Astronomy Reports, Vol. 46, No. 10, 2002, pp. 779–790. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 867–878.
Original Russian Text Copyright © 2002 by Holovatyy, Melekh.

The Chemical Composition of HII Regions in Blue Compact Dwarf Galaxies

V. V. Holovatyy[1] and B. Ya. Melekh[2*]

[1]*Chair of Astrophysics, Ivan Franko National University of Lviv, Kyryla and Mephodia str. 8, Lviv, 79005 Ukraine*

[2]*Astronomical Observatory, Ivan Franko National University of Lviv, Kyryla and Mephodia str. 8, Lviv, 79005 Ukraine*

Received December 3, 2001; in final form, May 23, 2002

Abstract—A set of empirical relationships for the ionization correction factors used by various authors to determine the chemical compositions of the gas in nebular objects is tested. New expressions for the ionization correction factors are used to find the nebular-gas compositions in HII regions in blue compact dwarf galaxies. The abundances of He, N, O, Ne, S, and Ar in 41 HII regions are determined. The derived elemental abundances are compared with the results of other studies. The $Y-$O/H, $Y-$N/H, and $Y-Z$ dependences are analyzed in detail. The primordial helium abundance Y_p and its enrichment dY/dZ are also determined. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Blue compact dwarf galaxies (BCD galaxies) are characterized by active star-formation processes, with giant HII regions observed around the sites of star formation. These galaxies have low metallicities, and evolve from material that has not yet been processed by stellar evolution. Therefore, BCD galaxies are important objects for studies of the primordial helium abundance Y_p and the rate of its enrichment dY/dZ during the chemical evolution of matter in our Metagalaxy.

The determination of accurate elemental abundances in HII regions in BCD galaxies is a large task. Only a small number of ion lines of a single element are observed in real nebulae, so that the numbers of ions in most ionization stages cannot be derived directly from observations. So-called ionization correction factors (ICFs) are usually used to determine the abundance of an element based on its ion abundance. Some of these factors (which are empirical expressions) are based only on the proximity of the ionization potentials of the corresponding ions and do not take into account differences in their effective ionization cross sections. Others are based on calculations using grids of photoionization models (PhM) for HII regions (see, for example, [1]). However, calculated total relative ion abundances are usually used to find ICFs. It is obvious that using ICFs derived from PhM is a more accurate method for determining heavy-element abundances. However, we believe that ion abundances obtained from diagnostics of integrated PhM spectra should be used when searching for expressions for ICFs. The relative ion abundances in real objects are determined using precisely this method. New expressions for ICFs obtained taking this into account are presented in [2].

The most comprehensive study of the chemical compositions of HII regions in BCD galaxies is that discussed in the recent papers of Izotov *et al.* [3–5], Izotov and Thuan [6]. However, they use expressions for ICFs based on either empirical relationships (to determine the O, N, Ne abundances) or calculations using the grid of PhM for HII regions of Stasińska [1] (to determine the He, S, Ar abundances).

In the current paper, we present and discuss new estimates of the chemical compositions of HII regions in BCD galaxies. We have adopted the spectra of Izotov *et al.* [3–5], Izotov and Thuan [6] as an observational basis for this study.

2. TESTING EXPRESSIONS FOR DETERMINING ICFs

Before turning to our derivation of the chemical compositions of HII regions in BCD galaxies using the new ICFs [2], it is of interest to test the ICF expressions used for this purpose by various authors.

The main expressions used most frequently to determine the compositions of HII regions and planetary nebulae are listed in Table 1. Our analysis of these expressions will enable us to test the suitability

*E-mail: gol@astro.franko.lviv.ua;melekh@astro.franko.lviv.ua

1063-7729/02/4610-0779$22.00 © 2002 MAIK "Nauka/Interperiodica"

Table 1. Empirical expressions for determining the compositions of nebular gas and values of $\left(\frac{(A/\mathrm{H})_{\mathrm{calc}}}{(A/\mathrm{H})_{\mathrm{mod}}}\right)$ for HII regions in BCD galaxies

