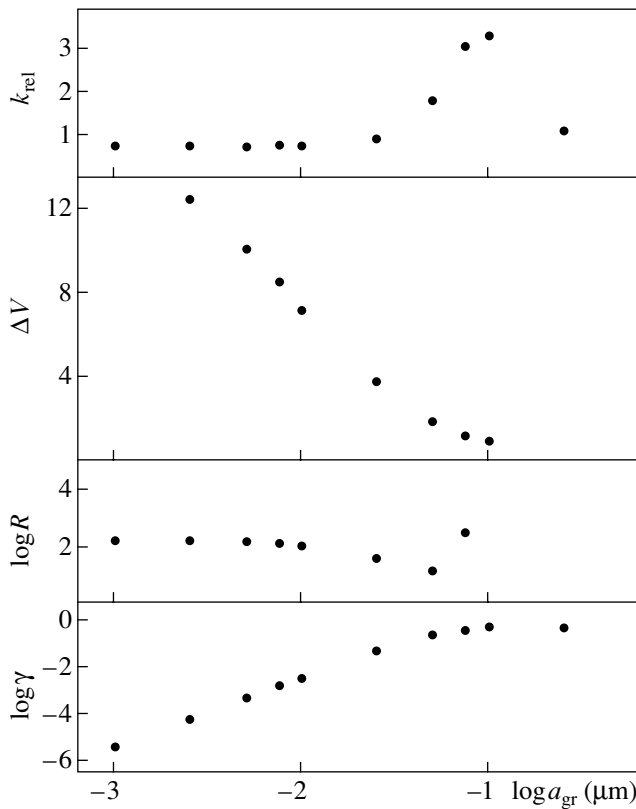
Fig. 2. The optical properties of spherical graphite grains as a function of their radii. $\gamma(a_{gr})$ is the albedo; $R(a_{gr}) = A(V)/E(B-V)$ is the absorption-to-reddening ratio; $\Delta V = V_{env} - V_{obj}$ is the difference in the magnitudes of the dust envelope and the object as a whole; $k(a_{gr}) = 3Q_{abs}/4a_{gr}\rho_{gr}$ is the specific absorption coefficient. The parameters γ and k refer to 0.55 μm. See text for details.

recorded in July 1998 (a local maximum; Table 1), though the visual brightness of V4334 Sgr was then considerably lower than in 1997 (Fig. 1).

The bolometric flux of V4334 Sgr approximately doubled from October 1996 to June 1997 (Table 1), due solely to the developing IR excess, i.e., to the emission of the condensed dust envelope. The envelope's contribution to the observed bolometric flux was ~50%. In this case, if the star was seen through a hole in a dust blanket in June 1997, then the blanket had to absorb the radiation from the central star nearly completely to reproduce this value. In turn, this implies that there was a single hole in the line of sight, with the star completely covered by optically thick clouds in all other directions. This model for the dust envelope is clearly not realistic, and it is completely ruled out for November 1997, when the IR excess additionally increased by a factor of ~1.8.

Thus, we have every reason to believe that, during all of 1997 and in July 1998 (a local maximum of the visual brightness), the dust envelope of V4334 Sgr can be considered to offer complete cloud cover, so that we can use a spherically symmetrical model for the dust

envelope to calculate the radiation from the object (central star plus dust envelope). All the same, we will consider the first deep minimum of October 1998 (Fig. 1) in terms of both models—complete cloud cover (the CCC model) and a cloud in the line of sight (the CLS model). The reason is that the CLS model cannot be rejected *a priori* for the first deep minimum.

After the second deep minimum in March 1999, the optical emission of V4334 Sgr disappeared altogether. Even if the first deep minimum was due to the appearance of a dust cloud in the line of sight, it is quite logical to suppose that, between minima, such clouds should also have appeared in other directions. This is usual for R CrB stars [20], since there are no reasons for a preferential direction for the ejection of material. Note that the planetary nebula around Sakurai's star is spherical, and that it was formed essentially at the evolutionary stage to which V4334 Sgr returned after its "second birth."

Figure 1 shows that the visual brightness of V4334 Sgr between the minima was much lower than prior to them. This implies that the first dust cloud (in the CLS model) did not completely dissipate before the appearance of a second, denser, cloud in the line of sight. For reasons of symmetry, the same should have occurred in 1999 in all other directions. In turn, this implies that the cloud cover should be considered complete, so that the CCC model can be used to estimate the dust envelope's parameters.

An important property of the dust envelope of V4334 Sgr is that it leads to a relatively small reddening of the object's *B–V*. It is clear from Fig. 1 that V4334 Sgr was becoming gradually redder before the first deep minimum. However, the absence of a clear correlation between the visual brightness and *B–V* makes it impossible to ascribe this reddening solely to the influence of the dust envelope, since, in principle, some (or all) of the reddening could be due to reddening of the star itself in the course of its evolution backward along the post-AGB track. We cannot estimate this effect after 1997, when V4334 Sgr became a carbon star and a dust envelope formed around it. The star's reddening was clearly recorded in the first deep minimum. It started as soon as the brightness began to decrease, and the dependence of *B–V* on *V* was initially steeper. After smoothing with a second order polynomial, we find that $R = A(V)/E(B-V) \approx 1/(0.47-0.02V)$. Thus, at the onset of the brightness decrease, $R \approx 4.3$, and at the visual minimum, $R \approx 7$.

Figure 2 shows the (R, a_{gr}) and $(\gamma(\lambda), a_{gr})$ diagram for spherical graphite grains, where $\gamma(\lambda, a_{gr}) = Q_{sca}(\lambda, a_{gr})/(Q_{abs}(\lambda, a_{gr}) + Q_{sca}(\lambda, a_{gr}))$ is the albedo calculated for wavelength $\lambda = 0.55$ μm and a_{gr} is the grain radius. The efficiencies of absorption, $Q_{abs}(\lambda, a_{gr})$, and scattering, $Q_{sca}(\lambda, a_{gr})$, were computed using formulas for Mie theory, and the refractive index for graphite was taken from [21]. It is apparent from Fig. 2 that, for small graphite grains ($a_{gr} = 0.001 - 0.01$ μm), $R \approx 2.15$.

With increasing grain size, R first decreases and then abruptly increases, and, for $a_{gr} \geq 0.08$ μm, the absorption in the B and V bands becomes neutral.

Figure 2 shows that the albedo of small graphite grains grows with their size, so that, for $a_{gr} \geq 0.05$ μm, the dust envelope will contribute appreciably to the object's radiation in the visible, due to scattering. Figure 2 shows the calculated dependence of the difference between the visual magnitudes of the dust envelope and of the object ($\Delta V = V_{env} - V_{obj}$) on the grain size. The optical depth of the dust envelope was taken to be $\tau(V) = 2$. The value of ΔV shows the decrease in the object's visual brightness that occurs if the direct light from the star is completely screened by a dust cloud in the line of sight. In the first minimum, the visual brightness of V4334 Sgr decreased by $\sim6^{m}.5$. For this reason, in the CLS model, the dust envelope with complete cloud cover that describes the state of Sakurai's star in the preceding local maximum must contain grains with sizes ≤0.01 μm; otherwise, the CLS model is not able to provide a minimum that deep. If the star's radiation is screened by a dust cloud, we will see it in reflected light. Note that V4334 Sgr was $\sim2^{m}$ fainter in the second deep minimum than in the first one. Thus, in the CLS model, the dust envelope surrounding the star and creating the IR excess must consist of small grains. The cloud itself must contain large grains in order not to lead to substantial reddening.

4. PROBLEM FORMULATION AND COMPUTATIONS

We computed the spectral energy distribution of the star surrounded with the dust envelope using the CSDUST3 code, described in detail by Egan *et al.* [22]. The envelope was described as a dust layer with inner radius R_{in}, outer radius R_{out}, and thickness ΔR_{env}.

The optical depth at a given wavelength, in our case at $\lambda = 0.55$ μm (V band), is one of the input parameters of the problem, and is determined as

$$\tau(\lambda) = (Q_{abs}(\lambda) + Q_{sca}(\lambda))\pi a_{gr}^2 n_{gr}\Delta R_{env}, \qquad (1)$$

where n_{gr} is the number of grains per unit volume. We can use (1) to estimate the mass of the dust envelope:

$$M_{dust} = 4\pi R_{in}^2 f(R_{in}/R_{out}, \alpha)\tau(V)/k(V). \qquad (2)$$

Here, the function $f(R_{in}/R_{out}, \alpha)$ depends on the ratio R_{in}/R_{out} and the index α in the power law for the distribution of dust in the layer, $n_{gr} \propto (R_*/r)^{\alpha}$, with R_* being the star's radius. The normalized $k(V)$, a_{gr} relation is shown in Fig. 2.

In our case, $\Delta R_{env}/R_{in} < 1$, so that changes in α have virtually no influence on the computed spectral energy distribution, and only a small effect on the estimate of the mass of the dust envelope. For example, when $\Delta R_{env}/R_{in} = 0.4$, the mass estimate decreases by a factor

of ~1.3 as α changes from two to four. Note that this value of $\Delta R_{env}/R_{in}$ was characteristic of the dust envelope of V4334 Sgr in mid-1999, so that $\Delta R_{env}/R_{in} < 0.4$ in 1997–1998, and the influence of α on the mass estimate for the dust envelope is negligible. The luminosity of V4334 Sgr increased after its flare (Fig. 1), while its temperature decreased (at least until mid-1997), so that we can assume that the mass-loss rate was increasing. For this reason, we computed the parameters of the dust envelope presented below for $\alpha = 3$.

The grain temperature at the inner boundary of the dust envelope is

$$T_{gr, in} \propto T_*(R_{in}/R_*)^{-\beta}, \qquad (3)$$

where T_* and R_* are the star's temperature and radius. When $T_* \approx 5000$ K, the power-law index β is from 0.34 to 0.40 for graphite grains with $a_{gr} \leq 0.25$ μm. The calculations show that, given this $T_{gr, in}(R_{in})$ relation, we can estimate the inner radius of the dust envelope with accuracy to $\sim30\%$. Thus, for a given time, taking into account possible variations of the factors in (2), the mass of the gas envelope can be estimated within a factor of two. Relative changes in the course of the evolution of the gas envelope can be determined even more accurately.

In our model for the dust envelope, we assumed that the grains were formed at a fixed distance from the star ($R_{in} = $ const) of gas layers flowing into the condensation zone. In this case, the thickness of the dust layer, ΔR_{env}, gradually increases. We assumed 50 km/s for the expansion rate of the envelope of V4334 Sgr. Note that $V_{esc} \geq 35$ km/s. The star's luminosity increased during 1997, so that $T_{gr, in}$ was growing. When the star's luminosity increases by a factor of ~2.3, corresponding to the range of variations recorded for V4334 Sgr from Spring 1997 until July 1998, the temperature of the graphite grains at a given radius, for instance at R_{in}, should increase by $\sim15\%$.

In computations of the radiation from a star surrounded with a dust envelope, the condensation temperature is often fixed. In this case, the inner radius of the dust envelope of V4334 Sgr should have gradually increased until mid-1998, at a rate of ~50 km/s. In this type of model, the material for new grains can originate only from gas swept up by the dust layer. Because the mass-loss rate of Sakurai's star increased during its backward motion along the post-AGB track, this scheme with increasing mass and optical depth for the dust envelope seems less realistic than the scheme described above. However, it does not obviously contradict any specific facts. Therefore, we will also analyze this model for the dust envelope's evolution.

Since the spectral energy distribution for the star 4334 Sgr itself is not known, we represented the star as a blackbody with a temperature of 4600–5000 K [8]. The particular value of the temperature was selected so
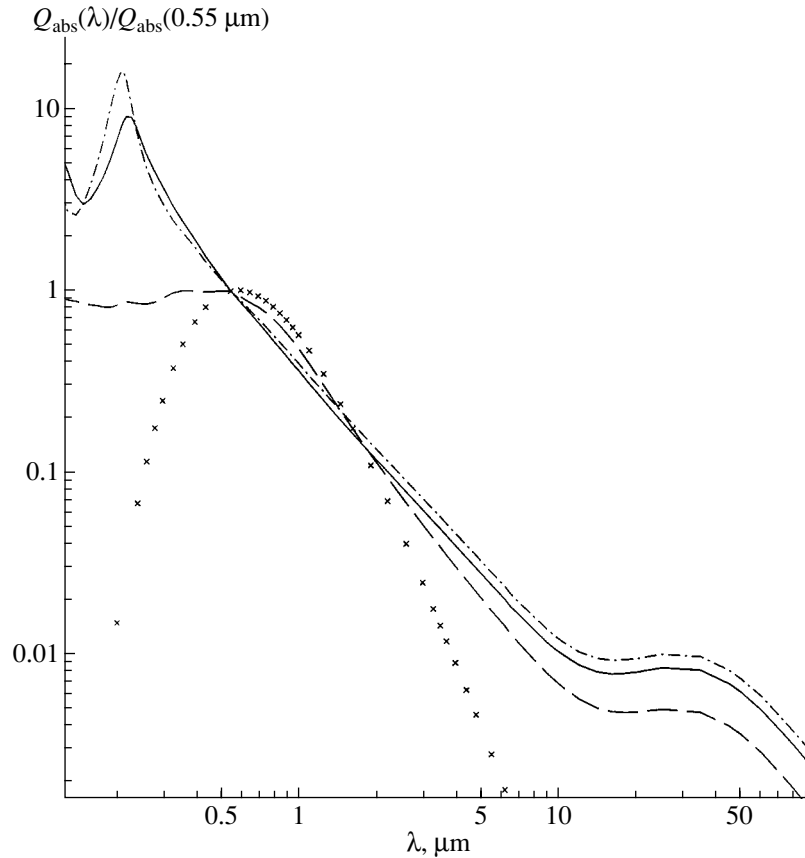
$Q_{abs}(\lambda)/Q_{abs}(0.55\ \mu m)$

Fig. 3. $Q_{abs}(\lambda)/Q_{abs}(0.55\ \mu m)$, λ diagram for spherical graphite grains with radii $a_{gr} = 0.005\ \mu m$ (dot–dashed), 0.025 μm (solid), and 0.1 μm (dashed). The crosses correspond to the normalized blackbody curve for $T = 5000$ K.

that the *B–V* for the computed spectral energy distribution for June 1997 coincided with the observed value. We did not change the star's temperature in our computations of the spectral energy distribution for subsequent dates, and the computed changes of *B–V*—due solely to changes of the optical depth of the dust envelope—were compared to the observed changes of this index. As noted above, such changes could be due, in part, to color changes of the star itself. For the observations of June 1997, the computed energy distribution was normalized to the observed radiation flux at 1.25 μm (the *J* band). When computing models for subsequent observations, we retained the normalization factor determined for June 1997.

In computations of the radiation for a star surrounded by a dust envelope, the luminosity-related input parameter is the flux at the inner boundary of the dust envelope: $L_{bol}/4\pi R_{in}^2 = T_*^4 (R_*/R_{in})^2 = F_{bol}(D/R_{in})^2$, where *D* is the distance to the star. Thus, if we change the estimate of the distance to V4334 Sgr, the linear size of the dust envelope and related parameters should be scaled according to the relation $D/R_{in} = $ const. Note also that the input parameter R_{in}/R_* considerably influences the shape of the spectral energy distribution; like the

optical depth of the envelope, this does not change with the distance to the star. We adopted ~5.5 kpc for the distance to V4334 Sgr [2].

We assumed in the CCC model that the envelope contained graphite grains with $a_{gr} = 0.05$, 0.1, and 0.25 μm. The properties of this mixture are discussed in detail in [18]. Grains with $a_{gr} = 0.05\ \mu m$ are introduced into the mixture to ensure reddening of the object in the optical, if required. The absorption by larger grains is neutral. However, the use of grains with $a_{gr} = 0.1\ \mu m$ instead of $a_{gr} = 0.05\ \mu m$ leaves the shape of the object's spectral energy distribution in the IR virtually unchanged. Adding grains with $a_{gr} = 0.25\ \mu m$ to the mixture makes it possible to alter this characteristic of the object's radiation, for a given optical depth and inner radius of the dust envelope.

In the CLS model, the dust envelope must consist of small grains in order to avoid appreciable scattering of the star's radiation. In our calculations, we chose graphite grains with $a_{gr} = 0.01\ \mu m$. Graphite grains with $a_{gr} \leq 0.02\ \mu m$ have nearly identical optical properties (Figs. 2 and 3), with the exception of the albedo, so that there is no need to compose a mixture of the different grains. Their albedos are so low that they do not influ-

F_λ, 10^{-13} erg cm^{-2} s^{-1} Å^{-1}
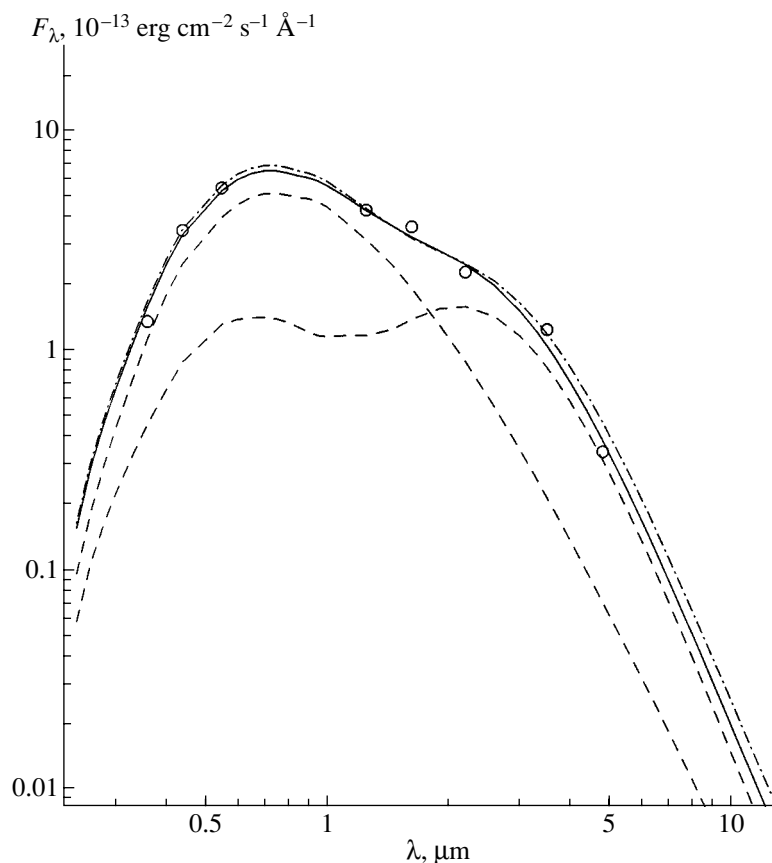
Fig. 4. Spectral energy distribution of V4334 Sgr in June 1996 (circles). The solid curve shows the computed spectral energy distribution of the object (star + dust envelope) for model G97Jn1; the dashed curves are spectral energy distributions of the star and the dust envelope. The computed energy distribution for model G97Jn2 is shown as a dot–dashed curve.

ence the shape of the spectral energy distribution of the object.

Near 11 μm, the energy distribution computed for a dust envelope containing only graphite grains is much lower than the point corresponding to the observed radiation flux. To compensate for this, we supplemented the graphite grains with silicon carbide grains, leading to an intense emission band at $\lambda \approx 11.3$ μm, as is observed in the spectra of carbon stars.