No.	Reference	Expression	$\frac{(A/\mathrm{H})_{\mathrm{calc}}}{(A/\mathrm{H})_{\mathrm{mod}}}$
T1	[16, 17]	$\mathrm{He/H} = [\mathrm{He}^{+} + \mathrm{He}^{++}]/\mathrm{H}^{+}$	0.87
T2	[18]	$\mathrm{He/H} = (\mathrm{He}^{+}/\mathrm{H}^{+})[1 - 0.25\,(\mathrm{O}^{+}/\mathrm{O})]^{-1}$	0.95
T3	[19]	$\mathrm{C/H} = (\mathrm{C}^{+} + \mathrm{C}^{2+} + \mathrm{C}^{3+})/\mathrm{H}^{+}$	0.71
T4	[20]	$\mathrm{C/H} = ((\mathrm{He}^{+} + \mathrm{He}^{++})/\mathrm{He}^{+})^{1/3}(\mathrm{C}^{+} + \mathrm{C}^{2+} + \mathrm{C}^{3+})/\mathrm{H}^{+}$	0.71
T5	[21]	$\mathrm{N/O} = \mathrm{N}^{+}/\mathrm{O}^{+}$	0.88
T6	[3–5, 17]	$\mathrm{N/H} = (\mathrm{O}/\mathrm{O}^{+})(\mathrm{N}^{+}/\mathrm{H}^{+})$	0.87
T7	[15, 17, 18]	$\mathrm{O/H} = [(\mathrm{He}^{++} + \mathrm{He}^{+})/\mathrm{He}^{+}][(\mathrm{O}^{+} + \mathrm{O}^{++})/\mathrm{H}^{+}]$	1.02
T8	[3–5, 22]	$\mathrm{O/H} = (\mathrm{O}^{+} + \mathrm{O}^{++})/\mathrm{H}^{+}$	1.01
T9	[20]	$\mathrm{O/H} = [1/[1 - 0.95\mathrm{N}^{4+}/(\mathrm{N}^{+} + \mathrm{N}^{++} + \mathrm{N}^{3+} + \mathrm{N}^{4+})]]\times$ $\times(\mathrm{O}^{+} + \mathrm{O}^{++} + \mathrm{O}^{3+})/\mathrm{H}^{+}$	1.01
T10	[20]	$\mathrm{O/H} = [(\mathrm{He}^{+} + \mathrm{He}^{++})/\mathrm{He}^{+}]^{2/3}(\mathrm{O}^{+} + \mathrm{O}^{++})/\mathrm{H}^{+}$	1.01
T11	[3–5, 18]	$\mathrm{Ne/H} = (\mathrm{Ne}^{++}/\mathrm{H}^{+})(\mathrm{O}/\mathrm{O}^{++})$	1.27
T12	[21]	$\mathrm{Ne/O} = \mathrm{Ne}^{++}/\mathrm{O}^{++}$	0.87
T13	[3–5]	$\mathrm{S/H} = [0.013 + x(5.10 + x[-12.78 + x(14.77 - 6.11x)])]^{-1}\times$ $\times(\mathrm{S}^{+} + \mathrm{S}^{++})/\mathrm{H}^{+}$, where $x = \mathrm{O}^{+}/\mathrm{O}$	1.49
T14	[22]	$\mathrm{Ar/H} = [(\mathrm{Ar}^{++} + \mathrm{Ar}^{3+} + \mathrm{Ar}^{4+})/\mathrm{H}^{+}][(\mathrm{S}^{+} + \mathrm{S}^{++})/\mathrm{S}^{++}]$	1.03
T15	[20, 21]	$\mathrm{Ar/H} = 1.87\mathrm{Ar}^{++}/\mathrm{H}^{+}$	0.86
T16	[3–5]	$\mathrm{Ar/H} = (0.99 + x[0.091 + x(-1.14 + 0.077x)])^{-1}\times$ $\times(\mathrm{Ar}^{2+} + \mathrm{Ar}^{3+})/\mathrm{H}^{+}$, where $x = \mathrm{O}^{+}/\mathrm{O}$	1.10
T17	[3–5]	$\mathrm{Ar/H} = [0.15 + x(2.39 - 2.64x)]^{-1}(\mathrm{Ar}^{2+}/\mathrm{H}^{+})$, where $x = \mathrm{O}^{+}/\mathrm{O}$	1.09

of each for estimating the compositions of HII regions in BCD galaxies. We used calculated values obtained using the PhM grid of [2] for this purpose. One feature distinguishing our approach is that, in our models, the ionizing radiation field is represented by the energy distribution beyond $\lambda \leq 912$ Å instead of by the effective temperature of a star (or stars); in addition, we used relative ion abundances derived from analyses of nebular PhM spectra to determine the chemical compositions.

Our testing method was as follows. We specified a chemical composition $(A/\mathrm{H})_{\mathrm{mod}}$ in each PhM. We first found the relative ion abundances via an analysis of the integrated spectrum for each of the PhM, which were then used to calculate the compositions $(A/\mathrm{H})_{\mathrm{calc}}$ using each of the analytical expressions listed in Table 1 (labeled T1–T17). A comparison of the corresponding compositions $(A/\mathrm{H})_{\mathrm{mod}}$ and the calculated values $(A/\mathrm{H})_{\mathrm{calc}}$ provided the test criterion: the closer these two values, the more accurate the expression. Note that some of the expressions in Table 1 used to determine the abundance of a particular element depend on the abundance of some other element. If the expressions for both elements are given in the corresponding reference, we tested that formula; otherwise we eliminated it from the analysis. We also excluded expressions for which the method for analyzing nebular gas implemented in the DIAGN code [7] could not be used to determine the corresponding ion abundance.