5. RESULTS OF THE COMPUTATIONS

Figure 4 shows the spectral energy distribution of V4334 Sgr on June 24, 1997 (Table 1) together with two computed distributions for the CCC model: models G97Jn1 and G97Jn2, for which parameters are presented in Table 2. The model code used in Table 2 indicates the model's chemical composition (G: graphite only; GS: graphite with silicon carbide), the date (year and month), and the model's number among models computed for the given date. With the exception of the ratio R_{in}/R_*, all remaining input parameters were the same for models G97Jn1 and G97Jn2: $R_{in}/R_{out} = 0.8$, $\tau(V) = 1.05$, $\alpha = 3$, $T_* = 4600$ K, $L_{bol} = 10.3 \times 10^3 L_\odot$,

with the fraction by number (mass) of graphite grains with $a_{gr} = 0.05$, 0.1, and 0.25 μm being ~85% (25%), ~14% (32%), and ~1% (43%), respectively. In the models for later dates, the number of grains with $a_{gr} = 0.05$ μm remained unchanged, while the numbers of grains with $a_{gr} = 0.1$ and 0.25 μm were varied such that their combined fraction of the total by number remained 15%; the number of grains with $a_{gr} = 0.25$ μm never exceeded 2.5%. We also did not vary the star's temperature.

The reddening introduced by such an envelope is $\Delta(B-V) \approx 0.16$. However, this parameter increased by $\sim 0^{m}\!.3$ from November 1996 to June 1997 (Table 1). Increasing the fraction of small grains in the mixture, we can correspondingly increase the reddening. Figure 4 indicates that the dust envelope scatters the radiation of the star rather effectively. For this reason, even if we completely block the star's radiation with a dust cloud, we will still not obtain a deep optical eclipse. The radiation flux from the dust envelope (thermal + scattered) is ~45% of the bolometric flux from the object.

Models G97Jn1 and G97Jn2 have $R_{in}/R_* \approx 27$ and 34 ($R_* \approx 162 R_\odot$), respectively. This ~25% increase in this parameter increases the estimate of the dust-enve-

Table 2. Parameters of the dust envelope of V4334 Sgr

Date	$\tau(V)$	R_{in}/R_*	R_{in}/R_{out}	$R_{in}, \times 10^{14}$ cm	$T_{gr, in}$, K	$M_{dust}, \times 10^{-8} M_\odot$	Model
June 24, 1997	1.05	27	0.80	3.0	1190	1.2	G97Jn1
June 24, 1997	1.05	34	0.80	3.6	1110	1.7	G97Jn2
Sept. 9, 1997	1.30	28	0.75	3.6	1190	2.7	G97S1
Sept. 9, 1997	1.25/0.05	28	0.75	3.6	1190	2.7	GS97S2
July 12, 1998	2.35	24	0.62	3.6	1270	5.7	G98Jl1
July 12, 1998	2.35	34	0.80	5.1	1110	7.8	G98Jl2
July 12, 1998	2.35	47	0.80	7.2	1000	15	G98Jl3
July 12, 1998	2.25/0.10	24	0.62	3.6	1270	5.7	GS98Jl4
Oct. 7, 1998	8.45	24	0.45	2.9	1300	12	G98O1
Apr. 30, 1999	13.5	28	0.40	2.9	1260	18	G99A1
July 31, 1999	17.5	29	0.40	2.9	1250	22	G99Jl1

For models GS97S2 (September 9, 1997) and GS98Jl4 (July 12, 1998), the first and second values in the $\tau(V)$ columns refer to the graphite and silicon carbide components of the dust envelope, respectively.

lope mass by a factor of ~1.5. Below, we adopt G97Jn2 as the dust-envelope model for June 1997. It has $R_{in} \approx 3.6 \times 10^{14}$ cm and $M_{dust} \approx 1.7 \times 10^{-8} M_\odot$. As noted above, the condensation of the dust envelope began in mid-January 1997. For the chosen envelope expansion rate, $R_{in}/R_{out} \approx 0.8$ in June 1997. This is the value of the corresponding input parameter used in our computations.

The value R_{in}/V_{env} is about 2.4 yr. Consequently, matter that reached the condensation radius by June 1997 should have left the star in February 1995. If we assume that the carbon abundance in the envelope is ~5% by mass [23] and that all the carbon is bound in grains in the condensation region, the mass-loss rate for Sakurai's star at that time was ~$6 \times 10^{-7} M_\odot$/yr.

If the distance to V4334 Sgr is decreased by a factor of 1.5, the estimates of R_*, R_{in}, R_{in}/V_{env}, and $\dot{M}$ will decrease by the same factor. For $D \approx 3.7$ kpk, matter that reached the condensation radius by August 1998 left the star in January 1997. Recall that V4334 Sgr became a carbon star in the winter of 1997, and its luminosity increased significantly. This could lead to a significant increase in the mass-loss rate and, accordingly, to the appearance of a density discontinuity in the circumstellar envelope. Thus, by decreasing the distance estimate for Sakurai's star, we can consider these events to be cause and effect.

Figure 5 presents the spectral energy distributions of V4334 Sgr on September 9, 1997 and July 12, 1998 (Table 1), along with several computed curves (models G97S1, GS97S2, G98Jl1, G98Jl2, G98Jl3, and GS98Jl4, Table 2). The data points on the spectral energy distribution curves near 11 μm are derived from [17, 24, 25].

In model GS97S2, we added silicon carbide grains to the graphite grains to compensate for the difference between the observed and computed (in model G97S1) radiation fluxes near $\lambda \approx 11$ μm. The SiC grains contrib-

ute ~4% of the optical depth of the dust envelope in the V filter, but they create a substantial emission band near 11 μm. The largest deviation of the computed energy distribution for model GS97S2 from the spectral energy distribution of V4334 Sgr occurs in the K and M filters, and amounts to ~15%. The flux from the dust envelope is ~58% of the object's bolometric flux.

Figure 5 shows energy distributions for July 1998, computed using three models with pure graphite dust envelopes, differing only in their values of R_{in}/R_* and R_{in}/R_{out}. Model G98Jl1 corresponds to the assumption of constant inner radius and increasing thickness of the dust envelope. The luminosity of V4334 Sgr increased from September 1997 to July 1998 (Table 1), so that the temperature of the grains at the inner boundary of the model dust envelope became higher (Table 2). The increased luminosity also leads to a considerably higher K-band brightness for model G98Jl1, so that the computed energy distribution goes much higher than the data point for the radiation flux of Sakurai's star at 2.2 μm (Fig. 5). The corresponding flux difference is ~30%.

Between June 1997 and July 1998, the mass of the dust envelope increased by a factor of ~3.3, and the mass-loss rate increased by ~25%. Recall that the estimated mass-loss rate refers to some time in the past, rather than to the time of the observations. For the given D and V_{env} values, this was March 1996, or, if we reduce D by a factor of 1.5, December 1996. The $B–V$ of model G98Jl1 is ~$0\overset{m}{.}2$ higher than that of model G97Jn2. However, the observed increase between June 1997 and July 1998 was ~$0\overset{m}{.}35$ (Table 1). As noted above, this difference could be due to cooling of V4334 Sgr, or could indicate that the fraction of grains with $a_{gr} = 0.05$ μm should be increased.

The model G98Jl2 is based on the assumption of constant grain temperature at the inner radius of the

F_λ, 10^{-13} erg cm^{-2} s^{-1} Å^{-1}
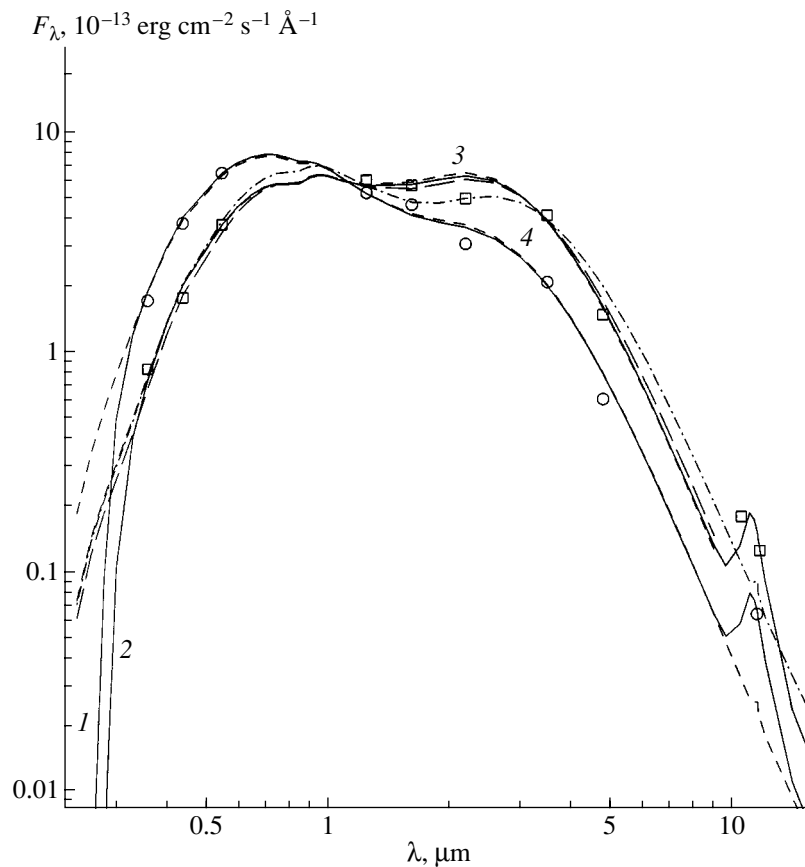
Fig. 5. Spectral energy distribution of V4334 Sgr in September 1997 (circles) and July 1998 (squares). The solid curves are computed energy distributions for models GS97S2 (curve *1*) and GS98Jl4 (curve *2*). The short-dashed curves display the computed energy distributions for models G97S1 (curve *3*) and G89Jl1 (curve *4*); the long-dashed curve, for model G98Jl2; and the dot–dashed curve, for model G98Jl3. For clarity, the energy distribution curves of models G98Jl1 and G98Jl2 were truncated at 9 μm.

dust envelope (grain-condensation radius) and constant envelope thickness. In this case, the ratio R_{in}/R_* increases (Table 2), reducing the model's K brightness and increasing its M brightness by ~10% (Fig. 5). The computed radiation fluxes at 2.2 and 4.8 μm for model G98Jl2 are higher than the observed values by ~20%.

To bring the computed K radiation flux into agreement with the observed value, it is necessary to additionally increase R_{in}/R_* by a factor of ~1.4 (model G98Jl3, Table 2). However, fitting the computed energy distribution in the K band brings about a significant discrepancy in the M band (Fig. 5).

The inner radius of the dust envelope in model G98Jl2 is larger than in model G97Jn2 by ~1.5×10^{14} cm. The corresponding displacement rate is ~50 km/s. This rate is needed to keep the grain temperature at the dust envelope's inner boundary constant as the luminosity of the star increases. As noted above, in this model, the only material for new grains is gas swept up by the dust layer. In this case, between June 1997 and July 1998, the mass of the dust layer could increase by only a factor of ≤1.5, even if we assume $\alpha \approx 0$ (in the $n_{gr} \propto (R_*/r)^\alpha$ law); i.e., if we suppose that the mass-loss rate of Saku-

rai's star decreased in time. Note that this mass increases by a factor of ~4.6 for models G97Jn2 and G98Jl2 (Table 2).

Of course, one could argue that not all the carbon in the condensation zone had become solid in June 1997, and that this process continued throughout the following year. However, the visual dips of R CrB stars, FG Sge, and V4334 Sgr are abrupt, testifying that the grain-condensation process occurs rapidly and in a limited volume.

After the local maximum of the visual brightness of V4334 Sgr in July 1998, the first deep minimum occurred (Fig. 1). The spectral energy distribution on October 7, 1998 is displayed in Fig. 6, together with the curve computed in model G98O1 (Table 2). The CCC model associates the deep and rather abrupt dip in visual brightness to a density discontinuity in the circumstellar envelope of Sakurai's star, which reached the grain-condensation zone by August 1998. For this reason, the model envelope for October 1998 consists of two layers. A geometrically thin ($R_{in}/R_{out} = 0.9$), dense layer is added to the fairly extended dust envelope ($R_{in}/R_{out} \approx 0.6$) formed by August 1998 (models
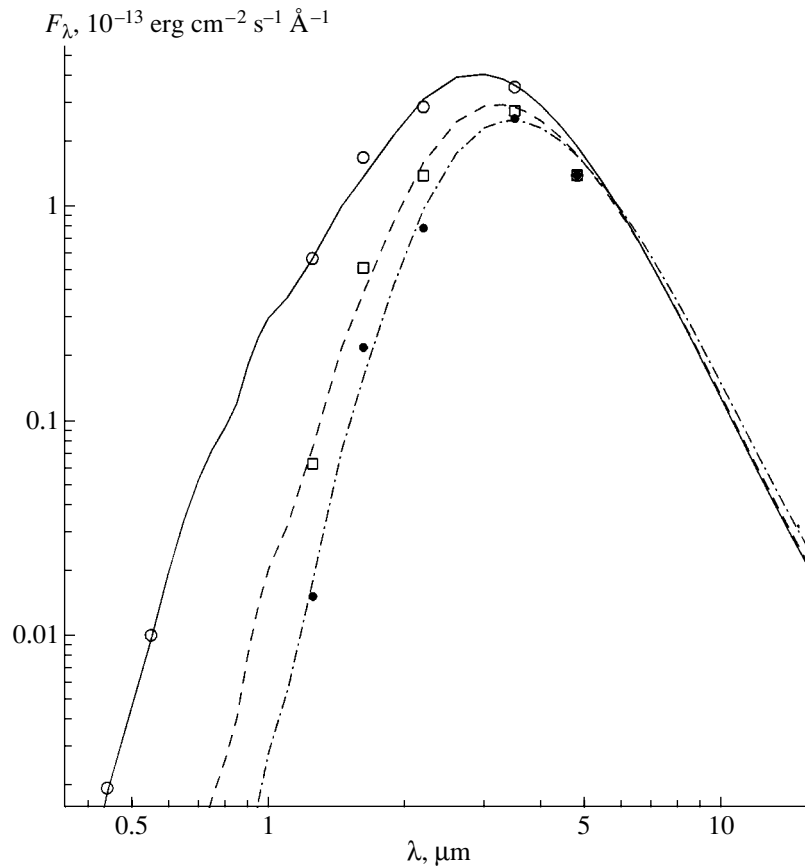
Fig. 6. Spectral energy distributions of V4334 Sgr for October 1998 (circles), April 1999 (squares), and July 1999 (dots). The solid, dashed, and dot–dashed curves are the computed energy distributions for models G98O1, G99A1, and G99Jl1, respectively.

G98Jl1 or GS98Jl4, Table 2). Thus, R_{in} in model G98O1 refers to the inner radius of the layer and R_{out} to the outer radius of the whole dust envelope. The layer's optical depth, $\tau(V)$, and mass are ~70% and ~50% of the corresponding parameters for the entire envelope presented in Table 2.

After the deep minimum in October 1998, the visual brightness of V4334 Sgr, without returning to its original level, dropped to $\sim 20^m$ in March 1999 (Fig. 1). No further estimates of the visual brightness of V4334 Sgr in 1999 are available. Sakurai's star effectively became an IR source. Figure 6 shows the spectral energy distributions for April and July 1999, along with the computed energy distributions for the models whose parameters are presented in Table 2.

The dust envelopes of models G99A1 and G99Jl1, like that of model G98O1, consist of two layers, with the density discontinuity that had formed in August 1998 (at the beginning of the first deep brightness decrease). The discontinuity's outer radius in April 1999 was computed using the adopted expansion rate of the dust envelope. The same technique was used to estimate the position of the outer boundary of the dust envelope as a whole. We used the values $\Delta R_{env}/R_{in} \approx 0.25$ and 0.3 for the inner, dense layer in models G99A1 and G99Jl1,

respectively. This parameter was taken to be 0.1 for model G98O1, though it could actually be lower. However, as already noted, further reduction of this parameter does not influence the estimated mass of the dense layer of the dust envelope, while the estimated mass-loss rate changes considerably.

We can estimate the thickness of the dense layer for models G99A1 and G99Jl1 rather confidently based on our assumptions. Therefore, we can also estimate the mass-loss rates for these models. In model G99A1, $\dot{M} \approx 2 \times 10^{-6} M_\odot/$yr. The value of $\dot{M}$ for model G98Jl1 is a factor of ~5.5 lower than for model G99A1. Thus, the density discontinuity in the envelope of V4334 Sgr was due to an abrupt increase of the mass-loss rate of Sakurai's star, by a factor of about four. As noted above, if $D \approx 3.7$ kpc, this time coincides with the transformation of V4334 Sgr into a carbon star in early 1997.

In the CLS model describing the first deep minimum, the dust envelope prior to the brightness decrease, in July 1998, consists of fine ($a_{gr} = 0.01$ μm) graphite grains, and has optical depth $\tau(V) = 1.9$. Its geometrical parameters are similar to those of the envelope of model G98Jl3 (Table 2): $R_{in}/R_* \approx 47$, $R_{in}/R_{out} = 0.8$. Among the CCC models for July 1998, model G98Jl3 has

F_λ, 10^{-13} erg cm^{-2} s^{-1} Å^{-1}

λ, μm
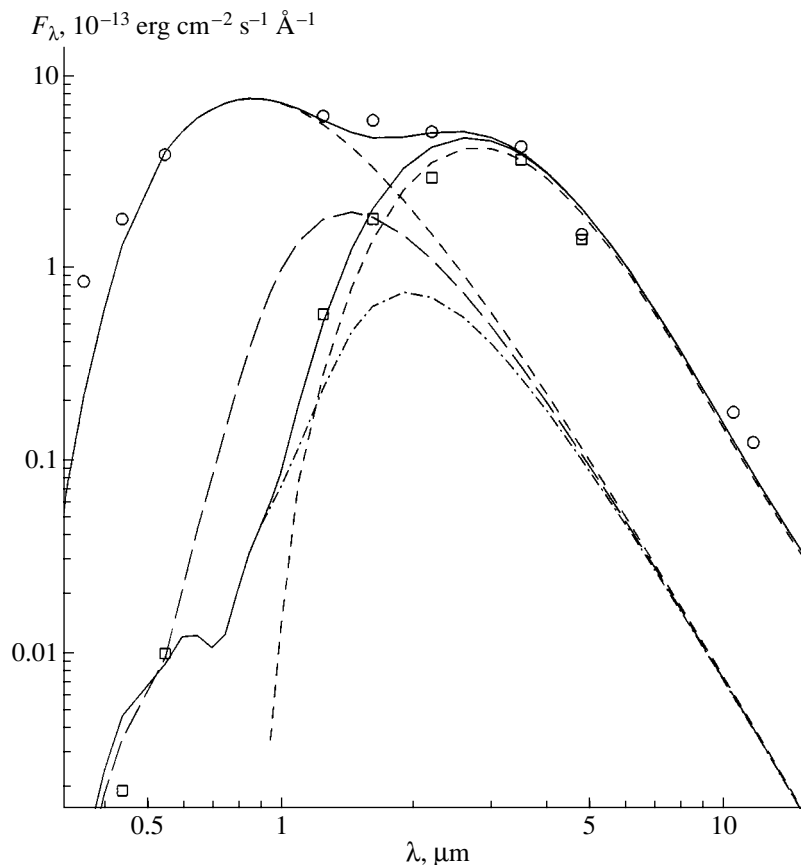
Fig. 7. Spectral energy distributions of V4334 Sgr for July 1998 (circles) and October 1998 (squares). The solid curve is the computed spectral energy distribution of the object (star + dust envelope) for July 1998 for the CCC model with fine ($a_{gr} = 0.01$ μm) graphite grains; the short-dashed curves are the energy distributions of the star and dust envelope. The spectral energy distributions for the case when the star is obscured by a dust cloud with $\tau(V) = 6$ and graphite grains with $a_{gr} = 0.1$ and 0.15 μm are shown by the long-dashed and dot–dashed curves, respectively. The computed energy distribution in the CLS model for October 1998 is shown by the solid curve.

the highest R_{in}/R_* and, for the reasons described above, seems more realistic than model G98Gl1.

In the CLS model, we cannot reduce R_{in}/R_* because the dust envelope must both be faint in reflected light in the optical and also emit rather weakly in the J band (1.25 μm). Only in this case are we able to reproduce the spectral energy distribution of V4334 Sgr in the deep minimum via the occultation of the star by a dust cloud, with the J flux governed by R_{in}/R_*.

Figure 7 demonstrates the spectral energy distribution for July 1998, along with the computed spectral energy distribution. We also show the spectral energy distributions of the star and dust envelope. For a dust envelope made of fine graphite grains, the increase of $B–V$ between June 1997 and July 1998 would have been $\sim0\overset{m}{.}43$, $\sim0\overset{m}{.}1$ higher than observed. The model with fine graphite grains does not rule out a contribution from intrinsic reddening of Sakurai's star due to cooling.

In the CLS model, the emission from the dust envelope remains the same in the visual brightness minimum, whereas the star's brightness decreases due to the passage of its light through a dust cloud that has crossed the line of sight. The cloud's optical depth is $\tau(V) \approx 6.0$, and it must consist of fairly large ($a_{gr} \geq 0.8$ μm) grains in order not to strongly redden the object.

In addition, the cloud must absorb fairly strongly in the J band, to ensure that the combined flux from the star and dust envelope does not exceed the observed value. This condition is met by grains with $a_{gr} \geq 0.15$ μm. This result was not expected *a priori*. In Fig. 7, we display the star's spectral energy distributions after passage of its radiation through a dust cloud consisting of grains with $a_{gr} = 0.1$ and 0.15 μm. The grains with $a_{gr} = 0.1$ μm do not yet produce sufficient J absorption. In other words, even grains that large, which absorb neutrally in the optical, are not "neutral" enough from 0.55 to 1.25 μm (the $\tau(\lambda)$ relation is not flat enough).

Figure 7 shows the spectral energy distribution of V4334 Sgr for October 1998, along with the computed

energy distribution formed as the sum of the emission from a dust envelope with fine grains and the star's radiation after passage through a dust cloud with grains with $a_{\text{gr}} = 0.15$ μm. This is the combination of grain sizes capable of reproducing the deep minimum in the CLS model.

6. CONCLUSIONS

We see that the CCC model can reproduce all the characteristics of Sakurai's star observed in 1997–1999 fairly well, without considerably changing the inner radius of the dust envelope (grain-condensation zone). The dust envelope appeared early in 1997 and had a mass of $M_{\text{dust}} \geq 2 \times 10^{-7} M_{\odot}$ by the Fall of 1999.

A gas-envelope formation scheme with fixed grain-condensation temperature leads to somewhat better agreement between the computed curve and the spectral energy distribution of V4334 Sgr in July 1998. However, it is not able to explain the mass increase of the dust envelope.

In the CCC model, the first deep minimum is associated with a density discontinuity in the envelope of Sakurai's star, which reached the grain-condensation zone by August 1998. This discontinuity probably formed early in 1997, when V4334 Sgr became a carbon star, and its luminosity considerably increased. If this is the case, an envelope expansion rate of ~50 km/s implies an estimated distance to the star of ~3.7 kpc.

Creation of the necessary density discontinuity requires an increase of the mass-loss rate by a factor of about four. Prior to the first deep brightness dip in August 1998, the mass-increase rate for the dust envelope was $\sim 10^{-10} M_{\odot}$/day.

The CCC model requires that the dust envelope's optical depth be dominated by large grains that absorb neutrally in the optical. However, this does not imply an absence of fine grains. The number of fine grains could exceed that of large grains by more than an order of magnitude.

The CLS model applied to the first deep visual dip of V4334 Sgr leads to restrictions on parameters of the dust envelope and the dust cloud in the line of sight that seem improbable. The subsequent visual dimming and disappearance of Sakurai's star in the optical cannot be explained in this model.

In conclusion, we would like to comment that the Japanese amateur astronomer Sakurai discovered this unique object, a manifestation of the second birth of a red giant from the already "dying" nucleus of a planetary nebula, when the "newborn" was more than one year old. In the case of an additional 2.5-year delay, this discovery could not have been made.

REFERENCES

1. Yu. Sakurai, IAU Circ., No. 6322 (1996).
2. H. W. Duerbeck and S. Benetti, Astrophys. J. **468**, 111 (1996).
3. I. Iben, Astrophys. J. **277**, 333 (1984).
4. F. Kerber, J. Koppen, M. Roth, and S. C. Trager, Astron. Astrophys. **344**, L79 (1999).
5. D. Pollaco, Mon. Not. R. Astron. Soc. **304**, 127 (1999).
6. F. Kerber, H. Gratl, and M. Roth, IAU Circ., No. 6601 (1997).
7. S. P. S. Eyres, A. Evans, T. R. Geballe, *et al.*, Mon. Not. R. Astron. Soc. **298**, 37 (1998).
8. V. P. Arkhipova, V. F. Esipov, R. I. Noskova, *et al.*, Pis'ma Astron. Zh. **24**, 297 (1998) [Astron. Lett. **24**, 248 (1998)].
9. S. Kimesvenger, H. Gratl, F. Kerber, *et al.*, IAU Circ., No. 6608 (1997).
10. W. Liller, M. Janson, H. W. Duerbeck, *et al.*, IAU Circ., No. 6825 (1998).
11. G. Jacoby and O. De Marco, IAU Circ., No. 7065 (1998).
12. V. P. Arkhipova, R. I. Noskova, V. F. Esipov, and G. V. Sokol, Pis'ma Astron. Zh. **25**, 711 (1999) [Astron. Lett. **25**, 615 (1999)].
13. G. Jacoby, IAU Circ., No. 7155 (1999).
14. M. Asplund, B. Gustafsson, N. K. Rao, *et al.*, Astron. Astrophys. **332**, 651 (1998).
15. A. M. Tatarnikov, V. I. Shenavrin, B. F. Yudin, *et al.*, Pis'ma Astron. Zh. **26**, 580 (2000) [Astron. Lett. **26**, 506 (2000)].
16. H. W. Duerbeck, S. Benetti, A. Gautschy, *et al.*, Astron. J. **114**, 1657 (1997).
17. F. Kerber, J. A. D. I. Blommaert, M. A. T. Groenewegen, *et al.*, Astron. Astrophys. **350**, L27 (1999).
18. A. M. Tatarnikov, V. I. Shenavrin, and B. F. Yudin, Astron. Zh. **75**, 428 (1998) [Astron. Rep. **42**, 377 (1998)].
19. H. W. Duerbck, W. Liller, C. Sterken, *et al.*, Astron. J. **119**, 2360 (2000).
20. M. W. Feast, Mon. Not. R. Astron. Soc. **285**, 339 (1997).
21. B. T. Drain, Astrophys. J., Suppl. Ser. **57**, 587 (1985).
22. M. P. Egan, C. M. Leung, and G. F. Spagna, Comput. Phys. Commun. **48**, 271 (1988).
23. M. Asplund, D. L. Lambert, T. Kipper, *et al.*, Astron. Astrophys. **343**, 507 (1999).
24. D. K. Lynch, R. W. Russel, and C. J. Rice, IAU Circ., No. 6952 (1998).
25. H. U. Kaeufl and B. Stecklum, IAU Circ., No. 6938 (1998).

Translated by N. Samus'

Detection of Prolonged, Extremely Faint Decimeter Bursts on the Sun

V. M. Bogod[1] and L. V. Yasnov[2]

[1]*Nizhniĭ Arkhys, Special Astrophysical Observatory, Russian Academy of Sciences, Russia*

[2]*Research Institute of Radiophysics, St.Petersburg State University,*
Ul'yanovskaya ul. 1, Staryĭ Peterhof, St. Petersburg, 198904 Russia

Abstract—We report the detection of long-lived sources of radio bursts accompanied by polarized background emission in solar active regions. Both types of radio sources were detected at several decimeter wavelengths in observations on the RATAN-600 radio telescope in one-dimensional scans in intensity and circular polarization with a sensitivity of about 5–10 Jy. The degree of polarization is from 70 to 100%. The microburst and background sources exist for several days and appear at sites of prolonged energy release. The typical duration of an individual microburst is about 1–2 s, and the time interval between microbursts is about 3–5 s. A negative microburst frequency drift of about 100 MHz/s or more is also observed. Some interpretations of the microburst and background sources are discussed. The most probable microburst model involves the generation of radio emission via the plasma mechanism, with the upward propagation of fast electrons above an active region. In this case, the required energy of the Langmuir waves is 2×10^{-8} of the heat energy of the background plasma. Microbursts appear in different places in an active region. New methods for determining the magnetic-field intensity in the regions of generation of the decimeter-wavelength emission are proposed. © *2001 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

For the first time, we investigate here prolonged (lasting several days) nonthermal energy release in a solar active region using radio observations. Earlier, conclusions about the possible prolonged existence of nonthermal electrons were based on the unusual radio spectra observed for some parts of active regions, such as "halo" [1] and "peculiar" sources, which are usually located above the magnetic-polarity separatrices of active regions [2]. However, it was difficult to reconcile the long-lived stability of the emission of such sources with the presence of fast electrons. For this reason, until now, observations at centimeter and decimeter wavelengths have provided no direct evidence for the prolonged generation of energetic particles in active regions.

In our RATAN-600 radio observations, we recorded pulsed emission from one active region over a prolonged time (up to several days). This emission emerges against the background of the fairly stable active-region emission with ordinary polarization, which has a steeply growing spectrum at wavelengths from 20 to 30 cm. The fluxes of the radio microbursts are rather low (0.001 sfu), and the emission is 100% circularly polarized in the sense corresponding to the ordinary mode. The duration of a single microburst is about 1–2 s.

The time scales for variations in the intensity of the emission of solar nonthermal electrons are quite varied. Weak manifestations of burst activity, similar to the radio microbursts studied here, include quasi-periodic pulses of decimeter-wavelength flare emission with periods from 0.1 to 5 s [3–6]. The characteristics of periodic pulses of decimeter-wavelength flare emission are given in [7]. These pulses may result from MHD disturbances of the emission region, relaxation oscillations due to the interaction of plasma waves with plasma particles, quasi-periodic acceleration, and the injection of energetic particles [8–11]. The average durations of hard X-ray and type-III burst pulses are shorter than 1 s, but events lasting longer than 1 s are also observed fairly frequently. The durations of decimeter-wavelength pulses, and especially of millisecond spikes, are considerably shorter [12]. However, all these types of variations have been observed during bursts, that is, during transient events.

The prolonged radio microbursts and accompanying background source described below have parameters similar to those of noise storms. Noise storms include both brief bursts with durations from 0.1 to 3 s and background emission. Noise storms can last several days. The polarization of type-I bursts corresponds to that of ordinary waves. The background emission of noise storms has the same polarization. Noise storms are associated with various magnetic structures, as a rule, active regions with strong magnetic fields [13–16]. The distinction between noise storms and prolonged radio microbursts is, first and foremost, associated with the emission peak frequencies (300 MHz for noise storms and 1000 MHz for prolonged microbursts), and possibly also with the
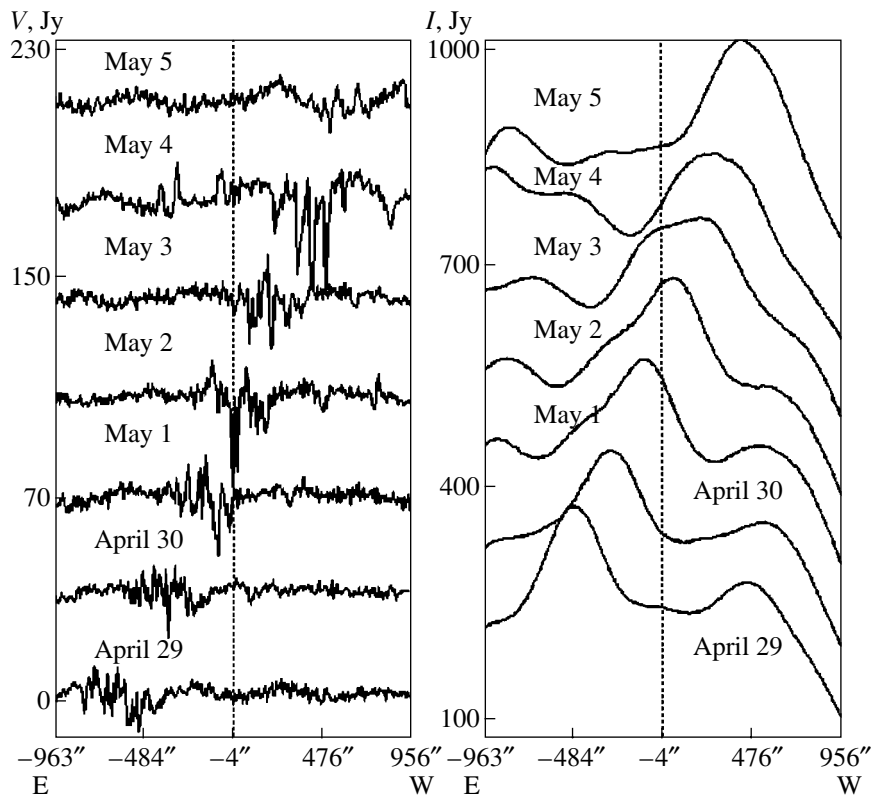
Fig. 1. Records of the polarized (*V*) and unpolarized (*I*) emission at 30.46 cm from April 29 to May 5, 1998. The *I* emission of the active region is accompanied by pulsed emission in the circular-polarization channels.

frequency drifts (up to tens of MHz/s at 200 MHz for noise storms and up to 100 MHz/s for prolonged microbursts).

The spectra of the background intensity and polarization also have interesting properties. At long wavelengths (about 30 cm), a flux increase and reversal of the sign of circular polarization are observed in the transition from the extraordinary to the ordinary mode.

Our observations of the polarized and unpolarized components over a broad wavelength range (1.8–32 cm) have enabled us to study the emission regions for the microburst and background sources and to determine the radio emission mechanisms in these regions.

In Section 2, we describe decimeter-wavelength observations of prolonged radio microburst sources and the background emission of the associated active regions using the RATAN-600 radio telescope. Section 3 analyzes some characteristics of the prolonged microburst emission: frequency drift, quasi-periodicity, height, and the random arrangement of the emission regions. In Section 4, we consider interpretations of the prolonged microburst emission based on the plasma mechanism for the radio emission. Section 5 addresses the spectral characteristics of the emission of the active region in which the prolonged microbursts arise. We briefly summarize our conclusions in Section 6.
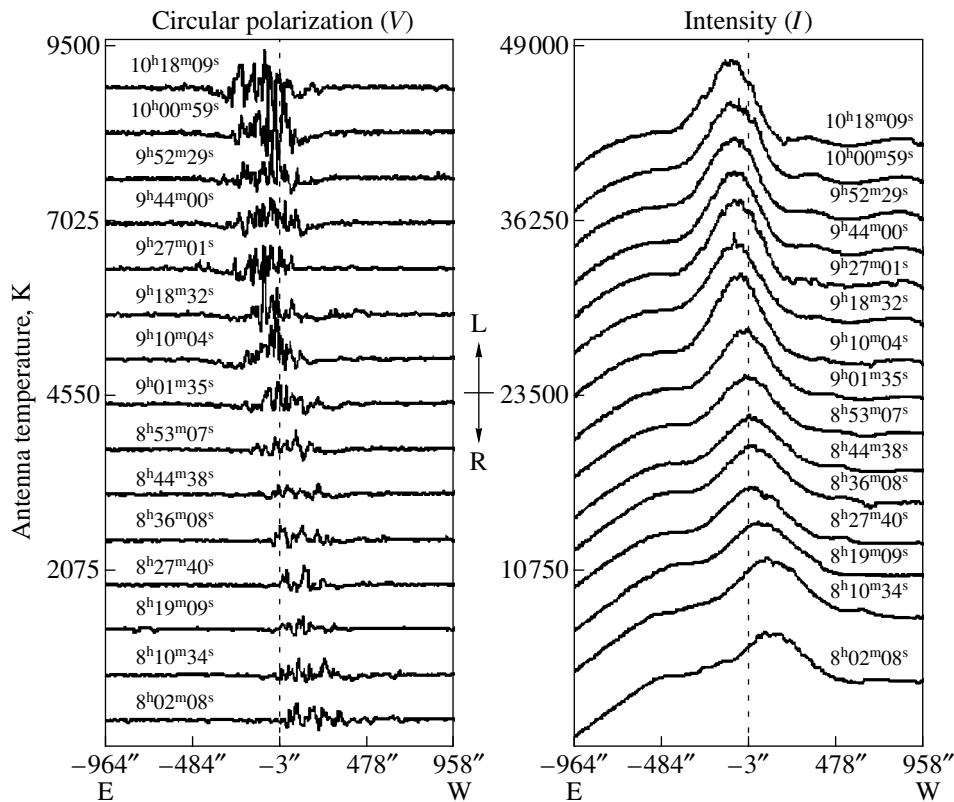
2. OBSERVATIONS

The observations were carried out using the broadband panoramic spectrum analyzer of the RATAN-600 radio telescope at 0.95–16.7 GHz in 35 frequency channels [17]. We used the radio telescope in the "southern sector + periscope reflector" mode [18], in which the telescope beam is one-dimensional with sizes θ_x [arcsec] = 0.85 λ [mm], θ_y [arcmin] = 0.75 λ [mm]. In our observations of the solar disk with a collecting area of 500 to 1000 m^2 and a time constant of 0.2 s, the flux sensitivity was 5–10 Jy throughout the wavelength range observed. This is the highest sensitivity for solar observations for existing instruments in the world. All receivers recorded both intensity and circular polarization.

Our search for faint bursts at several of the panoramic spectrum analyzer frequencies in the 30-cm band was successful. We detected prolonged radio microbursts and an accompanying background source. We present results for two observing sessions on the RATAN-600 telescope in azimuthal mode, on April 29–May 5, 1998 and May 19–20, 1999, from 7:00 to 11:00 UT.

2.1. Observations of April 29–May 5, 1998

Figure 1 shows records of solar radio scans obtained at 30.46 cm on the "southern sector + periscope" antenna system in intensity *I* and circular polarizations *V*. The observations were carried out in the transit mode once

Fig. 2. Repeated azimuthal observations of microbursts on May 19, 1999 from 08:00 UT to 10:20 UT. The curve labels indicate the times of observational culmination in each azimuth. For clarity, the constant component of the signal has been subtracted.

per day at local noon (about 9:00 UT) from April 29–May 5, 1998.