Table 1 lists the results of our tests for those ICF expressions that can reproduce the elemental abundances to accuracies $\leq 50\%$. Note that the expressions T6, T7, T8, T11, T13, T16, T17 in Table 1

Table 2. Composition of HII regions in BCD galaxies, found by us using the ICFs listed in Table 1

		B72		Sm96		$\frac{N}{H}\times10^6$	$\frac{O}{H}\times10^5$				$\frac{Ne}{H}\times10^6$	$\frac{S}{H}\times10^6$	$\frac{Ar}{H}\times10^7$		
		$\frac{He}{H}\times10^2$													
No.	HII region	T1	T2	T1	T2	T6	T7	T8	T9	T10	T11	T13	T14	T16	T17
1	0723+692A	9.69	9.76	11.6	11.7	2.06	6.13	6.09	6.09	6.12	8.83	2.52	1.76	1.61	3.50
2	0749+568	7.84	8.08	9.51	9.84	1.50	4.25	4.18	4.18	4.23	7.92	1.43	1.49	1.16	2.12
3	0749+582	10.5	10.9	12.7	13.1	4.07	12.3	12.3	12.3	12.3	25.1	2.89	1.66	1.36	3.16
4	0907+543	10.1	10.1	12.2	1.22	1.95	8.15	7.98	7.98	8.09	16.2	2.24	2.36	2.12	4.19
5	0917+527	8.78	9.04	10.6	1.10	1.67	5.84	5.73	5.73	5.80	12.0	1.23	1.27	0.98	1.77
6	0926+606	8.43	–	10.0	–	1.40	7.85	7.74	7.74	7.81	–	–	–	–	–
7	0930+554	7.47	7.46	8.86	8.89	0.35	1.25	1.22	1.22	1.24	1.82	0.32	0.25	0.20	0.48
8	0943+561A	7.27	7.54	8.40	8.71	–	4.40	4.40	–	4.40	8.77	1.80	1.53	1.18	2.67
9	1030+583	8.68	–	10.2	–	0.85	6.09	5.96	5.96	6.04	–	–	–	–	–
10	1054+365	9.02	9.31	10.9	11.3	2.67	8.57	8.57	8.57	8.57	14.2	3.13	2.59	2.42	4.55
11	1116+583B	7.68	7.63	9.26	9.25	1.04	3.45	3.37	3.37	3.42	5.28	1.38	0.94	0.77	2.13
12	1128+573	7.10	7.29	8.74	8.97	1.30	4.59	–	4.59	4.59	6.84	0.24	–	0.60	1.61
13	1205+557	8.15	8.59	9.96	10.5	2.16	6.59	6.49	6.49	6.56	1.09	1.46	1.30	1.11	1.67
14	1222+614	8.65	8.76	10.4	10.6	2.22	7.75	7.64	7.64	7.71	1.29	2.43	2.09	1.85	3.64
15	1223+487	9.03	9.10	10.7	10.8	1.91	4.98	4.93	4.93	4.97	7.61	2.16	1.54	1.40	3.18
16	1256+351	9.30	9.52	11.1	11.4	2.51	8.60	8.53	8.53	8.58	13.3	2.65	2.39	2.17	4.37
17	1319+579A	9.63	9.85	11.4	11.6	3.00	10.4	10.4	10.4	10.4	16.1	2.80	2.16	1.95	3.14
18	1319+579B	8.31	9.01	10.1	11.0	3.30	7.14	–	7.04	7.10	10.9	0.91	–	–	–
19	1319+579C	8.53	9.34	10.4	11.4	4.39	13.0	12.9	12.9	12.9	25.4	2.99	2.58	2.36	2.92
20	1358+576	7.86	8.14	9.23	9.58	3.75	6.60	6.52	6.52	6.57	11.2	1.65	1.82	1.51	2.41
21	1441+294	9.45	9.78	10.8	11.2	3.15	8.74	–	8.61	8.70	12.3	0.46	–	2.23	3.30
22	1533+574A	8.28	8.93	9.98	10.8	2.84	5.88	5.88	5.88	5.88	10.6	1.92	2.24	1.76	2.61
23	1533+574B	9.60	10.1	11.5	12.1	3.05	11.6	11.5	11.5	11.6	19.3	2.80	3.38	3.06	3.95
24	1535+554	7.90	8.51	9.27	9.99	2.96	10.4	10.4	10.4	10.4	17.8	3.13	2.37	2.19	3.26
25	0832+699	8.26	8.61	10.1	10.5	0.86	2.66	2.66	2.66	2.66	4.17	0.95	0.73	0.63	1.30
26	0940+544N	8.65	8.82	10.3	10.5	0.57	1.51	1.51	1.51	1.51	2.86	0.64	0.59	0.52	1.05
27	0946+558	9.47	9.73	11.4	11.8	2.75	8.79	8.68	8.68	8.75	15.3	3.23	2.56	2.39	3.95
28	0948+532	9.51	10.2	11.4	12.3	0.60	0.97	0.97	0.97	0.97	4.28	0.47	0.99	0.87	1.21
29	1135+581	9.25	9.53	11.1	11.4	3.19	8.46	8.34	8.34	8.42	16.6	2.17	2.93	2.61	4.22
30	1152+579	9.65	9.76	11.5	11.7	2.51	5.15	5.10	5.10	5.13	8.88	1.81	1.48	1.31	2.65
31	1159+545	8.15	8.31	9.80	10.0	1.66	2.15	2.14	2.14	2.14	3.67	0.83	0.80	0.71	1.51
32	1211+540	8.40	8.32	10.0	9.97	1.18	3.67	3.58	3.58	3.64	5.66	1.75	0.83	0.76	1.86
33	1249+493	9.07	9.29	10.7	11.0	1.09	4.04	3.99	3.99	4.02	6.92	1.25	0.74	0.66	1.22
34	1331+493N	8.26	8.48	10.0	10.3	1.42	5.00	5.00	5.00	5.00	8.53	1.68	0.94	0.87	1.76
35	1331+493S	7.52	8.19	8.81	9.59	2.30	5.87	–	5.87	5.87	10.8	0.57	–	–	–
36	1437+370	8.68	8.69	10.5	10.5	3.29	7.04	6.87	6.87	6.98	12.3	2.70	2.82	2.43	5.02
37	1415+437	8.72	8.96	10.1	10.5	0.69	2.39	2.34	2.34	2.37	–	0.68	0.88	0.71	1.06
38	1420+544	11.6	11.7	13.4	13.5	1.30	4.50	4.48	4.48	4.49	–	1.42	1.45	1.34	2.30
39	1533+469	9.18	9.75	11.1	11.8	6.93	8.45	8.38	8.38	8.42	12.8	2.06	3.18	2.14	3.28
40	IZw18(NW)	7.14	6.98	8.41	8.28	0.36	1.19	1.15	1.15	1.18	1.69	0.46	0.26	0.23	0.65
41	IZw18(SE)	9.04	9.42	10.8	11.2	0.69	2.21	2.20	2.20	2.21	3.63	0.55	0.42	0.36	0.69

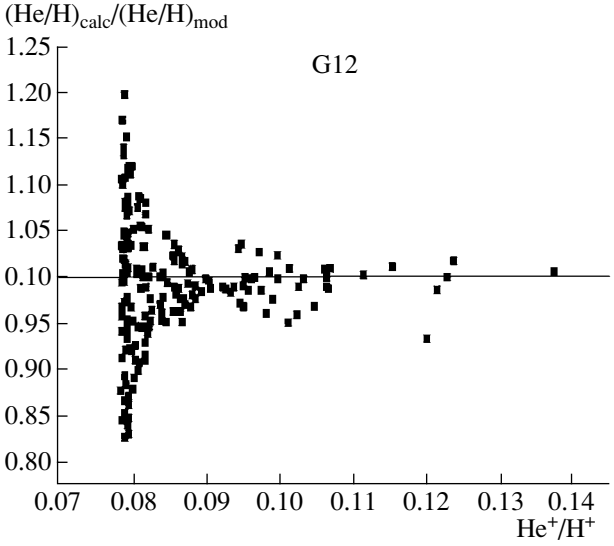

Fig. 1. Dependence of $(He/H)_{calc}/(He/H)_{mod}$ on He^+/H^+ according to the PhM grid calculations, constructed using the ICF G12.

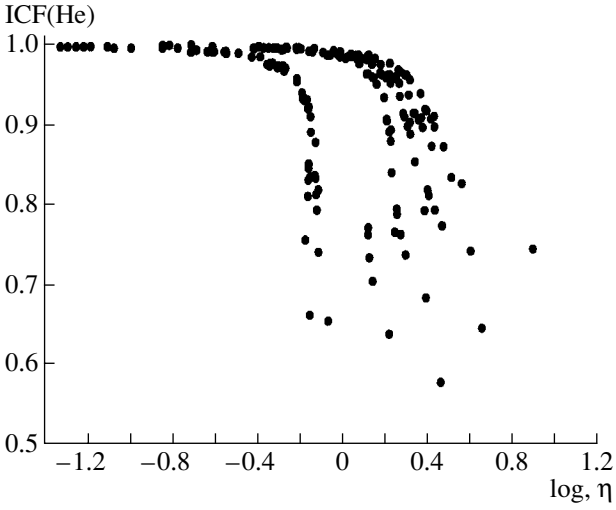

Fig. 2. Dependence of ICF(He) on the parameter $\eta = (O^+/S^+)(S^{++}/O^{++})$ according to the PhM grid calculations.

were used by Izotov *et al.* [3–5] to derive the A/H abundances in the same HII regions. The table shows that not all of these expressions (especially T13) are sufficiently accurate to be used to determine A/H ratios in these objects, supposing that an accuracy $\leq 10\%$ is necessary for this purpose. Note also that the abundance ratios $(A/H)_{calc}/(A/H)_{mod}$ in Table 1 were derived by averaging over the number of corresponding points on the PhM grid. The compositions found for 41 HII regions in BCD galaxies using these expressions are presented in Table 2.