The right-hand plot in Fig. 1 presents scans of the Sun from the Eastern to the Western limb in intensity I and left-hand polarization V (left-hand circular polarization has a positive sign). The vertical scales are in flux units (1 Jy $=10^{-4}$ sfu). A comparison of the I and V scans shows that, as the active region observed in I moves across the disk over one week, the region of the pulsed sources in the polarized emission also moves.
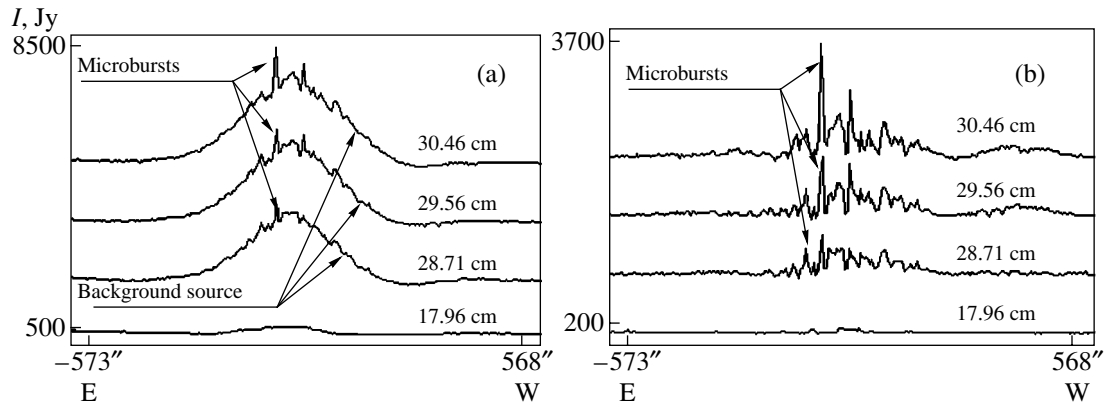
2.2. Observations of May 19–20, 1999

Certain technical improvements of the antenna–receiver complex were realized in the May 1999 observations, including the possibility of repeated azimuthal observations in the "southern sector + periscope" system, improved calibrations of spectral observations over the entire frequency band of the panoramic spectrum analyzer, and the addition of a new wavelength, 31.41 cm.

Figure 2 presents several azimuthal observations of an active region located close to the meridian in longitude and at a latitude of about 70° north on May 19, 1999. The left plot shows the behavior of the radio microburst polarization at 31.41 cm over a 2.20-h interval. The right plot presents the intensity scans. The shift

of the pulse packet and center of gravity of the source intensity is due to the azimuthal rotation of the beam during the observations. These observations demonstrate the continuous existence of microbursts in an active region over more than 2 h, and Fig. 1 suggests an even longer duration of one week. This is the first time that the prolonged existence of nonthermal processes at a such low levels of emission has been recorded.

The activity of this region was much higher on May 19 than on May 20, in both intensity and circular polarization (left-hand circular polarization has a positive sign). The emission consists of a left-hand polarized background component, corresponding to the ordinary mode, and a weak pulsed component. Further, we performed Gaussian filtration and subtraction of the large-scale sources in order to study the pulsed component of the polarized emission. The result is presented in Fig. 3 for several wavelengths for observations at the central azimuth at about 9:00 UT on May 19, 1999; this shows a clear detection of very faint microbursts.

We can see a high correlation of the pulsed sources at adjacent wavelengths in Fig. 3. However, the pulsed intensity sharply decreases at 17.96 cm. Unfortunately, the existing receiving complex does not cover the wavelength band from 18 to 28 cm, and we could not trace the growth of the microburst spectrum in this interval.

Fig. 3. Simultaneous observations of microbursts at several decimeter wavelengths on May 19, 1999, at azimuths +2° from the meridian (a) without and (b) with the background source subtracted. It is clear that, with increasing wavelength, a background source with superposed weak microburst pulses appears in the intensity channels. Note the high correlation of the microbursts at the long wavelengths.
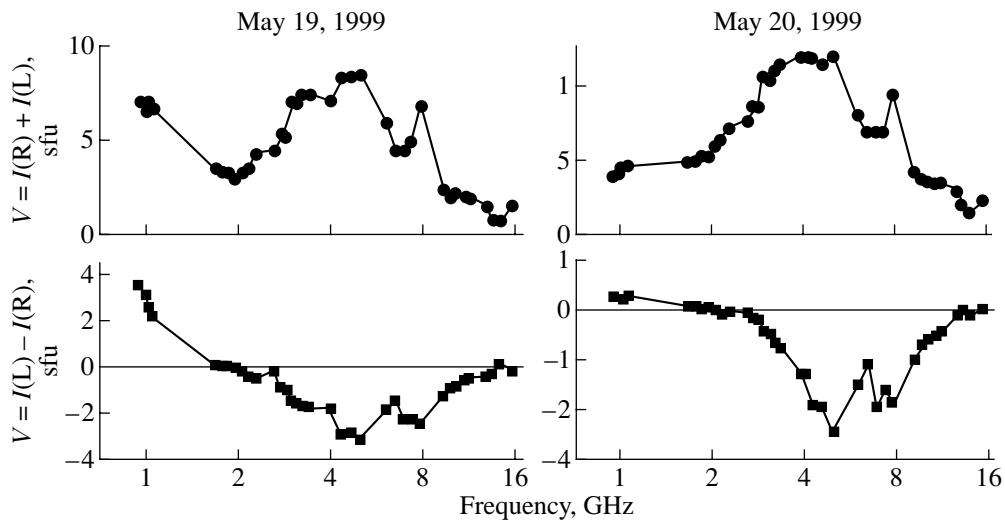

Fig. 4. Flux spectrum of the active region for May 19 and 20, 1999 in intensity (top plots) and circular polarization (bottom plots).

Figure 4 presents the spectrum of the entire active region throughout the observed frequency band from 0.95 to 16.7 GHz in intensity I and circular polarizations V for May 19 and 20, 1999. The I and V spectra for May 20 display a frequency dependence typical for thermal emission, peaking at 5–6 GHz, with the sign of the circular polarization corresponding to the extraordinary mode. May 20 provides an example of radio emission from a stable active region. On May 19, however, we note a growth of the intensity at low frequencies (owing to the appearance of the background source) and a sharp increase of the signal in the left circular polarization at wavelengths from 17.96 to 31.41 cm, which was completely absent from the observations of May 20. Note also that the polarization at shorter wavelengths is right-handed, corresponding to the extraordinary wave. The degree of polarization of the background

source at about 1 GHz is 60–70%, and the degree of polarization of the microbursts reaches 100%.

We compared the observations of May 19 and 20 with magnetic-field maps (Mt. Wilson Observatory) and meter-wavelength observations (Nancay radioheliograph). A small change of the magnetic-field structure on the photosphere during these two days was noted, and an intense noise storm above the active region studied (in the northern part of the disk) was recorded on May 19, at both 327 and 164 MHz. However, its intensity decreased sharply on May 20, and the solar background emission became appreciable. Thus, the meter-wavelength observations suggest a correlation with our observations at decimeter wavelengths, and probably indicate the detection of noise-storm emission at high frequencies, up to 1.1 GHz.
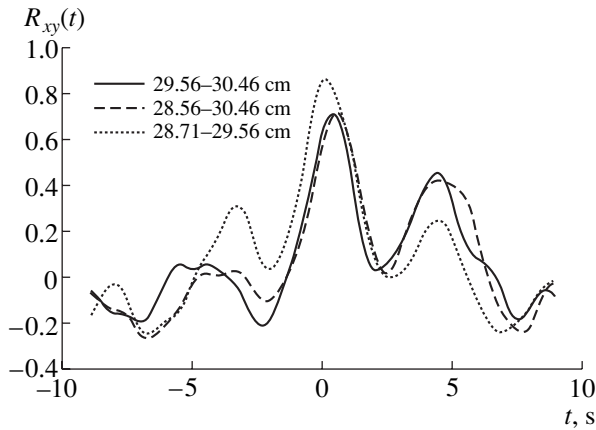
$R_{xy}(t)$

Fig. 5. CCF $R_{xy}(t)$ of the microburst polarized emission at 28.71, 29.56, and 30.46 cm for May 3, 1998.

Position, arcsec

Fig. 6. Positions of the maxima of the microburst emission and active-region intensity in diurnal scans as the active region moves over the disk. The horizontal axis plots the source heliolongitude in degrees.

3. RESULTS OF DATA REDUCTION

Let us analyze the observational data for April 29–May 5, 1988 in more detail. First, we compute the cross-correlation function (CCF) of the polarized microburst recordings at 28.71, 29.56, and 30.46 cm on May 3, 1998. The results are given in Fig. 5, which shows that a frequency drift is present in the bursts.

Fitting Gaussians to these curves, we can refine the position of the CCF peak and estimate the frequency-drift parameters. The frequency drift rate was –120 MHz/s for the 28.71–29.56-cm band, –90 MHz/s for the 28.71–30.46-cm band, and –60 MHz/s for the 29.56–30.46-cm band. However, our analysis shows that the most powerful radio microbursts in this session had a frequency drift exceeding 600 MHz/s.

Figure 5 also shows that the period for the radio microbursts is 3–5 s. Figure 1 indicates a change of the relative locations of the peaks in I and V as the active region moves across the disk. Let us analyze this phenomenon in more detail.

We determined the maxima of the prolonged microburst emission from the scans for the polarization V, and the maxima of the emission for the entire active region from the scans for the intensity I (Fig. 1) as the active region moves across the solar disk. In most cases, we determined the peak position through Gaussian fitting. The systematic decrease of the offset between the peak positions (Fig. 6) may reflect the difference between the height of the polarized source (microburst source) and the mean height of the dominant source (from the entire active region). Note, however, that this interpretation is valid only if the source is small, which is not the case here.

Note also the presence of a break in the curves between scans three and four. The active region underwent substantial restructuring in this period. It became more compact, as was immediately reflected in the position of the prolonged microburst source. Thus, the
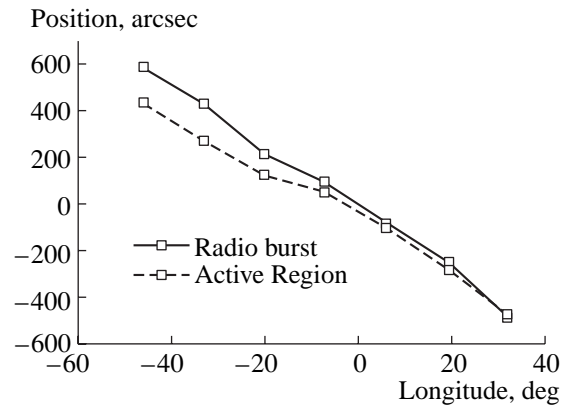
sources are closely related (probably, via the overall structure of the magnetic field).

We calculated the absolute height of the microburst source taking into account the differential rotation of the Sun. We assumed that the photospheric part of the active region with which the prolonged microburst source is associated rotates in the same way as recurrent sunspot groups, with angular speed $(14.38°–2.77°\sin^2\varphi)$, where φ is the heliographic latitude of the source. We found the height of the prolonged microburst source to be 15 000 km.

For some microburst sources (e.g., for May 3, 1998), the frequency drift is rather slow. In this case, we can assume that the emission comes from a site near the top of a coronal loop with a fairly uniform electron-density distribution. The mean microburst duration, estimated from the CCF, is about 1–2 s.

We observed similar microburst periods on May 19, 1999. Figure 7 shows the correlation function for the

$R(t)$
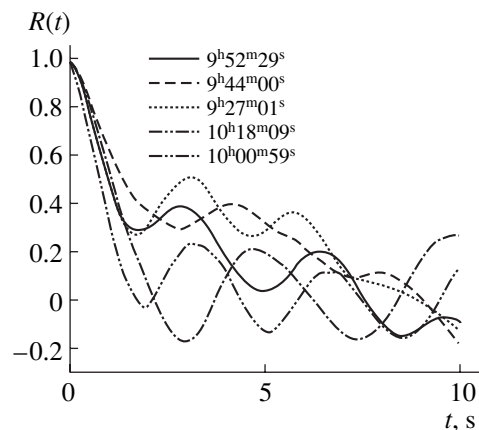
Fig. 7. Correlation function of the microburst emission at 31.4 cm for various times.

microburst emission at 31.4 cm for various times. The microburst repetition period is 3–5 s. Note that there is an almost discrete change of the microburst period from 3 to 5 s. For this observing session, in all cases, the CCF peaked at zero; for the conditions of our observations, this implies a frequency drift rate higher than 800 MHz/s for all the microbursts.

For the observations of May 19, 1999, we can determine the spatial location of the region in which the prolonged microburst is generated. We use the fact that the source of the prolonged microburst is displaced relative to the vertical circle passing through the disk center as the Sun rotates with respect to a horizontal coordinate system. The magnitude of this displacement and the calculated CCF yield this information.

On May 19, 1999, the active region was at a heliolatitude of 70° in the northern hemisphere. In the case of cospatial sources generating microbursts at different times, the difference between the CCF peaks should be 46″, or 3.2 s. However, for the times of highest microburst intensity ($10^h18^m09^s$–$10^h00^m59^s$ and $10^h00^m59^s$–$99^h52^m29^s$), this difference is only about 1 s, or 14″. Thus, we can firmly conclude that the microbursts arise at different sites of the active region.

4. INTERPRETATION OF THE MICROBURST EMISSION

The parameters of the prolonged microburst sources (in particular, their similarity to type-III bursts), as well as their connection with noise storms, suggest the plasma mechanism for the radio emission at the sites of the prolonged microbursts. This mechanism supposes a continuous pulsed injection of electrons above a magnetic loop. The electrons stream upward toward the lower-density region, and radio emission can arise at either the fundamental or the second harmonic of the plasma frequency. Let us first consider emission at the second harmonic, then at the fundamental plasma frequency.

4.1. Emission at the Second Harmonic of the Plasma Frequency

We used the formulas for the emission at the first and second harmonics of the plasma frequency given in [19]. We also took into account thermal bremsstrahlung absorption of the radio emission, which can be significant at the second harmonic of the plasma frequency. In these calculations, we assumed that Langmuir waves (L-waves), being in resonance with the electron beam, have phase velocities close to the beam velocity. The speed of the electrons was taken to be $0.3c$. We took the size of the source along the line of sight for emission at a fixed frequency to be equal to the size of the region l along the path travelled by the radio signal, assuming a constant electron density (see, for example, [20]).

The calculations show that the emission region is optically thin in this case (due to the low intensity of the emission). The resulting formula for the flux in solar units (for $\tau < 1$) is

$$F = \frac{2.74 f^2 H s T_e^2 v_p^3 w^3}{10^{71}}, \qquad (1)$$

where f is the frequency in Hz, T_e the electron temperature in K, v_p the phase velocity of the L-waves in cm/s, w the ratio of the energy of the L-waves to the plasma heat energy $n_e k_b T_e$ (k_b is the Boltzmann constant), s the source area in cm², H the electron-density scale height in cm, and n_e the electron density in cm⁻³.

Assuming $s = l^2$, we can write

$$F = \frac{6.03 f^2 H^3 T_e^4 w^2}{10^{48} v_p}. \qquad (2)$$

Adopting $T_e = 2 \times 10^6$ K, $v_p = 0.3c$, and $H = 10^{10}$ cm (a typical value for the corona), with $f = 10^9$ Hz, we can obtain an estimate for the L-wave energy from the formula

$$F = 1.07 \times 10^{16} w^2. \qquad (3)$$

When $F = 0.001$ sfu, the energy of the plasma waves relative to the heat energy is $w = 3 \times 10^{-10}$.

In spite of the low L-wave energy, the emission at the second harmonic of the electron plasma frequency encounters a certain difficulty: for an isotropic plasma-wave spectrum, the sign of the resulting polarization corresponds to an extraordinary wave [21]. At the same time, a comparison of the background polarizations for the active region and microburst shows that the ordinary mode dominates in both cases.

Another difficulty connected with this mechanism is that the Langmuir waves become isotropic due to scattering on ion-density fluctuations, which, in particular, creates the conditions required for radio emission at the second harmonic of the plasma frequency. Let us estimate the isotropization time scale using the expression [22]

$$\tau_{iso} = \frac{10}{2\pi f_p w}. \qquad (4)$$

For $f_p = 1000$ MHz and $w = 3 \times 10^{-10}$, we find $\tau_{iso} = \frac{50}{3\pi}$ s, which is considerably longer than the duration of a microburst. This implies that emission with the observed duration cannot arise at the second harmonic of the plasma frequency if w is so small.

4.2. Emission at the Plasma Frequency

The difficulties associated with emission at the second harmonic of the plasma frequency can be overcome if the emission is at the fundamental harmonic.

4.2.1. Polarization of the emission and magnetic-field intensity in prolonged microburst sources. The

primary emission at the first harmonic is weakly polarized; nevertheless, a high degree of polarization can be provided at a much lower magnetic-field intensity. Here, we must take into account the different conditions for the escape from the corona for ordinary and extraordinary waves. If the well-known inequality [23]

$$\frac{f_p^2}{f^2} > 1 - \frac{f_b |\cos(\theta)|}{f},$$ (5)

is satisfied in the emitting regions, the emission will be completely ordinary-polarized. Since the frequency of the plasma emission is

$$f^2 = f_p^2 \left(1 + \frac{3v_t^2}{v_p^2}\right),$$ (6)

where v_t is the thermal velocity of the electrons, this inequality will become

$$\frac{f_b |\cos(\theta)|}{f} > \frac{3v_t^2}{v_p^2}.$$ (7)

Hence, we find a lower limit for the magnetic-field intensity in the source. For $f = 1000$ MHz, $T_e = 2 \times 10^6$ K, $v_p = 0.3c$, and $\cos(\theta) = 1$, we find $B > 4$ G; i.e., this case requires considerably lower magnetic-field intensities than does emission at the second harmonic of the plasma frequency.

According to our observations, microbursts occur at frequencies where the polarization for the entire active region changes sign. Since this sign change is stable and does not depend on the position of the active region on the solar disk, it is probably associated with a smaller height gradient for the magnetic field compared to that for the electron density, and the fact that $f_b > f_p$ in the region of emission at the first harmonic of the gyrofrequency for sufficiently low frequencies. In this case, according to (5), the emission has ordinary polarization. Determining the frequency at which the ordinary-wave emission of the main source begins to dominate provides a new method for estimating the magnetic-field intensity in the active region in which the prolonged microbursts are generated (if $f_b \geq f_p$).

4.2.2. Energy of langmuir waves. Below we give formulas for calculating the emission at the plasma frequency. In this case, the bremsstrahlung optical depth significantly exceeds unity, and the flux in sfu will be

$$F = \frac{1.7 f s T_e^{3/2} v_p^2 w}{10^{49} \ln(4.3 \times 10^6 T_e^{2/3} f^{-2/3})}.$$ (8)

Assuming, as earlier, $s = l^2$, $T_e = 2 \times 10^6$ K, $v_p = 0.3c$, and $H = 10^{10}$ cm, we obtain for $f = 10^9$ Hz the following estimate for the L-wave energy:

$$F = 4.8 \times 10^4 w.$$ (9)

Thus, we have for 0.001 sfu $w = 2 \times 10^{-8}$; as expected, this is considerably greater than for emission at the second plasma-frequency harmonic.

4.2.3. Period, duration, and frequency drift. Let us estimate the radio microburst parameters assuming emission at the fundamental harmonic of the plasma frequency.

It is noted in [24] that the plasma emission can be modulated by fast magnetosonic waves with period $T \approx L_t/v_a$, where L_t is the transverse size of the coronal arch and v_a is the Alfvén velocity. Assuming $f_p = 1000$ MHz and $L_t = 10^9$ cm, we obtain $T = 555/B$. Thus, if $B = 100$–150 G, the microburst period is estimated to be 3–5 s, in agreement with the observational data.

We cannot exclude the possibility of quasi-periodic generation of fast electrons. In particular, Aschwanden *et al.* [7] reported the generation of fast electrons with an average period of 1.6 s for the flare of June 22, 1989, 14:45 UT. It is quite possible that this period could be somewhat longer for quieter active regions.

The duration of a microburst can be determined by the time for an individual electron to travel through the region of efficient plasma emission at a fixed frequency, by the time scale for quasi-linear relaxation of the Langmuir waves, or by the duration of an event generating fast electrons. Our calculations show that the duration of a microburst is primarily determined only by the length of the time during which fast electrons are generated.