3. THE CHEMICAL COMPOSITIONS OF HII REGIONS IN BCD GALAXIES

3.1. Determination of the Helium Abundance

To derive the He/H ratios in the BCD galaxy HII regions we are studying, we must know the relative ion abundances He^+/H^+, He^{++}/H^+ and an appropriate ICF expression. The abundances He^+/H^+ and He^{++}/H^+ are determined from the intensity ratios of the recombination lines HeI/H_β and $HeII/H_\beta$ and are given by Izotov *et al.* [4]. We adopted the values for these quantities from [4], which were derived from the recombination coefficients of Brocklehurst (B72) [8] and Smits (Sm96) [9] in various combinations with the collision rates of Kingdon and Ferland [10]. The ICF expressions we have calculated with the PE model grid (designated G1–G32) are presented in [2]. However, for various reasons, we have not used them here to determine the He/H ratios, since they are not entirely accurate. Figure 1 shows $(He/H)_{calc}/(He/H)_{mod}$ as a function of $(He^+/H^+)_{mod}$, derived from the data of [2] using the ICF expression G12 [2, Table 3]. The results of Izotov *et al.* [3, 4] suggest that, for most of the objects considered, the values of He^+/H^+ found from observations are in the range 0.072–0.090. However, as we can see from Fig. 1, there is a wide scatter of $(He/H)_{calc}/(He/H)_{mod}$ in this range of He^+/H^+. A similar conclusion can be drawn using the expressions G1 and G23 from [2, Table 3].

Note that the ICF for He in [3, 4] was derived from calculations for the PhM grid of Stasińska [1], and the corresponding ICF(He) – η relationship is shown in Fig. 6 of [3] [formula (30)], where $\eta = (O^+/S^+)(S^{++}/O^{++})$. The PhM grid of Stasińska shows that $ICF(He) \geq 1$. This means that the H^+ region is larger than the region of ionized helium. However, as shown by Armour *et al.* [11] and Ballantyne *et al.* [12], in low-metallicity HII regions, a situation can arise when, on the contrary, the region of ionized hydrogen is smaller than that of ionized helium. Figure 2 shows the ICF(He) values as a function of the parameter η, calculated using our PhM grid. We can see that ICF(He) has values close to unity for $\log\eta \leq -0.9$, whereas $\log\eta \leq -0.2$ for most HII regions of [3, 4]. The situation for the interval $-0.9 \leq \log\eta \leq -0.2$ remains unclear, since we could not find a clear dependence of the scatter of the points on any parameter of the PhM grid. In this interval, $ICF(He) < 1$, implying that the ionized-hydrogen region is smaller than the ionized-helium region. This means that the approximation for ICF(He) used by Izotov *et al.* [3, 4] is inaccurate, and a new relationship must be determined.

Table 3. Composition of HII regions in BCD galaxies, found by us with ICFs from the PhM

No.	HII region	B72 $\left(\frac{He}{H}, \Delta\frac{He}{H}\right) \times 10^2$		Sm96 $\left(\frac{He}{H}, \Delta\frac{He}{H}\right) \times 10^2$		$\left(\frac{O}{H}, \Delta\frac{O}{H}\right) \times 10^5$		$\left(\frac{N}{H}, \Delta\frac{N}{H}\right) \times 10^6$		$\left(\frac{Ne}{H}, \Delta\frac{Ne}{H}\right) \times 10^6$		$\left(\frac{S}{H}, \Delta\frac{S}{H}\right) \times 10^6$		$\left(\frac{Ar}{H}, \Delta\frac{Ar}{H}\right) \times 10^7$	
1	0723+692A	6.80	0.18	8.00	0.20	4.89	0.28	–	–	8.42	0.00	1.10	0.10	1.63	0.06
2	0749+568	6.98	0.32	7.98	0.61	3.56	0.26	1.64	0.25	–	–	–	–	1.36	0.01
3	0749+582	–	–	–	–	10.1	0.78	4.43	0.75	–	–	–	–	1.69	0.03
4	0907+543	7.36	0.45	8.25	0.55	6.46	0.35	2.05	0.37	14.8	0.12	1.19	0.06	2.17	0.08
5	0917+527	7.66	0.24	8.16	0.24	4.88	0.35	1.83	0.28	–	–	–	–	1.14	0.01
6	0926+606	7.09	0.24	7.87	0.31	6.80	0.37	1.58	0.20	12.3	0.49	1.56	0.06	1.72	0.09
7	0930+554	4.98	0.16	6.02	0.17	1.00	0.08	0.37	0.06	–	–	–	–	0.25	0.00
8	0943+561A	–	–	–	–	3.63	0.28	–	–	–	–	–	–	1.45	0.02
9	1030+583	7.00	0.19	7.80	0.20	4.83	0.25	0.92	0.14	7.19	0.08	1.18	0.04	1.85	0.10
10	1054+365	–	–	–	–	7.01	0.37	2.90	0.50	13.2	0.19	1.93	0.04	2.45	0.10
11	1116+583B	6.87	0.52	7.86	0.62	2.71	0.22	1.09	0.20	–	–	–	–	1.00	0.02
12	1128+573	–	–	–	–	3.71	0.30	1.40	0.25	–	–	–	–	0.78	0.02
13	1205+557	7.63	0.40	7.53	0.44	5.74	0.39	2.41	0.33	–	–	–	–	1.24	0.00
14	1222+614	7.34	0.25	7.94	0.26	6.28	0.34	2.37	0.41	12.0	0.19	1.36	0.02	1.84	0.07
15	1223+487	7.10	0.19	8.00	0.20	4.01	0.22	0.00	–	7.21	0.07	0.97	0.04	1.39	0.05
16	1256+351	–	–	–	–	7.19	0.39	2.70	0.46	13.0	0.32	1.59	0.01	2.09	0.08
17	1319+579A	7.26	0.25	8.05	0.26	8.30	0.45	3.23	0.56	14.4	0.02	1.90	0.16	2.19	0.10
18	1319+579B	6.79	0.79	7.07	0.79	6.44	0.43	3.75	0.46	–	–	–	–	–	–
19	1319+579C	7.42	0.21	7.61	0.21	11.9	0.79	5.01	0.59	–	–	–	–	2.44	0.02
20	1358+576	7.69	0.29	7.98	0.30	5.47	0.28	4.11	0.64	10.1	0.24	1.13	0.00	1.52	0.07

Table 3. (Contd.)

No.	HII region	B72 $\left(\frac{He}{H}, \Delta\frac{He}{H}\right) \times 10^2$		Sm96 $\left(\frac{He}{H}, \Delta\frac{He}{H}\right) \times 10^2$		$\left(\frac{O}{H}, \Delta\frac{O}{H}\right) \times 10^5$		$\left(\frac{N}{H}, \Delta\frac{N}{H}\right) \times 10^6$		$\left(\frac{Ne}{H}, \Delta\frac{Ne}{H}\right) \times 10^6$		$\left(\frac{S}{H}, \Delta\frac{S}{H}\right) \times 10^6$		$\left(\frac{Ar}{H}, \Delta\frac{Ar}{H}\right) \times 10^7$	
21	1441+294	7.70	0.90	8.09	0.64	7.09	0.36	3.44	0.53	10.4	0.19	–	–	2.39	0.11
22	1533+574A	–	–	–	–	5.21	0.36	3.22	0.44	–	–	–	–	1.95	0.00
23	1533+574B	6.97	0.28	7.66	0.30	9.39	0.48	3.39	0.49	15.3	0.23	2.29	0.04	3.54	0.20
24	1535+554	–	–	–	–	9.23	6.35	3.35	0.46	–	–	–	–	2.43	0.00
25	0832+699	–	–	–	–	2.22	1.67	0.95	0.15	–	–	–	–	0.75	0.01
26	0940+544N	–	–	–	–	1.21	0.07	–	–	2.69	0.00	0.32	0.03	0.54	0.02
27	0946+558	7.38	0.23	8.07	0.25	7.16	0.37	2.99	0.48	13.6	0.25	2.33	0.02	2.47	0.11
28	0948+532	7.50	0.19	8.20	0.27	0.87	0.59	0.67	0.09	–	–	–	–	0.94	0.00
29	1135+581	7.58	0.24	8.06	0.25	6.96	0.36	3.47	0.55	14.8	0.34	1.56	0.00	2.65	0.12
30	1152+579	6.90	0.19	8.10	0.20	4.12	0.23	–	–	8.25	0.05	0.90	0.05	1.34	0.05
31	1159+545	6.67	0.25	7.84	0.26	1.75	0.96	1.78	0.32	3.47	0.05	0.42	0.00	0.71	0.03
32	1211+540	6.97	0.19	8.17	0.26	2.93	0.17	–	–	5.38	0.07	0.72	0.02	0.74	0.03
33	1249+493	–	–	–	–	3.36	0.18	1.18	0.20	6.57	0.17	0.81	0.00	0.64	0.03
34	1331+493N	–	–	–	–	4.07	0.22	1.53	0.27	8.01	0.09	0.91	0.03	0.87	0.03
35	1331+493S	–	–	–	–	5.27	0.36	2.62	0.34	–	–	–	–	–	–
36	1437+370	7.09	0.23	7.55	0.24	5.76	0.31	3.48	0.60	11.7	0.27	1.48	0.00	2.34	0.09
37	1415+437	–	–	–	–	1.93	0.10	0.75	0.12	–	–	0.47	0.00	0.76	0.04
38	1420+544	–	–	–	–	3.55	0.21	–	–	–	–	0.87	0.12	1.44	0.06
39	1533+469	–	–	–	–	7.36	0.51	7.78	1.09	–	–	–	–	2.39	0.00
40	IZw18(NW)	–	–	–	–	0.92	0.08	0.38	0.07	–	–	–	–	0.30	0.01
41	IZw18(SE)	–	–	–	–	1.85	0.14	0.76	0.12	–	–	–	–	0.42	0.00