Let us summarize the conclusions following from the frequency-drift measurements for both observing sessions.

For the radio microburst series of May 1998, weaker microbursts arise near the top of a coronal arch. In this region, the electron density is nearly uniform, and the microbursts have a small frequency drift, of the order of 100 MHz/s. We observed only one more powerful microburst that arose due to electrons escaping upward from the coronal arch toward the region with a height-dependent electron density.

In the second microburst series (May 19, 1999), we are dealing with a more open magnetic structure, such as a streamer, and the electrons escape from the region in which they are generated upward toward the propagation region, with its varying background-plasma density. Here, we can estimate the frequency drift using the barometric formula for the height dependence of the electron density. In this case,

$$\frac{df}{dt} = -\frac{10^6 f_p}{T_e}.$$ (10)

Assuming a typical coronal temperature $T_e = 1.5 \times 10^6$ K, we obtain –670 MHz/s for the frequency drift. This value is close to the estimated limits derived from the observations. Since the frequency-drift limit for May 19, 1999 was slightly higher than this value, this
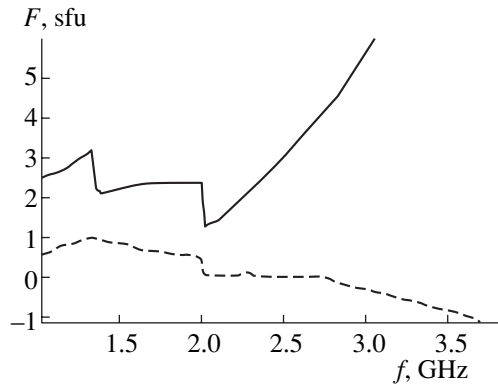
Fig. 8. Radio flux spectrum $F(f)$ (the solid curve is L + R, and the dashed curve L − R) calculated for $B_0 = 50$ G, $r = 7.5 \times 10^9$ cm, $n_e = 10^9$ cm^{-3}, and $T_e = 2 \times 10^6$ K.

may suggest the propagation of fast electrons in an atmosphere with a smoother variation of the background-plasma electron density than specified by the barometric formula (as, for example, in a streamer).

To conclude, we note that the magnetobremsstrahlung mechanism for the radio emission encounters difficulties explaining a frequency drift of 100 MHz/s, since the speeds of the electrons responsible for the decimeter-wavelength emission are close to the velocity of light. Another difficulty is that the emission should always have the extraordinary polarization.

5. INTERPRETATION OF THE DECIMETER-WAVELENGTH SPECTRUM OF THE ACTIVE-REGION BACKGROUND SOURCE

Our 1–2-GHz observations indicate the existence of a background source with a rather unusual spectrum (Fig. 4), whose intensity increases with decreasing frequency. This spectral behavior could be associated with nonthermal emission of mildly relativistic electrons, especially since the presence of prolonged microbursts suggests that such electrons could be generated in the active region. The difficulty with this interpretation is the short lifetime of the electrons, which is limited by either the thermalization time scale in the corona (no more than tens of seconds), or the time scale for escape from the generation region (a fraction of a second). In this connection, we will investigate the possibility of interpreting this spectral behavior in the framework of thermal magnetobremsstrahlung as the mechanism for the radio emission.

Consider a magnetic tube whose shape and magnetic-field intensity are described by a dipole approximation. We will assume that the dipole is perpendicular to the solar surface and is embedded at some depth beneath the photosphere. In this case, due to the effect of projection, the apparent area of the gyroresonant levels at the top of the loop decreases, and the angle

between the line of sight and the magnetic-field lines changes sign; this should lead to an intensity minimum and a change in the sign of the polarized emission. The loop parameters are the magnetic-field intensity B_0, loop width d_0 at the point that is most distant from the dipole axis, and the distance $2r$ of this point from the dipole axis. We calculated the radio flux from each gyroresonant level taking into account absorption in overlying layers. For the ascending branch of the loop, we included levels with larger gyrofrequency harmonic numbers s, and for the descending branch, levels with smaller values of s. We also took into account the degree of overlap of the gyroresonant layers. The fluxes from all layers were then added. We restricted our consideration to the calculation of four gyroresonant layers, since the effect of the fifth gyroresonant layer is insignificant. We adopted constant values for the electron density and temperature in the loop, $n_e = 10^9$ cm^{-3}, and $T_e = 2 \times 10^6$ K. Figure 8 shows the results of these calculations.

We can see that the flux has a minimum at 2 GHz. This feature of the flux spectrum depends only weakly on the loop parameters, provided, of course, that the optical depth of the gyroresonant layer near the top of the loop remains greater than unity. The loop parameters affect mainly the frequency at which the flux minimum occurs (which is determined by B_0, or the magnetic-field intensity at the loop top), as well as by the overall intensity and flux level.

Assuming that we observe a gyroresonant layer with $s = 3$ at the loop top at the frequency of the spectral minimum f_{mn} (as is precisely the case in our model), we can estimate the magnetic-field intensity. Thus, we have developed a fairly efficient method for measuring the magnetic field at the top of a loop. In the actual observations, the frequency of the minimum is $(1.5–2) \times 10^9$ Hz, yielding $B_{top} = 180–240$ G.

6. CONCLUSION

In our observations on the RATAN-600 radio telescope, we have for the first time detected the prolonged generation (lasting several days) of nonthermal decimeter-wavelength radio emission in the form of microbursts with fluxes of about 0.001 sfu. The microbursts had durations of 1–2 s, and recurred with periods of 3 and 5 s. The microburst frequency drift rate in the May 1998 session was about −100 MHz/s. The frequency drift rate for the 1999 session exceeded 800 MHz/s.

The peaks of the radio emission from the entire active region and for the microburst source were offset in position (in 1998). The maximum offset was reached at the eastern limb of the Sun (150 arcsec). As the sources moved toward the western hemisphere, the offset decreased. At the end of the observing session (six days later), the offset was close to zero. The microburst source height (in 1998) measured via comparison with the motion of recurrent sunspot groups was 15 000 km.

We have analyzed possible mechanisms for the generation of the microbursts. The best agreement with the observations is achieved for radio emission at the plasma frequency. To obtain the observed level of emission at the plasma frequency, the Langmuir turbulence level w must be 2×10^{-8}. In this case, the emission has ordinary polarization, due to propagation effects in the emission region. The only important thing is that the magnetic-field intensity be higher than 4 G. The observations of the polarized background emission for the entire active region suggest that this condition is satisfied with a large margin.

The prolonged microbursts arise in different places in the active region. Our analysis of the frequency drift shows that, in the May 1998 observing session, the microburst source was located near the top of a coronal loop, with the energetic electrons moving upward toward the top. In the May 1999 observing session, the microburst source was in an open magnetic structure, such as a streamer, with the density varying with height more smoothly than specified by the barometric formula for a coronal temperature of 1.5×10^6 K.

We have constructed a model for the main source based on a dipolar magnetic tube. We have proposed a method for estimating the magnetic-field intensity at the top of a magnetic tube using the frequency of the minimum in the intensity spectrum of the active region.

ACKNOWLEDGMENTS

REFERENCES

1. Sh. B. Akhmedov, V. M. Bogod, G. B. Gel'freĭkh, *et al.*, Soln. Dannye, No. 5, 68 (1985).

2. K. R. Lang, R. F. Willson, V. M. Bogod, *et al.*, Astrophys. J. **419**, 398 (1993).

3. B. L. Gotwols, Sol. Phys. **25**, 232 (1972).

4. O. Elgaroy and O. P. Sveen, Sol. Phys. **32**, 23 (1973).

5. H. J. Wiehl, A. O. Benz, and M. J. Aschwanden, Sol. Phys. **95**, 167 (1985).

6. M. J. Aschwanden and A. O. Benz, Astron. Astrophys. **158**, 102 (1986).

7. M. J. Aschwanden, A. O. Benz, B. R. Dennis, and V. Gaizauskas, Astrophys. J. **416**, 857 (1993).

8. M. J. Aschwanden, Sol. Phys. **111**, 123 (1987).

9. M. Pick and G. Trottet, Sol. Phys. **60**, 353 (1978).

10. G. Trottet, A. Kerdraon, A. O. Benz, and R. Treumann, Astron. Astrophys. **93**, 129 (1981).

11. B. R. Dennis, L. E. Orwig, and A. L. Kiplinger, in *Rapid Fluctuation in Solar Flares*; NASA Conf. Publ. **2449**, 123 (1987).

12. M. J. Aschwanden, B. R. Dennis, and A. O. Benz, Astrophys. J. **497**, 972 (1998).

13. K. Kai, Proc. Astron. Soc. Aust. **1**, 186 (1969).

14. J. Wild, Proc. Astron. Soc. Aust. **1**, 137 (1968).

15. O. Elgaroy, *Solar Noise Storms* (Pergamon, Oxford, 1977).

16. T. E. Gergely and M. R. Kundu, Sol. Phys. **41**, 163 (1975).

17. V. M. Bogod, S. M. Vatrushin, V. E. Abramov-Maximov, *et al.*, Astron. Soc. Pac. Conf. Ser. **46**, 306 (1993).

18. D. V. Korol'kov and Yu. N. Parijskij, Sky Telesc. **57**, 4 (1979).

19. V. V. Zaitsev and A. V. Stepanov, Sol. Phys. **88**, 297 (1983).

20. N. Gopalswamy, Astrophys. J. **402**, 326 (1993).

21. E. Ya. Zlotnik, Astron. Astrophys. **101**, 750 (1981).

22. A. G. Emslie and D. F. Smith, Astrophys. J. **279**, 882 (1984).

23. V. V. Zheleznyakov, *Radio Emission from the Sun and Planets* [in Russian] (Nauka, Moscow, 1964).

24. Yu. M. Rozenraukh and A. V. Stepanov, Astron. Zh. **65**, 300 (1988) [Sov. Astron. **32**, 153 (1988)].

Translated by G. Rudnitskiĭ

Coherent Structures in the Dynamics
of the Large-Scale Solar Magnetic Field

A. V. Mordvinov and L. A. Plyusnina

Institute of Solar–Terrestrial Physics, Siberian Division, Russian Academy of Sciences,
P.O. Box 4026, Irkutsk, 664033 Russia
Received November 8, 2000

Abstract—Variations in the mean solar magnetic field (MSMF) are studied in both the frequency–time and longitude–time domains. A wavelet analysis of the MSMF clearly demonstrates that variations in the mean field are not stationary. Combined with longitude–time diagrams for the background solar magnetic field (BSMF), the analysis reveals the emergence of the background field, which occurs discretely at intervals of 1.5–2 years. Based on an analysis of the fine structure in MSMF variations, we develop a numerical technique to study time-dependent heliographic-longitude distribution of the large-scale magnetic field. A detailed picture of the rotation of the large-scale magnetic field is derived for activity cycles 20–23. Coherent structures are detected in longitude–time diagrams obtained by deconvolving the MSMF series. These structures are related to discrete rigid-rotation modes of the large-scale magnetic fields. Various rotational modes coexist and replace one another. During the phase of activity growth, modes with periods of 27.8–28.5 days dominate, whereas a mode with a rotational period of about 27 days dominates during the decline phase. Occasionally, modes with periods of 29–30 days appear. Most structures in the longitude–time MSMF distribution correspond to similar structures in the BSMF distribution for the northern or southern hemisphere. Chronologically, the emergence of the BSMF has frequently been accompanied by changes in the solar rotational regime and has been correlated with variations in the polarity asymmetry in the course of the 11-year activity cycle. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The mean solar magnetic field (MSMF) is one of the basic physical parameters characterizing the Sun as a star, and largely determines space weather in the Solar System. This global index is defined to be the convolution of the distribution of the line-of-sight magnetic-field component over the solar disk with the weight function of the magnetograph [1, 2]. The mechanisms for the generation, advection, and diffusion of solar magnetic fields produce MSMF variations on a wide range of time scales [3]. Apparent variations in the MSMF result from the solar rotation and displacements of the Earth relative to the solar equator in the course of its orbital motion [4]. Time variations in the MSMF, their periodograms, and power spectra have already been studied in detail [5, 6] using measurements obtained at the Crimean, Stanford, and Sayan observatories [1, 2, 7]. Here, we investigate spectral and temporal regularities in variations of the MSMF using wavelet analysis, which is highly efficient for studies of complex time-dependent processes [8].

Although advances in helioseismology have brought about great progress in investigations of the solar rotation, the pattern of external manifestations of rotation inferred from various data is complex, inconsistent at some points, and depends on the level of magnetic activity [9–11]. Discrete rigid-rotation modes with periods of 27–30 days and torsional oscillations are observed against the background of the differential rotation [12].

According to [13] and our own studies, time variations in the MSMF are mainly determined by the evolution of the background solar magnetic field (BSMF). The solar rotation derived from the MSMF can be used to estimate the total contribution of the large-scale magnetic field. A separate analysis of BSMF data for the northern and southern hemispheres enabled the determination of their contributions to the net MSMF signal.

2. WAVELET ANALYSIS OF THE MSMF

We analyze here a time series of MSMF measurements obtained at the Stanford observatory, supplemented with early data from the Crimean observatory. We applied a continuous wavelet analysis [8] to study the combined sequence of MSMF data for 1968–2000. Wavelet spectra provide a detailed, informative picture of spectral and temporal variations in the MSMF on a wide range of time scales.

Figures 1a and 1b present the modulus of the continuous Morbet wavelet transform of the MSMF, the combined series of MSMF measurements, and its cumulative sum, which characterizes the apparent asymmetry in the polarities. The data from two observatories for 1975–1976 were averaged with the weights estimated in [14]. The wavelet spectrum indicates the distribution of oscillational energy on various time scales at various times. Distortions arise near the spectrum edges due to the fact that the analyzing wavelets partially extend
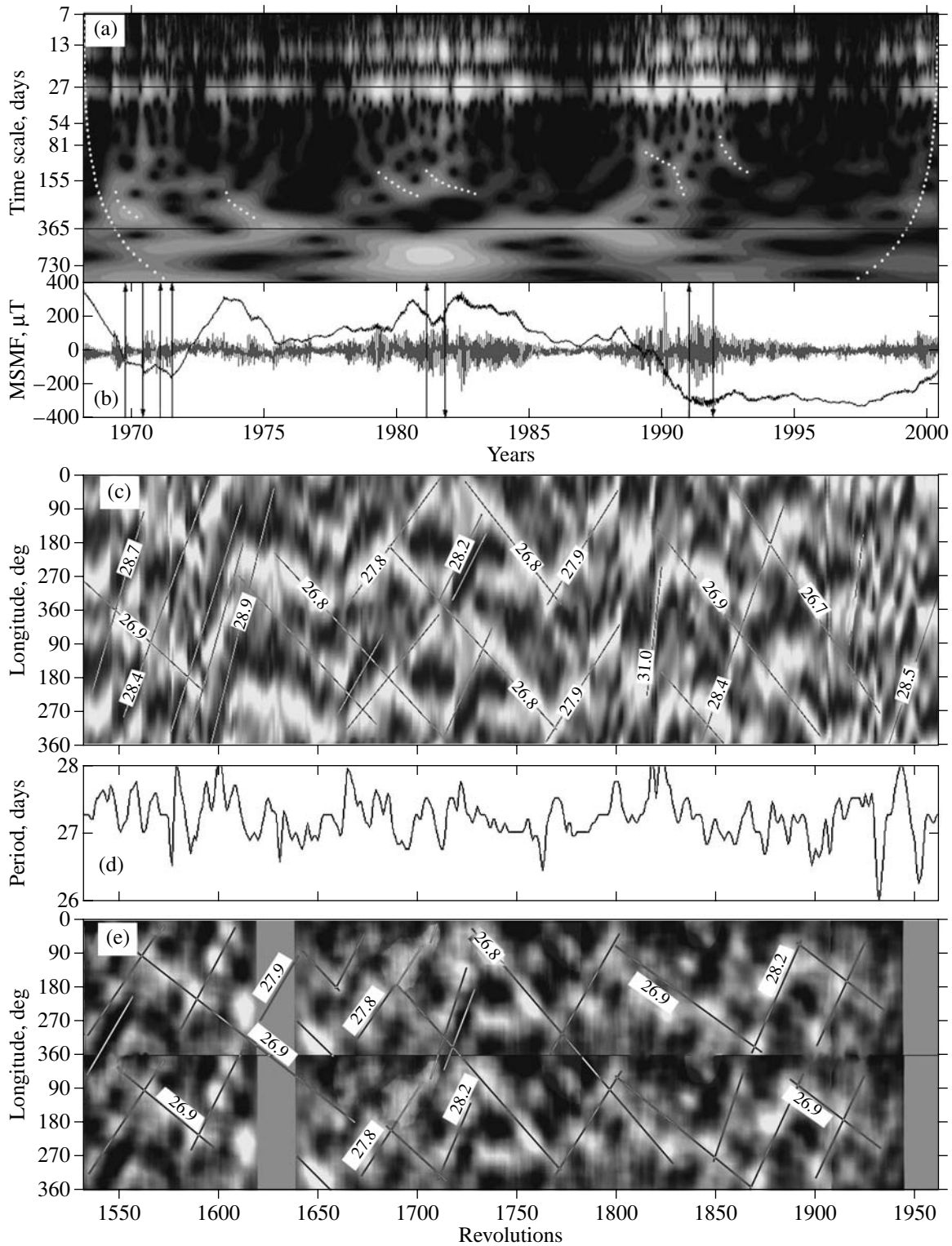
Fig. 1. (a) Normalized wavelet spectrum and (b) MSMF during 1968–2000. The spectral maximum corresponds to 0 dB and is shown by white areas. The light dots delimit the edge distortions of the spectrum and energy cascades. The upward and downward arrows mark polarity reversals in the northern and southern hemispheres, respectively. (c) Normalized values of the filtered MSMF component. (d) Rotational correlation period of the large-scale magnetic field. (e) BSMF diagram for latitudes 40° N–40° S.

beyond the record. Missing MSMF values were reconstructed based on an autoregression model [15]. The pattern of spectral–time MSMF variations is nonstationary and displays multiple scales. The energy in the spectrum is concentrated near the solar rotational period. However, the crude frequency resolution of our technique does not allow different rotational modes to be resolved, which produce a set of peaks in the MSMF power spectrum [5]. A period close to 13.5 days can also be seen; this is a manifestation of the four-sector magnetic-field structure. A marked increase in the spectral energy is observed on time scales near nine days, apparently due to the convolution of the magnetic-field distribution with the limb-darkening function [16].

In a horizontal section of the spectrum, on a time scale of about 27 days, we can see structures that manifest their presence in the amplitude modulation of the signal component with period close to the solar rotational period. The characteristic time separation between these structures is 1.5–2 years. Four such structures were observed in cycle 21, and five in cycle 22. These structures are related to events of emergence of the BSMF, which occur at intervals of about two years [17]. The amplitude of the 27-day MSMF variation increases with the emergence of a new portion of the large-scale magnetic field.

The annual variation is well defined in the MSMF spectrum in conjunction with the other long-term changes. At epochs of minimum activity, the annual variation stands out clearly against the weak MSMF variations, whereas it can barely be distinguished among other signal components during periods of high activity. The orbiting Earth changes its position with respect to the solar equator, producing annual variations in the line-of-sight contribution of the dipole component of the global magnetic field [4]. An interesting feature of the MSMF spectrum is the presence of structures on a time scale of about 155 days. The spectral energy grows in this range during the maximum-activity epochs of 1980–1982 and 1989–1991, and energy is transferred to larger time scales, so that an inverse energy cascade takes place. Similar structures can be seen in the wavelet spectra of the total solar irradiance and of the Ca II index. Oscillations with characteristic periods of about 155 days have also been detected in other solar magnetic indices, and may be related to the relaxation of large-scale magnetothermal disturbances in the course of the 11-year activity cycle [18].

3. SPATIO–TEMPORAL ANALYSIS OF THE MSMF

In the study of nonstationary and multimodal rotation of the Sun, with coherent patterns rigidly rotating against the background of the normal differential-rotation profile, the use of longitude–time diagrams [19] proves to be efficient, and enables reconstruction of the complete picture of the solar rotation from data on time variations of the MSMF and BSMF. In our study, the technique employed in [19] is further developed to analyze the distributions of the MSMF and BSMF in heliographic longitude.

To construct a longitude–time diagram from the MSMF signal, we used wavelet filtering to identify the component containing the effect of the rotational modulation due to the spatial nonuniformity of the large-scale magnetic field. We performed the wavelet filtering using an orthogonal wavelet expansion, which is best suited for analyses of time-dependent processes. Daubechies wavelets were chosen as the analyzing functions; they have good frequency–time localization, and the time scales of the wavelet functions used in the expansion are discrete and equal to integer powers of two. To analyze the MSMF signal as a function of time, $f(t)$, we performed the discrete wavelet transformation [8]

$$f(t) = \sum_{j,k=-\infty}^{\infty} c_{jk} \psi_{jk}(t),$$

where $\psi_{jk}(t)$ are wavelet functions forming an orthonormal basis and the coefficients of the wavelet expansion are determined by the relations $c_{jk} = \langle f \psi_{jk} \rangle$. The time scales of the main rotational-modulation effects range from 13 to 30 days. For this reason, during the wavelet filtering, only coefficients corresponding to discrete scales of 8, 16, and 32 days should be retained in the inverse discrete wavelet transformation. Then, the filtered component containing the main effects of the rotation has the form

$$f_{\text{rot}}(t) = \sum_{j=3}^{j=5} \sum_{k=-\infty}^{\infty} c_{jk} \psi_{jk},$$

with the values 2^j ($j = 3, 4, 5$) corresponding to the scales on which rotational modulation is manifest.

Further, the filtered component was subdivided into partial arrays corresponding in time to individual Carrington revolutions, and two-dimensional gray-scale diagrams were constructed for each rotation. The observation time within a Carrington revolution corresponds to the heliographic longitude of the event. If the process of measuring the MSMF is regarded as the convolution of the magnetic-field distribution with the magnetograph weight function, derivation of spatial parameters of the magnetic field from the time variation of the global index is the inverse procedure—the deconvolution of the MSMF time series.

We show the longitude–time diagram constructed via deconvolution of the MSMF twice in Fig. 1c, to trace the behavior of structures at the edges of the diagram. In the diagram, areas of positive polarity are lighter than the 50% gray background, while areas of negative polarity are darker. Since the intensity of repeated events is not important when studying rotation, we normalized the values of the filtered MSMF component in each revolution to the maximum value for that revolution. This made it possible to detect rota-

tion-related structures for various solar-activity levels in a form that enabled comparisons. Some features form diagonal chains in the diagram, indicating the presence of discrete rigid-rotation modes with periods of 27–30 days. In cycle 21, the distinct pattern of positive polarity is related to the development of activity. During the growing-activity phase, in 1977–1980, a rotational mode with a period of about 27.8 days dominates. At the epoch of activity maximum, the solar-rotation regime changes: starting from 1982, rotation with a period of about 28.2 days is occasionally observed; later, a rotational mode with a period of 26.8 days dominates during the entire declining-activity phase of cycle 21. At the epoch of activity minimum, the rotational regime changed again: the mode with a 26.8-day period disappeared by 1986. Positive-polarity structures again exhibit modes with periods of 27.9 and 31.0 days against the background of growing activity in cycle 22, and the pattern of rotation is more complex than in the preceding cycle. In the declining-activity phase of cycle 22, rotational modes with periods of 26.9 and 26.7 days are dominant. In cycle 20, the spatial distributions are smeared due to the high noise level during the initial period of observation. Nevertheless, a rotational mode with a period of 26.9 days can be distinguished, which dominates in the declining-activity phase. In addition, a mode with a period of 28.4 days emerged in 1972–1973. The negative polarity sometimes forms parallel structures, shown as dark areas.

Previous analysis of the rotation of the large-scale magnetic field based on a separate consideration of the two-sector and four-sector components led Obridko and Shelting [10, 11] to conclude that a two-sector structure with rotational periods of 27.5–28.5 days dominates at the maximum of the solar-activity cycle, while a four-sector structure with a 27-day recurrence period dominates in the declining phase. At the epoch of secondary maximum in 1972, the previously established order broke down, and a two-sector structure with a period of 28–28.5 days emerged. Thus, despite differences in the techniques used to form the data series and analyze them, Obridko and Shelting [10, 11] obtained essentially the same results that we have concerning variations in the rotation of the large-scale magnetic field in the course of the solar-activity cycle.

By estimating the displacements of the features from rotation to rotation, we can determine an average correlation period of the rotation of the large-scale magnetic field. We derived the rotational correlation period from the departure of the maximum of the cross-correlation function from zero lag, which corresponds to the Carrington rotational period of the Sun (Fig. 1d). The rotational period of the large-scale magnetic field exhibits complex time variations and quantifies the contributions from various rotational modes that vary in the course of the activity cycle.

On the whole, the pattern of the rotation varies both during a given 11-year cycle and from cycle to cycle.

As demonstrated in [10], during the last three cycles, modes with periods of 27.8–28.5 days were more pronounced in phases of activity growth and activity maximum, whereas the 27-day rotational mode dominated in the declining-activity phase. To identify the contributions from the positive and negative polarities in the different hemispheres to the spatial and temporal MSMF variations from rotation to rotation, we analyzed longitude–time diagrams for the BSMF separately.

4. LONGITUDE–TIME DIAGRAMS FOR THE BSMF AND MSMF

The quantitative characteristics of the distribution of the dominant BSMF polarities and their longitude–time diagrams were obtained from H_α and magnetic-field synoptic maps using the robust technique described in [20]. Figure 1e shows a doubled longitude–time diagram of the BSMF distribution in the equatorial zone (40° N–40° S). Positive polarity is represented by areas lighter than the gray background, and negative polarity by darker areas. The BSMF, as well as the MSMF, is comprised primarily of rigid-rotation modes with periods of about 27 and 28 days. Most MSMF features can be identified with similar features in the BSMF diagrams. The BSMF diagrams are noisier and more fragmentary, possibly due to the fact that the BSMF pattern reflects the longitudinal distribution of individual structures of the large-scale magnetic field, while the MSMF pattern represents the integrated flux from the entire solar disk.

The behavior of the BSMF can be substantially different in the northern and southern hemispheres. In our case, the coefficient of the two-dimensional correlation between the BSMF distributions in the two hemispheres is statistically insignificant for the entire sample. Thus, it is reasonable to compare the MSMF and BSMF diagrams for each hemisphere. Figure 2 displays the BSMF diagrams for the two hemispheres as compared with the MSMF. The numbers 1, 2, 3, etc. label the most prominent MSMF structures and the corresponding BSMF structures in both hemispheres. If a number is repeated in all the three diagrams, this means that the labeled feature is observed in both hemispheres and in the MSMF. We can see that most MSMF structures have counterparts in one or both hemispheres. An asymmetry is evident: the MSMF pattern is more similar to the BSMF distribution in the southern hemisphere, especially in large features. The MSMF diagram is in best agreement with the diagram for the BSMF distribution summed over the two hemispheres (Fig. 1e). Thus, the MSMF signal reflects the contributions from the BSMF of both hemispheres, with some dominance of the southern hemisphere during the time interval considered.

To obtain the mean characteristics of the solar rotation for the entire observation period, we applied a Radon transformation (widely used in tomography [21]) to the longitude–time MSMF diagram. The Radon transfor-
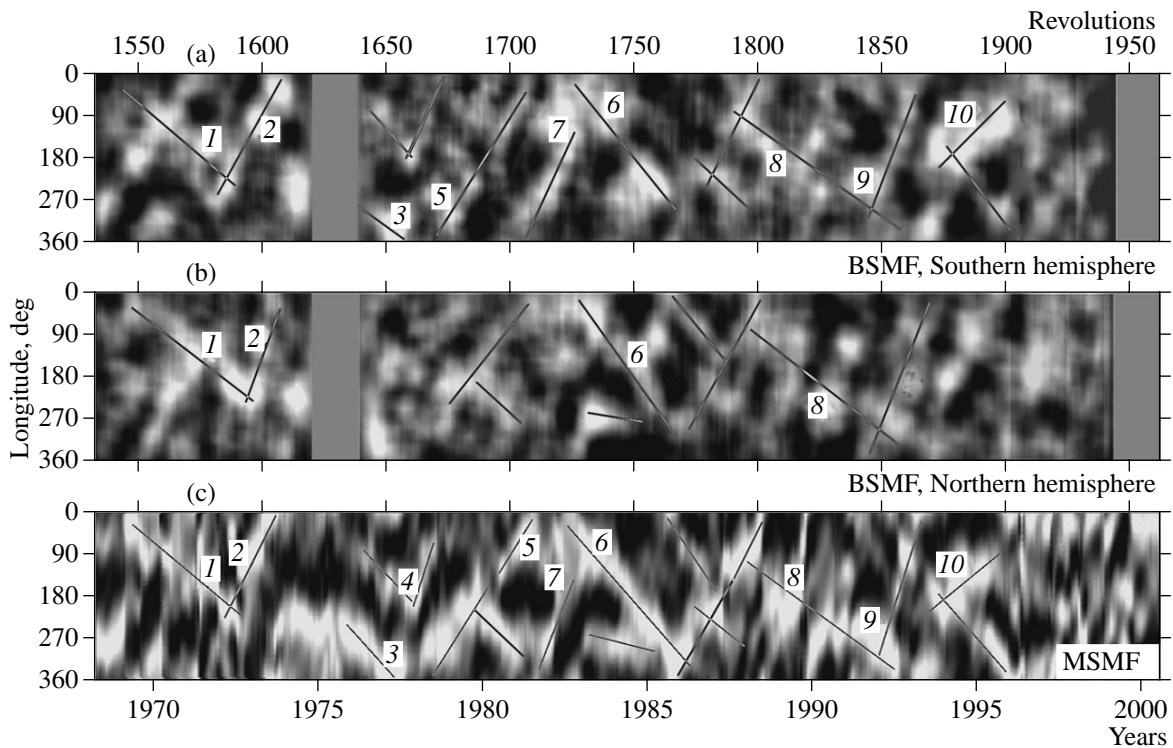
Fig. 2. Normalized BSMF values in the (a) southern and (b) northern hemispheres; (c) MSMF.

mation of a two-dimensional function $f(x, y)$ or image consists of integrating this function along the direction specified by an angle θ and determining the projection of this image onto the axis of the new spatial coordinate x' normal to the integration direction:

$$R_\theta(x') = \int_{-\infty}^{\infty} f(x'\cos\theta - y'\sin\theta, x'\sin\theta + y'\cos\theta)dy',$$
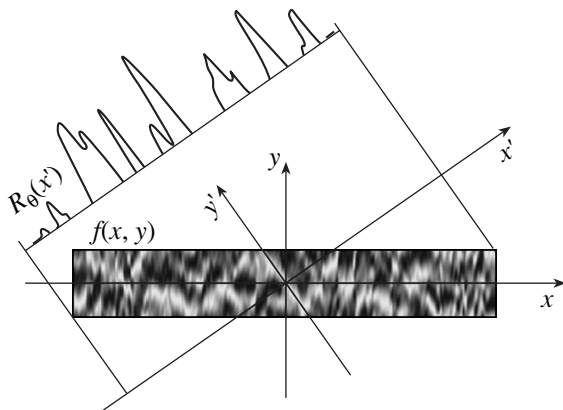

Fig. 3. The Radon transformation geometry and variables.

where the relationship between the two coordinate systems is

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The disposition of the original diagram $f(x, y)$ and its projection $R_\theta(x')$ in the different coordinate systems is illustrated in Fig. 3.

Integration of the function along various directions yields projections describing the internal structure of the object under study as observed from various sides. If internal patterns of the object are characterized by increased values of $f(x, y)$ (density, intensity) and if these are aligned in some direction, then, when the function is integrated over this direction, the Radon transform yields a local maximum at the corresponding projection after the integration over the same direction. We will make use of this property by applying the Radon transform to detect rigid-rotation modes and estimate their periods.

The Radon transform is usually represented by a two-dimensional diagram whose axes correspond to the angle θ and distance x' from the center of the projection. Figure 4 shows the Radon transform obtained for the squared normalized values of the longitude–time MSMF diagram (the sign here is not important). The direction of the main maximum $\theta = 90°$ corresponds to the Carrington rotation; in this case, the integration is

performed in the horizontal direction. Near $\theta = 90°$, we can see a core, whose complex structure is represented by contours. Obviously, these modes are not smeared by averaging over the entire diagram along various directions, but instead retain their discrete structure. This pattern contradicts the simplified notion that the transition from slow to fast rotation modes in the course of the 11-year cycle results from the drift of spot-formation zones toward the equator. Otherwise, the Radon transform would not exhibit such a distinct structure in the wings.

5. LONG-TERM MSMF VARIATION

The MSMF signal is a differential in the sense that it represents the difference between the contributions from the positive and negative polarities in the distribution of the line-of-sight magnetic-field component over the solar disk. The sign of the MSMF is determined by the contribution of the dominant polarity. To reveal long-term trends in the magnetic-field evolution, we computed the cumulative sum, equal, for a given time, to the sum of all preceding MSMF values. The graph of the cumulative sum CumSum rescaled according to the formula $\text{CumSum}(t_i) = (\sum_{k=1}^{i} f(t_k) + 9000)/25$ is shown in Fig. 1b by the smooth curve. MSMF fluctuations compensate each other in the summation, while the weakly pronounced asymmetry of the polarities in the original signal is additive, and exhibits fairly smooth, cyclic variations. The cumulative sum reflects long-term trends in the variations: domination of positive MSMF values increases the sum, while domination of negative values decreases it. This property of the integral parameter is well suited to describe the asymmetry in the behavior of the two polarities in the MSMF signal. The dominant magnetic field or a magnetic "monopole" exhibit slow variations during the solar-activity cycle [22].

For cycle 21, the cumulative sum can be described by a two-peaked curve between the epochs of minima, 1976 and 1986. The magnetic-field reversals in the northern and southern hemispheres are marked by upward and downward arrows, respectively. The reversals encompass the interval 1980.0–1981.8 [23] and occur between the two peaks of the cumulative sum. During the phase of growing activity, the mode with a period of 27.8 days dominates, and the cumulative sum increases. In the phase of declining activity, the 27-day mode dominates, the prevailing polarity is negative, and the cumulative sum decreases. At the very end of cycle 21 and the beginning of cycle 22, the 27.9-day mode and the positive polarity dominate. Starting in mid-1988, the negative polarity dominates in cycle 22, and we observe a different character in the time variation. In the growth phase and at the epoch of maximum activity, the cumulative sum decreases rapidly, and small-amplitude oscillations of the sum set in after the reversals in 1990.8–1991.8. Only since 1998 can we see
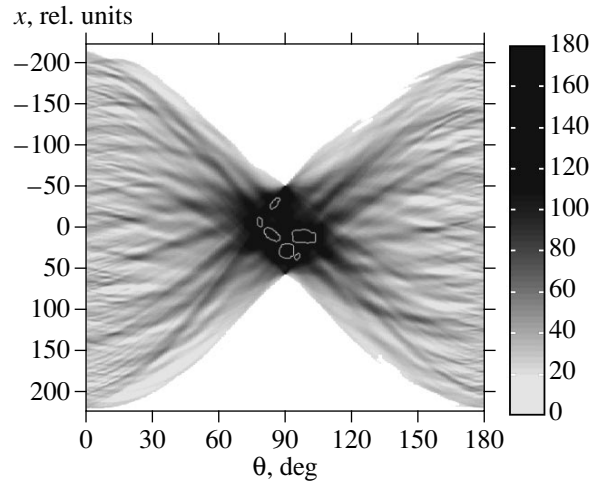
Fig. 4. The Radon transform of the squared normalized longitude–time distribution of the MSMF variations during 1968–2000.

evidence for increases in the cumulative sum, reflecting the beginning of a new activity cycle. During cycle 20, partially covered by measurements at the Crimean Astrophysical Observatory, we observe a rapid decrease in the cumulative sum, which changes to a rapid increase after the reversal period, 1969.7–1971.5.

6. CONCLUSION

Our spectral–time and longitude–time analyses of variations in the MSMF and BSMF clearly demonstrate the presence of coherent structures in the dynamics of the large-scale solar magnetic field, which appear as rigid-rotation modes. The rotational regime changes both in the course of an 11-year activity cycle and from cycle to cycle. Various rigid-rotation modes coexist and replace one another during the cycle. At the phase of activity growth, rotational modes with periods of 27.8–28.5 days dominate, while the 27-day rotational mode is predominant during the declining-activity phase. Rotational modes with periods of 29–30 days appear occasionally. On the whole, the solar rotational regime seems to be determined by the dynamics and nonlinear interactions of large-scale modes, which can be present even in a highly supercritical turbulent shear flow [24].

Deconvolution of the MSMF and BSMF time series enabled construction of a detailed picture of the solar rotation for activity cycles 20–23. Changes in the rotational regime are well correlated with changes in the apparent asymmetry of the MSMF polarities. The integrated characteristic of the MSMF exhibits long-term and cyclic variations determined by reversals of the global magnetic field. In two of three cases, a general decline in the cumulative MSMF sum changes to a rise after polarity reversals, and vice versa. A 22-year cyclic variation can be traced in the behavior of the cumulative MSMF sum. It is interesting that variations in the

cumulative MSMF sum during cycle 21 are balanced in the sense that, by the end of the cycle (1986), the sum returns to the value characterizing the beginning of the cycle (1976). In cycle 22, the character of time variations is different: the decrease in the cumulative sum is replaced by oscillations, and the variation of the integrated MSMF remains unbalanced over the cycle. It would be premature to draw final conclusions about cyclic regularities in MSMF variations; however, differences between even and odd cycles may be present in the behavior of the magnetic-field asymmetry, resulting from the physical coupling between even and odd cycles [25] or due to the global relict magnetic field of the Sun [26].

ACKNOWLEDGMENTS

REFERENCES

1. P. H. Scherrer, J. M. Wilcox, V. A. Kotov, *et al.*, Sol. Phys. **52**, 3 (1977).

2. V. A. Kotov and A. B. Severnyĭ, in *Total Magnetic Field of the Sun as a Star. Catalogue 1968–1976*, Ed. by É. E. Dubov [in Russian] (Mezhduvedomstv. Geofiz. Komitet Akad. Nauk SSSR, Moscow, 1983).

3. N. R. Sheeley, Jr., C. R. De Vore, and L. R. Shampine, Sol. Phys. **106**, 251 (1986).

4. D. I. Ponyavin, in *New Cycle of Solar Activity: Observational and Theoretical Aspects*, Ed. by V. I. Makarov, V. N. Obridko, *et al.* [in Russian] (Glavn. Astron. Obs. Ross. Akad. Nauk, St. Petersburg, 1998), p. 153.