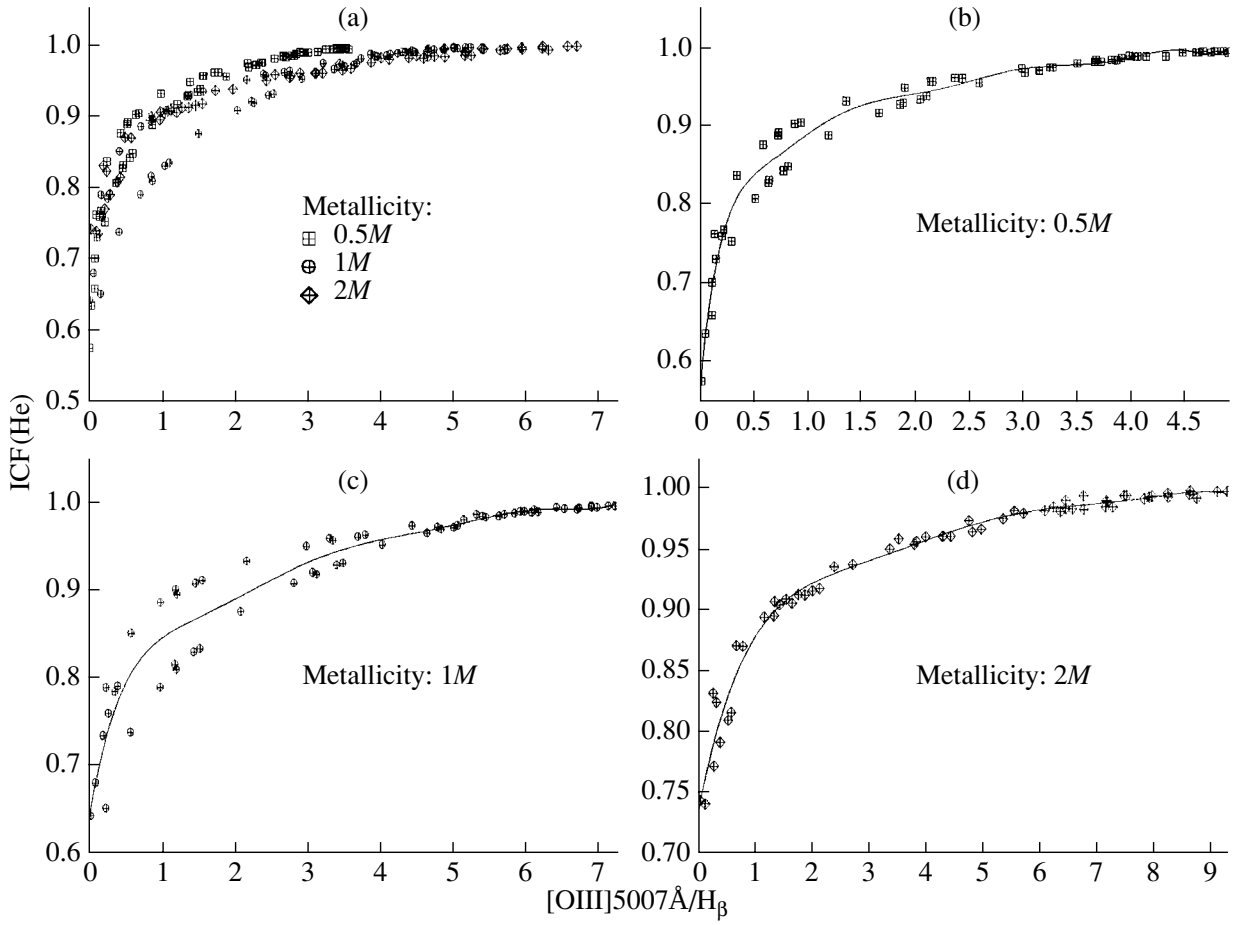

Fig. 3. Dependence of ICF(He) on the line intensity ratio [OIII]5007 Å/H_β, derived from the PhM grid calculations. The scatter is due to the presence of data for different metallicities; the dependence for each metallicity is shown separately.