5. V. A. Kotov, V. I. Khaneĭchuk, and T. T. Tsap, Astron. Zh. **76**, 218 (1999) [Astron. Rep. **43**, 185 (1999)].

6. V. N. Obridko and Yu. R. Rivin, Astron. Zh. **73**, 812 (1996) [Astron. Rep. **40**, 740 (1996)].

7. V. M. Grigoryev and M. L. Demidov, Sol. Phys. **114**, 147 (1987).

8. N. M. Astaf'eva, Usp. Fiz. Nauk **166**, 1145 (1996) [Phys. Usp. **39**, 1085 (1996)].

9. J. O. Stenflo, Astron. Astrophys. **210**, 403 (1989).

10. V. N. Obridko and B. D. Shelting, Astron. Zh. **77**, 124 (2000) [Astron. Rep. **44**, 103 (2000)].

11. V. N. Obridko and B. D. Shelting, Astron. Zh. **77**, 303 (2000) [Astron. Rep. **44**, 262 (2000)].

12. R. Howard and B. J. LaBonte, Astrophys. J. Lett. **239**, L33 (1980).

13. V. A. Kotov, N. N. Stepanyan, and Z. A. Shcherbakova, Izv. Krym. Astrofiz. Obs. **56**, 75 (1977).

14. V. I. Khaneĭchuk, Astron. Zh. **76**, 385 (1999) [Astron. Rep. **43**, 330 (1999)].

15. A. V. Mordvinov, Issled. Geomagn., Aéron. Fiz. Solntsa **99**, 132 (1992).

16. A. V. Mordvinov, Astron. Astrophys. **293**, 572 (1995).

17. D. V. Erofeev, Sol. Phys. **167**, 25 (1996).

18. R. C. Willson and A. V. Mordvinov, Geophys. Res. Lett. **26**, 3613 (1999).

19. A. V. Mordvinov and L. A. Plyusnina, Sol. Phys. **197**, 1 (2000).

20. L. A. Plyusnina, in *Large-Scale Structure of the Solar Activity: Advances and Perspectives*, Ed. by V. I. Makarov, V. N. Obridko, *et al.* [in Russian] (Glavn. Astron. Obs. Ross. Akad. Nauk, St. Petersburg, 1999), p. 237.

21. C. Chapman, in *Seismic Tomography. With Application in Global Seismology and Exploration Geophysics*, Ed. by G. Nolet (Reidel, Dordrecht, 1987; Mir, Moscow, 1990).

22. V. M. Grigoryev and M. L. Demidov, in *Solar Magnetic Fields and the Corona*, Ed. by R. B. Teplitskaya (Nauka, Novosibirsk, 1989), p. 108.

23. V. I. Makarov and V. V. Makarova, Sol. Phys. **163**, 267 (1996).

24. M. I. Rabinovich and M. M. Sushchik, Usp. Fiz. Nauk **160** (1), 3 (1990) [Sov. Phys. Usp. **33**, 1 (1990)].

25. V. V. Pipin, Astron. Astrophys. **293**, 572 (1999).

26. E. E. Benevolenskaya, in *New Cycle of Solar Activity: Observational and Theoretical Aspects*, Ed. by V. I. Makarov, V. N. Obridko, *et al.* [in Russian] (Glavn. Astron. Obs. Ross. Akad. Nauk, St. Petersburg, 1998), p. 19.

Translated by A. Getling

Anisotropy of Small-Scale Inhomogeneities in the Region of Solar-Wind Acceleration

I. V. Chasheĭ

Lebedev Physical Institute, Leninskiĭ pr. 53, Moscow, 117924 Russia
Received November 14, 2000

Abstract—A model is proposed to explain observational data on the scattering of radio signals, which indicate that small-scale plasma-density inhomogeneities in the region of solar-wind acceleration are strongly elongated in the radial direction, with the degree of elongation sharply decreasing at heliocentric distances of about six solar radii. The evolution of the energy spectra of the fluctuations of the magnetic field and plasma density is studied assuming that the plasma-density fluctuations are generated locally by nonlinear interactions of high-frequency Alfven waves, and that the gradients of the mean plasma parameters are smooth. The growth rates of the main nonlinear processes are estimated. The strong elongation of the inhomogeneities first arises when the Alfven waves travel through the chromosphere–corona transition layer, then survives to considerable distances from the Sun because the associated nonlinear relaxation processes are fairly slow. Estimates of the degree of elongation of the inhomogeneities and the characteristic distance for changes in the angular wave spectra are in good agreement with radio propagation data. © *2001 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The first data on the anisotropy of small-scale inhomogeneities in the inner regions of the solar wind were derived from observations of meter-wavelength radio signals scattered at heliocentric distances $r \geq 10R_S$ (R_S is the solar radius) [1]. These data show that density fluctuations with scales of about 1 km are elongated in the radial direction. However, the anisotropy is comparatively small, and the degree of elongation ξ (the axial ratio) does not exceed 1.5–2. The scattering of centimeter-wavelength signals observed by the Very Large Array reveals anisotropy for inhomogeneities with scales 10–30 km as well [2, 3]. The degree of elongation is fairly high ($\xi \geq 10$) at small heliocentric distances $r \leq 6R_S$, and decreases to $\xi \leq 2$ at greater distances ($6R_S < r < 10R_S$). The strong elongation of small-scale inhomogeneities in the outer corona and solar-wind acceleration region was also reported in [4, 5]. According to [3], even large plasma-density inhomogeneities (with scales of 1000 km) are anisotropic in the solar-wind acceleration region, though their degree of elongation is somewhat lower. Inhomogeneities of the magnetic field in the acceleration region have been studied using data on fluctuations in the Faraday rotation of radio signals emitted by the *HELIOS* spacecraft [6]. The large-scale ($10^4–10^5$ km) magnetic-field inhomogeneities are also strongly anisotropic, with their degree of elongation being about ten for heliocentric distances smaller than $10R_S$ [6]. The typical axial ratios for density inhomogeneities with scales of 10–100 km estimated from interplanetary scintillation data [7–10] do not exceed 1.5–2 in the fully formed solar wind ($r > 20R_S$). There is some decrease in ξ for $r < 40R_S$, and the degree

of anisotropy does not exhibit any appreciable dependence on heliocentric distance at greater distances [9, 10]. These experimental data suggest that inhomogeneities of the density and magnetic field are strongly anisotropic in the region of solar wind acceleration, $r < 10R_S$. For $r > 10R_S$, the anisotropy is considerably lower, at least for small-scale inhomogeneities.

The ratios ξ estimated from the radio-propagation data (scattering, scintillations, and Faraday fluctuations) correspond to a two-dimensional cross section of the density-inhomogeneity elongation ellipsoid made by the intersection with the plane of the sky along the line of sight. The spatial turbulence spectrum describing the wave-number distribution of the inhomogeneities is also anisotropic. The main physical reason for the anisotropy is the radial expansion of the solar-wind plasma and/or strong magnetic fields that are close to radial in the outer solar corona. However, the physical mechanism for the formation of the anisotropic turbulence spectra and their radial evolution remains open to question.

Here, we examine a model for the small-scale turbulence that can explain the anisotropy of the density fluctuations observed in the solar-wind acceleration region at heliocentric distances of $2–10R_S$. The main assumption of the model is that the density fluctuations are due to magneto-acoustic waves generated locally by nonlinear interactions of Alfven waves carried by the plasma flow [11]. The nonlinear excitation of magnetic sound waves during decay processes (decay instability of Alfven waves) was first studied in [12]. The appreciable role of such processes for the solar wind was noted in [11, 13–16]. The numerical simulations of [17–19] show that Alfven waves do excite density fluctuations in the outer corona, due to both the free propa-

gation of finite wave trains from the Sun and the presence of a driving source in the radiation region.

Let us suppose that the initial source of Alfven waves is at the base of the transition layer separating the chromosphere and corona, with the Alfven turbulence being strong and close to isotropic at the initial distance $r = r_* \approx R_S$ [20]. We assume that the plasma parameter $\beta_* \approx 1$ at the initial distance, but the magnetic field is strong ($\beta \ll 1$) in the corona and solar-wind acceleration region $r > r_*$. Traveling outward from the starting level, the Alfven waves pass through the transition layer and lower corona. During this passage, the wave energy flux and level of turbulence are partially attenuated by linear absorption in the plasma [21]. Of course, we consider those regions in the corona that are sources of the solar wind. In the case of the fast solar wind, these regions are associated with open magnetic configurations such as coronal holes. In the case of the slow solar wind, whose origin remains unknown, these regions are either non-stationary open magnetic configurations at low latitudes or the peripheral regions of open structures, whose escaping plasma fills the space above stationary closed structures. Below, we study the nonlinear evolution of the anisotropy of the turbulence spectra in the solar-wind acceleration region, assuming that linear damping of Alfven waves beyond the corona is negligible and that the initial anisotropy of the angular distribution of the Alfven waves forms in the transition layer and lower corona, where nonlinear processes can be neglected.

2. SPECTRAL ANISOTROPY OF ALFVEN WAVES ESCAPING FROM THE CORONA

Let us consider the formation of the angular spectrum of Alfven waves in the transition layer and inner corona, assuming that wave perturbations are generated at the base of the chromosphere–corona transition layer. The strong-turbulence regime in the upper chromosphere together with the condition $\beta \approx 1$ means that we can assume that the angular spectra of Alfven and magneto-acoustic waves at their place of generation are close to isotropic. On average, the wave vectors $\mathbf{K}$ (k, q) of these three types of MHD waves (Alfven, rapid, and slow magneto-acoustic) have approximately equal longitudinal (k) and transverse (q) components with respect to the regular magnetic field $\mathbf{B}$. The plasma temperature increases sharply in the transition layer, with the pressure remaining approximately constant. Consequently, the speed of sound v_S and Alfven velocity v_A undergo abrupt changes. Since rapid magneto-acoustic waves are strongly reflected in the transition layer, the perturbations escaping to the corona are mainly Alfvenic (a) and slow magneto-acoustic (s). Their dispersion relations for $\beta \ll 1$ ($v_A^2 \gg v_S^2$ in the corona and acceleration region) are

$$\omega^a = k v_A, \quad \omega^s = k v_S, \tag{1}$$

where $\omega^{a,s}$ are their frequencies. If we assume that the coronal plasma is quasi-stationary and quasi-homogeneous transverse to the regular magnetic field (with respect to the waves considered), the conditions $\omega^{a,s} =$ const and $\mathbf{q} =$ const are fulfilled in the linear approximation. Upon leaving the transition layer, the initially isotropic angular spectrum becomes strongly anisotropic due to abrupt changes in the wave phase velocities. Since the s waves suffer strong linear absorption in the corona [20], we shall further concentrate primarily on estimates for the spectrum of the Alfven waves. On average, the magnetic field in the corona decays with distance from the Sun much more slowly than the plasma density. Therefore, the Alfven velocity reaches its maximum $v_{Ac} \approx 10^8$ cm/s (number density $N \sim 10^7$–10^6 cm^{-3}, magnetic field $B \leq 1$ G) at heliocentric distances $r \approx r_c$, where $r_c \approx 1.5$–$2R_S$, and then slowly decreases with distance from the Sun. Further, we will take $r = r_c$ to be the outer boundary of the corona. The equations $qr =$ const [22] and $\omega^a =$ const yield the ratio of the wave vector components of the Alfven waves (axial ratio) for $r = r_c$

$$\xi_c^a = \frac{\langle q_c \rangle}{\langle k_c \rangle} = \frac{\langle q_* \rangle}{\langle k_* \rangle} \frac{v_{Ac} r_*}{v_{A*} r_c} \approx \frac{v_{Ac} r_*}{v_{A*} r_c}. \tag{2}$$

Relation (2) shows that the angular distribution of the Alfven waves escaping from the corona is strongly anisotropic, with the wave vectors being almost transverse to the magnetic field; that is, the inhomogeneities are strongly elongated in the radial direction. Namely, for $v_{A*} \approx 10^6$ cm/s, (2) yields $\xi_c^a \approx 50$.

We should also estimate the relative intensity of high-frequency Alfven waves escaping from the corona at $r \approx r_c$ (frequencies from $\omega_* \sim 2 \times 10^{-2}$ s^{-1}, corresponding to periods of about 5 min [20], to $\omega_1 \sim 10^{-1}$ s^{-1}, corresponding to wavelengths of 10^7 cm at the initial distance). Since the solar-wind region considered is assumed to be sub-Alfvenic ($v \ll v_A$), the spectral density W_ω of waves with frequency ω becomes

$$W_{c\omega}^a = W_{*\omega}^a \left(\frac{N_c}{N_*} \right)^{1/2} \exp[-\tau_\omega^a], \tag{3}$$

where N is the plasma number density, τ_ω^a is the optical depth of the corona to the linear absorption of Alfven waves with frequency ω, and the subscripts $*$ and c correspond to the initial distance $r = r_*$ and outer boundary of the corona $r = r_c$, respectively. Estimates for τ_ω^a are presented in [21]: for coronal magnetic fields of the order of several Gauss, $\tau_{\omega*}^a \approx 0.5$ and $\tau_\omega^a \approx 3$–4. Assuming in (3) that $N_* \approx 10^{11}$ cm^{-3} ($\beta_* \approx 1$ for $B_* \approx 2$ G

and $T_* \approx 10^4$ K), $N_c \approx 10^7$ cm^{-3}, $B_c \approx 0.5$ G, $W_{*\omega} \sim \omega^{-3/2}$, and $4\pi\omega_* W_{*\omega*}/B_*^2 \approx 1$, we obtain

$$\frac{4\pi\omega_* W_{c\omega*}}{B_c^2} \sim 5 \times 10^{-2}, \quad \frac{4\pi\omega_1 W_{c\omega 1}}{B_c^2} \sim 3 \times 10^{-4}. \quad (4)$$

3. KINETIC EQUATIONS FOR MHD WAVES BEYOND THE CORONA

Let us study the evolution of the Alfven turbulence beyond the corona, assuming, as noted above, that the plasma flow is sub-Alfvenic and $v_A^2 \gg v_S^2$ ($\beta \ll 1$) and taking into account the smooth inhomogeneity of the medium in addition to nonlinear processes. The spectrum of Alfven-wave turbulence with parameters (2) and (4) becomes a boundary condition for the problem. For a plasma with $v_A^2 \gg v_S^2$, the most efficient nonlinear process is the decay $a_1 = a_2 + s$, where the Alfven wave breaks down into Alfven and magneto-acoustic waves [11, 23]. During this process, the Alfven wave traveling from the Sun generates a slow magneto-acoustic wave, which is responsible for the density fluctuations, and an opposing Alfven wave. The spatial turbulence spectra are described by the kinetic equations [11, 23]

$$\frac{dW_1}{dt} = 4\pi\omega_1 \int d\mathbf{K}_2 d\mathbf{K}_3 \{ u_{123}(\omega_1 W_2 W_3 - \omega_2 W_1 W_3 \\ - \omega_3 W_1 W_2) + u_{213}(\omega_3 W_1 W_2 + \omega_1 W_2 W_3 - \omega_2 W_1 W_3) \}, \quad (5)$$

$$\frac{dW_3}{dt} = 8\pi\omega_3^2 \int d\mathbf{K}_1 d\mathbf{K}_2 u_{123}\{W_1 W_2\} - \gamma_L W_3, \quad (6)$$

where

$$W_{1,2} = W^a(\mathbf{K}_{1,2}), \quad W_3 = W^s(\mathbf{K}_3) \quad (7)$$

are the spatial energy spectra of the turbulence and the probability u_{123} of the decay process $a_1 = a_2 + s$ is determined by [23]

$$u_{123} = U\delta(\mathbf{K}_1 - \mathbf{K}_2 - \mathbf{K}_3)\delta(\omega_1 - \omega_2 - \omega_3),$$

$$U = \frac{\cos^2\varphi_{12}}{32\rho v_S^2}, \quad (8)$$

where φ_{12} is the angle between the transverse components of the wave vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ and ρ is the plasma density.

The equation for W_2 is similar to (5) except for an interchange of the subscripts 1 and 2. Equation (6) takes into account linear Landau damping of the slow magneto-acoustic wave by ions γ_L. This damping is fairly strong in the solar wind and outer corona, $\gamma_L = \nu\omega_3$, with the numerical factor $\nu \leq 1$ [11]. Equations (5)–(8) can be used if we neglect the smooth plasma

inhomogeneity compared to the characteristic scales for a nonlinear process, which is valid only for wave frequencies exceeding a certain critical frequency, which we determine below. We can neglect the nonlinearity of waves with lower frequencies. Their wave vectors and energy densities are described by expressions similar to (2) and (3) with $\tau_\omega^a = 0$ (very weak linear damping of Alfven waves in the supercorona).

The conservation of energy and momentum in the individual decay process gives the relations

$$\mathbf{K}_{1,2} - \mathbf{K}_{2,1} - \mathbf{K}_3 = 0, \quad \omega_{1,2} - \omega_{2,1} - \omega_3 = 0. \quad (9)$$

Using the dispersion relations (1), we can find exact solutions of (9) for the parallel components of the wave vectors

$$a_1 = a_2 + s: k_2 = -k_1(1-\alpha)/(1+\alpha),$$
$$k_3 = 2k_1/(1+\alpha),$$
$$a_2 = a_1 + s: k_2 = -k_1(1+\alpha)/(1-\alpha),$$
$$k_3 = -2k_1/(1-\alpha), \quad (10)$$

where $\alpha = v_S/v_A < 1$, and we assume that $k_1 > 0$. Expressions (10) show that s waves and related density fluctuations can be excited only via an interaction of opposing Alfven waves. The relation between the transverse components of the wave vectors

$$\mathbf{q}_{1,2} = \mathbf{q}_{2,1} + \mathbf{q}_3 \quad (11)$$

is not as restrictive as (10), and, for fixed $\mathbf{q}_{1,2}$, (11) permits various combinations of $\mathbf{q}_{2,1}$ and $\mathbf{q}_3$.

4. TURBULENCE REGIMES AND NONLINEAR GROWTH RATES

To solve (5) and (6), we assume that the turbulence spectra take the form

$$W^{a,s}(\mathbf{K}) = \frac{1}{v_{a,s}} W^{a,s}(\omega) f^{a,s}(\mathbf{q}),$$
$$\int f^{a,s}(\mathbf{q}) d\mathbf{q} = 1, \quad (12)$$

where $W^{a,s}(\omega)$ are the one-dimensional frequency spectra. The quasi-stationary turbulence spectra that are the solutions of (5) and (6) with zero right-hand side are formed by the nonlinear relaxation of initial spectra with the parameters (2) and (4). In our model, we can distinguish three main processes involved in relaxation to a quasi-stationary regime: (a) formation of the quasi-stationary frequency spectra $W^{a,s}(\omega)$, (b) formation of the quasi-stationary ratio W_1/W_2, and (c) angular redistribution of the initially anisotropic spectra so that they become isotropic. For each process, the relaxation to a quasi-stationary regime begins at the highest frequencies (of the order of the ion cyclotron frequency). As the plasma moves away from the Sun, the quasi-stationary spectral region is extended to lower frequencies. Let us

determine the nonlinear growth rates for processes (a)–(c) and characteristic forms of the quasi-stationary solutions $W^{a,s}(\mathbf{K})$.

When considering the formation of one-dimensional frequency spectra $W^{a,s}(\omega)$, we can neglect terms proportional to W_3 on the right-hand side of (5). We then find from (5) and (6) that the quasi-stationary spectra $W^{a,s}(\omega)$ take the identical forms

$$W^{a,s}(\omega) = A_{123}/\omega, \qquad (13)$$

independent of $f^{a,s}(\mathbf{q})$. The structure constants in (13) express the total energy of the turbulence (except the logarithmic factor of order unity) for the frequency interval of the quasi-stationary regime, with $A_1 \geq A_2 \gg A_3$. Using (5), we can estimate the rate γ_{fr} of nonlinear formation of the spectra (13), and find that it is proportional to the frequency ω_1:

$$\gamma_{fr} \approx \omega_1 \frac{A_1}{\rho v_A^2}. \qquad (14)$$

For the formation of the ratio W_1/W_2, we first neglect in (5) terms containing W_2 and retain those containing W_1 and W_3. Then, (5) indicates that the state with $W_2 < W_1$ becomes unstable. In fact, this describes the decay instability studied in [12] for an ensemble of Alfven waves with random phases. For the quasi-stationary spectra, the energy levels of the opposing Alfven waves are equally distributed: $W_1 = W_2$. We can similarly (14) estimate the nonlinear growth rate γ_{12} for relaxation to the equilibrium $W_1 = W_2$:

$$\gamma_{12} \approx \omega_1 \left(\frac{v_A}{v_S}\right)^3 \left(\frac{A_1}{\rho v_A^2}\right)^2. \qquad (15)$$

If we assume $W_1 = W_2$, (6) yields the following estimate for the quasi-stationary level of the slow magneto-acoustic waves:

$$A_3 \approx \frac{A_1^2}{\rho v_S v_A}. \qquad (16)$$

The quasi-stationary transverse wave-vector distributions $f^{a,s}(\mathbf{q})$ are determined by the integral equations

$$f^s(\mathbf{q}_3) = \int f^a(\mathbf{q}_1) f^a(\mathbf{q}_1 - \mathbf{q}_3) \cos^2 \varphi_{12} d\mathbf{q}_1, \qquad (17)$$

$$\int d\mathbf{q}_2 \cos^2 \varphi_{12} \left\{ (1 + \alpha) f^s(\mathbf{q}_1 - \mathbf{q}_2) \right.$$

$$\times \left[\frac{1+\alpha}{1-\alpha} f^a(\mathbf{q}_2) - f^a(\mathbf{q}_1) \right] + (1 - \alpha) f^s(\mathbf{q}_2 - \mathbf{q}_1) \qquad (18)$$

$$\left. \times \left[\frac{1-\alpha}{1+\alpha} f^a(\mathbf{q}_2) - f^a(\mathbf{q}_1) \right] \right\} = 0,$$

where $\alpha = v_S/v_A < 1$ as previously. We have derived Eqs. (17) and (18) from (5) and (6) with $W_1 = W_2$, taking into account (1), (10), (11), and (13). The rate of nonlinear relaxation to the spectra $f^{a,s}(\mathbf{q})$ (17), (18) can be found using the original Eq. (5):

$$\gamma_{an} \approx \omega_1 \frac{v_A}{v_S} \left(\frac{A_1}{\rho v_A^2}\right)^2. \qquad (19)$$

The relaxation to a quasi-stationary regime is described by (5), with its right-hand side proportional to the integral (18), which corresponds to a diffusion process. Simultaneous with the linear evolution of the transverse wave vectors, similar to (2), this process will decrease the initial anisotropy of the angular distribution of the Alfven waves.

A comparison of the nonlinear rates γ_{fr} (14), γ_{12} (15), and γ_{an} (19) shows that, for $A_1/\rho v_A^2 \ll 1$, the nonlinear evolution of the angular distribution of the Alfven waves is the slowest process among the three considered:

$$\gamma_{an} \ll \gamma_{12} \ll \gamma_{fr}. \qquad (20)$$

We can estimate the characteristic heliocentric distances for relaxation to the equilibrium distributions (13), (16), (17), and (18):

$$R_{fr}(\omega_1) \approx \frac{v}{\omega_1} \frac{\rho v_A^2}{A_1}, \qquad (21)$$

$$R_{12}(\omega_1) \approx \frac{v}{\omega_1} \left(\frac{v_S}{v_A}\right)^3 \left(\frac{\rho v_A^2}{A_1}\right)^2, \qquad (22)$$

$$R_{an}(\omega_1) \approx \frac{v}{\omega_1} \frac{v_S}{v_A} \left(\frac{\rho v_A^2}{A_1}\right)^2, \qquad (23)$$

equating the rates γ_{fr} (14), γ_{12} (15), and γ_{an} (19) with fixed wave frequency ω_1 to the reciprocal time for the linear evolution v/r, where v is the solar-wind velocity. By virtue of (20), a regime with the energy spectra (13) and turbulence levels $A_1 = A_2$, A_3 (16) $\ll A_{1,2}$ (16), but with the angular distribution of the Alfven waves determined mainly by linear processes, will be established for heliocentric distances $R_{fr} \leq R_{12} < r < R_{an}$ and frequencies $\omega > \omega_1$. In accordance with the decay conditions (10) and (11), the axial ratio ξ_s for the slow magneto-acoustic waves (and density fluctuations) will be approximately half the ratio ξ_a for the Alfven waves:

$$\xi_a \approx 2\xi_s. \qquad (24)$$

5. SCATTERING OF RADIO WAVES AND THE LEVEL OF TURBULENT FLUCTUATIONS

Let us determine the relation between the parameters of the scattered radio signals measured in [2, 3] and fluctuations of the density and magnetic field. For our

estimates, we assume that the spatial spectrum of the density fluctuations takes the form

$$\delta N^2(\mathbf{K}) = C(\xi^2 k^2 + q^2)^{-3/2}, \quad C = \frac{\xi \delta N_0^2}{2\pi \ln(\omega_m/\omega_{10})}, \quad (25)$$

where ξ is the axial ratio for the density fluctuations; ω_{10} is the Alfven-wave frequency determined by relation (21) for a given heliocentric distance; ω_m is the maximum Alfven-wave frequency, close to the ion cyclotron frequency; and δN_0^2 is the rms of the small-scale density fluctuations. For

$$\frac{C}{N^2} = \frac{A_3}{\rho v_S^2}, \quad W_3(k) = (k\xi)^{-1},$$

$$f_3(\mathbf{q}) = (2\pi)^{-1} k\xi (k^2 \xi^2 + q^2)^{-3/2} \quad (26)$$

expression (25) corresponds to the quasi-stationary energy spectrum (12), (13). We shall assume that the radial evolution of the spectrum (25) is governed by the r dependence of the axial ratio ξ and the structure constant C, with the function $\xi(r)$ related to the regular plasma inhomogeneity and determined by linear effects. The radio-interferometric observations [2, 3] detected the structure function $D_S(\mathbf{b})$ of the phase fluctuations of the radio signals studied ($\mathbf{b}$ is the baseline vector between the radio telescopes). This function is related to the spectrum (25) by the expression

$$D_S(\mathbf{b}) = 4\pi(\lambda r_e)^2$$
$$\times \int dx \int \delta N^2(\mathbf{K})[1 - \cos(\mathbf{K}\mathbf{b})]\delta(q_x)d\mathbf{K}, \quad (27)$$

where λ is the wavelength and $r_e = 2.8 \times 10^{-13}$ cm the classical electron radius. We assume that the line of sight is directed along the OX axis and is perpendicular to the regular magnetic field. Assuming that the scattering takes place in some layer of thickness ΔX and integrating in (27), we obtain the estimates

$$D_S(b_\parallel) = \frac{4\pi(\lambda r_e)^2 \delta N_0^2 \Delta X b_\parallel}{\xi \ln(\omega_m/\omega_{10})}, \quad (28)$$

$$D_S(b_\perp) = \xi D_S(b_\parallel), \quad (29)$$

where $b_\parallel$ and $b_\perp$ are the projections of the vector $\mathbf{b}$ parallel and perpendicular to the radial direction, respectively. Using (28) and (29), we can estimate the structure constant C (i.e., the variance of the density fluctuations δN_0^2) from the measured quantities $D_S(b_\parallel)$ and $D_S(b_\perp)$. This structure constant enables us to determine the relative level of high-frequency Alfven waves, which is, in accordance with (16) and (26),

$$\frac{A_1}{\rho v_A} = \frac{C^{1/2}}{N}\left(\frac{v_S}{v_A}\right)^{3/2}. \quad (30)$$

Using (30), we can estimate the nonlinear growth rates (14), (15), and (19) and corresponding heliocentric distances (21)–(23). In addition, the estimate (30) is of interest in its own right, since we have no direct measurements of high-frequency fluctuations of the magnetic field near the Sun.

6. COMPARISON WITH OBSERVATIONS

The observational data presented in [2, 3] show that the degree ξ of anisotropy of density fluctuations exceeds ten for regions close to the Sun ($r \leq 3r_*$), smoothly decreases to six to seven as r increases to $r \leq 6r_*$, then sharply falls to $\xi \leq 2$ and remains virtually constant as r continues to increase ($r \leq 15r_*$). For comparison with the data of [2, 3], we introduce a model with a radial magnetic field $B = B_*(r_*/r)^2$, where $B_* \approx 2$ G, and a certain plasma-density profile N [24, 25], with $N \approx 2 \times 10^5$ cm^{-3} for $r = 3r_*$ and $N \approx 5 \times 10^4$ cm^{-3} for $r = 5r_*$. These densities are typical for the low and middle heliolatitudes corresponding to the data of [2, 3]. Note that, at solar-activity minimum, the plasma density at near-polar latitudes associated with regions situated over coronal holes could be smaller by a factor of two to three [26]. For these densities and magnetic fields, the Alfven velocity becomes $v_A \approx 8 \times 10^7$ cm/s at $r = 3r_*$, and $v_A \approx 6.5 \times 10^7$ cm/s at $r = 5r_*$. We also assume that the sound speed is virtually independent of r and is equal to $v_S \approx 10^7$ cm/s.

To estimate the degree of anisotropy of plasma-density fluctuations for $r \leq 5r_*$, we use the relation $\xi^a = \xi_*^a \frac{r_* v_A}{r v_{A*}}$ for Alfven waves. Adopting the Alfven velocity obtained above, we find $\xi^a \approx 27$ for $r = 3r_*$ and $\xi^a \approx 13$ for $r = 5r_*$. Taking into account (24), we obtain for the anisotropy of the angular distribution of the slow magneto-acoustic waves

$$\xi = \xi^s \approx 13 \quad \text{at} \quad r = 3r_*,$$
$$\xi = \xi^s \approx 7 \quad \text{at} \quad r = 5r_*. \quad (31)$$

The estimates (31) are in good agreement with the data of [2, 3].

At the same time, linear effects cannot lead to an abrupt decrease in ξ within 1–$2r_*$. To estimate the role of nonlinear processes, we will determine the level of small-scale fluctuations of the density and magnetic field from the observations. The data presented in [2, 3] give for the structure function of the phase fluctuations of the radio signals $D_S(b_\perp = 10^6$ cm$) \approx 3$–4 and $D_S(b_\parallel = 10^6$ cm$) \approx 0.5$–0.6 for a wavelength of $\lambda = 20$ cm and heliocentric distances $r = 5$–$6r_*$. Substituting these

$D_S(b_{\parallel,\perp})$ into (28) and (29) and adopting $\Delta X \approx r \approx 4 \times 10^{11}$ cm for the effective thickness of the scattering layer, we obtain using (26) the following estimates for the rms density fluctuations δN_0^2 and relative level of slow magneto-acoustic waves:

$$\delta N_0^2 \approx 10^5 \ \text{cm}^{-6}, \quad \frac{A_3}{\rho v_S^2} = \frac{C}{N^2} \approx 10^{-5}. \tag{32}$$

The results (32) and relation (30) provide an estimate for the relative level of high-frequency Alfven waves

$$A_1/\rho v_A^2 \approx 3 \times 10^{-4}. \tag{33}$$

The estimate (33) is in good agreement with the relative level of Alfven waves escaping from the corona (4). The agreement between the estimates (33) and (4) is due to the fact that the nonlinear absorption of Alfven waves via slow magneto-acoustic waves is sufficiently weak that the relative level of the Alfven turbulence does not decrease (and even slowly increases [11]) with heliocentric distance.

Rewriting (23) in the form

$$R_{an}(b) = b \frac{v v_S}{v_A^2} \left(\frac{\rho v_A^2}{A_1} \right)^2, \tag{34}$$

we obtain for $b = 10^6$ cm and $v \approx v_S \approx 10^7$ cm/s (assuming that the supersonic flow is not formed for $r \approx 6r_*$)

$$R_{an} \approx 6r_*. \tag{35}$$

The estimate (35) is in very good agreement with the position of the region of abrupt decrease in the degree of anisotropy of the plasma-density fluctuations derived in [2, 3]. The estimates for R_{fr} (21) and R_{12} (22), similar to (35), show that $R_{fr} \leq R_{12} \leq 1r_*$ for $b = 10^6$ cm. Thus, our regime with a quasi-stationary frequency spectra and linear evolution of the angular distribution of the waves is established at fairly small distances from the outer boundary of the corona.

7. CONCLUSIONS

Our model for the local generation of small-scale density fluctuations by high-frequency Alfven waves in the region of acceleration of the solar wind can explain the anisotropy of density inhomogeneities and its radial evolution observed in radio-propagation experiments (adopting reasonable assumptions about the average magnetic field and plasma density). Our most important conclusion is that appreciable anisotropy first arises when the Alfven waves travel through the transition layer between the chromosphere and corona. Due to the fairly slow nonlinear angular redistribution of the waves, a high degree of anisotropy is preserved to heliocentric distances of about $6R_S$ for inhomogeneities with scales exceeding 10^6 cm. At the same time, the radial elongation of the inhomogeneities probably provides evidence for fairly strong radial magnetic fields. According to the radio-propagation data, the level of the high-frequency Alfven waves is rather low in the acceleration region. In contrast to low-frequency waves, these waves do not play a significant role in the energy balance of the supercorona. At the same time, the high-frequency Alfven waves are apparently important for the small-scale structure and its dynamics.

Note that a decrease in the anisotropy of inhomogeneities could be explained by other effects: (a) nonlinear processes associated with a cascade energy transfer from the low-frequency ($\omega \leq 10^{-2}$ s^{-1}) part of the energy spectrum of the Alfven waves; (b) an increase in the relative contribution of fast magneto-acoustic waves; (c) approach of the transverse scale of the inhomogeneities to the ion cyclotron radius with increasing distance from the Sun, with the associated manifestation of the kinetic properties of the Alfven waves (in particular, the presence of density fluctuations in a linear approximation [27, 28]). However, our estimates indicate that these effects become significant only for $r > 10r_*$, beyond the region where the sharp decrease in the anisotropy of the density inhomogeneities is observed.

The anisotropy of the inhomogeneities in the solar supercorona was explained in [4, 5] as the effect of strongly elongated small-scale stationary structures. Our wave interpretation also supposes the presence of non-stationary inhomogeneities, and enables investigation of the evolution of the anisotropy. At the same time, as in [4, 5], we conclude that the strong elongation of the inhomogeneities is associated with the physical conditions in the lowest layers of the solar corona.

Another argument supporting our model is the agreement of the power index of the turbulence energy spectra (13) and (25) with that implied by the radio-propagation measurements. Namely, the structure functions of the phase fluctuations detected in [2, 3] depend approximately linearly on the baseline length for the radio interferometer, independent of its orientation, which indicates a spatial spectrum of the form (25). The phase-fluctuation spectra for the radio signals detected in [29] for the solar-wind acceleration region are also in agreement with the spatial spectra (13) and (25) of the density fluctuations.

In [6], the relative level of the Alfven waves in the solar-wind acceleration region was derived from fluctuations of the Faraday rotation of polarized signals emitted by spacecraft. For low fluctuation frequencies (from 10^{-3} to 10^{-2} Hz), the relative level of the Alfven waves was higher than the estimate (33) by two orders of magnitude. This difference could be due to the weak linear absorption of low-frequency Alfven waves in the corona. In contrast to high-frequency Alfven waves, nonlinear processes in the solar-wind acceleration region become insignificant for such waves.

Our model is applicable for the inner regions of the solar wind, which display fairly smooth longitudinal and transverse spatial variations in the plasma parameters. We expect that the plasma inhomogeneities will also be anisotropic in the relatively small volumes occupied, for example, by streamers. However, the properties of the wave turbulence in such regions with sharp transverse gradients, including changes in the sign of the magnetic field, require special study.

ACKNOWLEDGMENTS

REFERENCES

1. V. F. Bedevkin and V. V. Vitkevich, Tr. Fiz. Inst. Akad. Nauk SSSR **38**, 96 (1967).
2. J. W. Armstrong, W. A. Coles, M. Kojima, and B. J. Rickett, Astrophys. J. **358**, 685 (1990).
3. R. R. Grall, W. A. Coles, S. R. Spangler, *et al.*, J. Geophys. Res. **102**, 263 (1997).
4. R. Woo, Nature **379**, 321 (1996).
5. R. Woo and S. Habbal, Astrophys. J. Lett. **474**, L139 (1997).
6. V. E. Andreev, A. I. Efimov, L. N. Samoznaev, *et al.*, Sol. Phys. **176**, 387 (1997).
7. W. A. Coles and S. Maagoe, J. Geophys. Res. **77**, 5622 (1972).
8. M. Kojima, Publ. Astron. Soc. Jpn. **31**, 231 (1979).
9. Y. Yamauchi, M. Kojima, M. Tokumaru, *et al.*, J. Geomagn. Geoelectr. **48**, 1201 (1996).
10. Y. Yamauchi, M. Tokumaru, M. Kojima, *et al.*, J. Geophys. Res. **103**, 6571 (1998).
11. I. V. Chasheĭ and V. I. Shishov, Astron. Zh. **60**, 594 (1983) [Sov. Astron. **27**, 346 (1983)].
12. A. A. Galeev and V. N. Oraevskiĭ, Dokl. Akad. Nauk SSSR **147**, 71 (1962) [Sov. Phys. Dokl. **7**, 988 (1963)].
13. J. V. Hollweg, Rev. Geophys. Space Phys. **13**, 263 (1975).
14. I. V. Chasheĭ and V. I. Shishov, Pis'ma Astron. Zh. **7**, 276 (1981) [Sov. Astron. Lett. **7**, 153 (1981)].
15. G. P. Zank and W. H. Matthaeus, J. Geophys. Res. **97**, 17189 (1992).
16. C. Y. Tu and E. Marsch, Space Sci. Rev. **73**, 1 (1995).
17. M. Velli, B. Buti, B. E. Goldstein, *et al.*, AIP Conf. Proc. **471**, 445 (1999).
18. B. Buti and L. Nocera, AIP Conf. Proc. **471**, 173 (1999).
19. B. Buti, V. L. Galinski, V. I. Shevchenko, *et al.*, Astrophys. J. **523**, 849 (1999).
20. I. V. Chasheĭ and V. I. Shishov, Astron. Zh. **61**, 474 (1984) [Sov. Astron. **28**, 279 (1984)].
21. I. V. Chasheĭ, Astron. Zh. **75**, 612 (1998) [Astron. Rep. **42**, 540 (1998)].
22. I. N. Toptygin, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **16**, 971 (1973).
23. *Plasma Electrodynamics,* Ed. by A. I. Akhiezer, I. A. Akhiezer, R. V. Polovin, *et al.* (Nauka, Moscow, 1974; Pergamon, Oxford, 1975).
24. D. O. Muhleman, P. B. Esposito, and J. D. Anderson, Astrophys. J. **211**, 943 (1977).
25. O. G. Badalyan and M. A. Livshits, Astron. Zh. **63**, 1029 (1986) [Sov. Astron. **30**, 609 (1986)].
26. M. Guhathakurta, A. Fludra, S. E. Gibson, *et al.*, J. Geophys. Res. **104**, 9801 (1999).
27. Yu. M. Voĭtenko, A. N. Krishtal', S. V. Kuts, *et al.*, Geomagn. Aéron. **30**, 901 (1990).
28. J. V. Hollweg, J. Geophys. Res. **104**, 14811 (1999).
29. R. Woo and J. W. Armstrong, J. Geophys. Res. **84**, 7288 (1979).

Translated by V. Badin