For this purpose, we considered the dependence of ICF(He) on the relative intensity of the [OIII]5007 Å/H_β lines (Fig. 3a), calculated using our PhM grid. We excluded from the analysis the results of grid calculations with anomalously low numbers of photons beyond $\lambda < 228$ Å (model E1 in [2]), since the absence of the HeII λ4686 line in the observed spectra of an HII region introduces uncertainty in the magnitude of the jump beyond $\lambda < 228$ Å (see the procedure for optimizing the L_c energy spectrum for $\lambda < 228$ Å presented in [2]). Relationships labeled with different symbols correspond to different metallicities M, which correspond to the relative atomic-number abundance of heavy elements adopted in the PhM grid calculations: $1M$ (H : He : O : C : N : Ne : Si : S : Ar : Fe $= 1 : 8.01 \times 10^{-2} : 7.27 \times 10^{-5} : 1.71 \times 10^{-5} : 2.46 \times 10^{-6} : 1.39 \times 10^{-5} : 2.41 \times 10^{-6} : 2.00 \times 10^{-6} : 4.00 \times 10^{-7} : 1.45 \times 10^{-6}$), $0.5M$(H : He : O : C : N : Ne : Si : S : Ar : Fe $= 1 : 8.01 \times 10^{-2} : 3.64 \times 10^{-5} : 8.55 \times 10^{-6} : 1.23 \times 10^{-6} : 6.95 \times 10^{-6} : 1.21 \times 10^{-6} : 1.00 \times 10^{-6} : 2.00 \times 10^{-7} : 7.25 \times 10^{-7}$), and $2M$(H : He : O : C : N : Ne : Si : S : Ar : Fe $= 1 : 8.01 \times 10^{-2} : 1.45 \times 10^{-4} : 2.42 \times 10^{-5} : 4.92 \times 10^{-6} : 2.78 \times 10^{-5} : 4.82 \times 10^{-6} : 4.00 \times 10^{-6} : 8.00 \times 10^{-7} : 2.90 \times 10^{-6}$). We can see that the scatter of the points in this dependence is due to the different metallicities. Figures 3b–3d show the same dependences separately for each metallicity, together with the corresponding approximation curves

for $0.5M$:

$$\mathrm{ICF(He)} = 0.563223 + 1.495963x - 3.558336x^2 + 4.892831x^3 - 3.965862x^4 + 1.961679x^5 - 0.5981941x^6 + 0.1096308x^7 - 0.01106230x^8 + 4.71976904 \times 10^{-4}x^9, \quad (1)$$

for $1M$:

$$\mathrm{ICF(He)} = 0.64165 + 0.49213x - 0.48428x^2 + 0.26694x^3 - 0.081291x^4 + 0.0137274x^5 - 0.0012049x^6 + 4.2861026 \times 10^{-5}x^7, \quad (2)$$

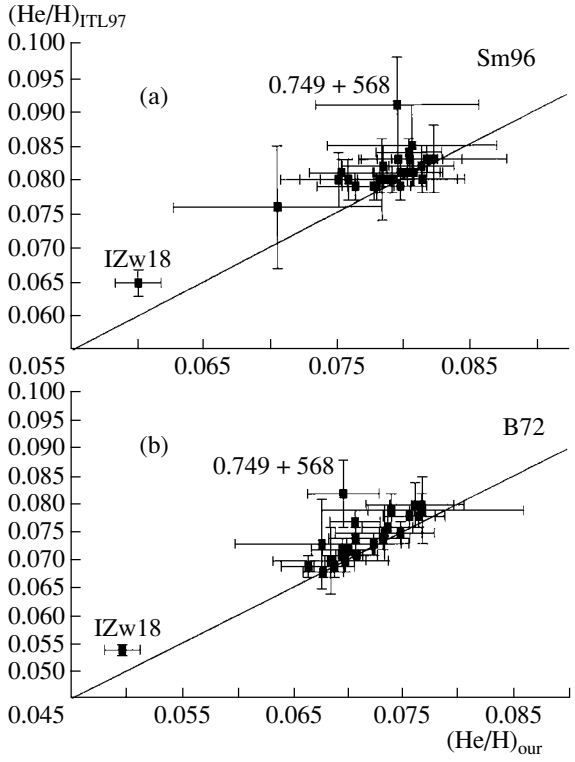

Fig. 4. Comparison of $(He/H)_{ITL97}$ and $(He/H)_{our}$ for the recombination coefficients Sm96 and B72.

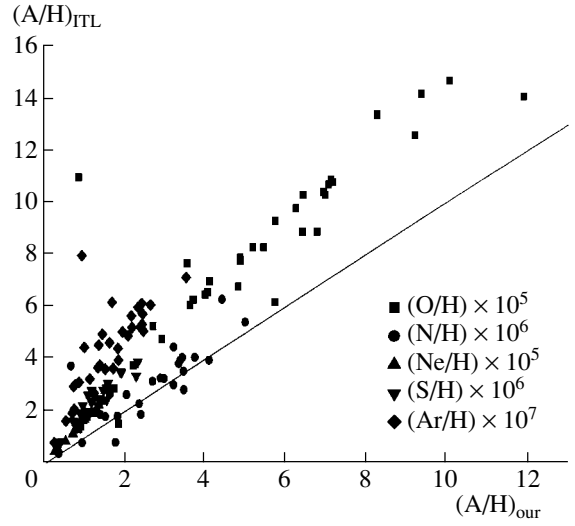

Fig. 5. Comparison of the heavy-element abundances $(A/H)_{our}$ and $(A/H)_{ITL}$. A straight line corresponding to identical values A/H is drawn.

and for $2M$:

$$\mathrm{ICF(He)} = 0.730985 + 0.244846x - 0.12993x^2 + 0.03771x^3 - 0.005835 * x^4 + 4.5410^{-4}x^5 - 1.39759 \times 10^{-5}x^6, \quad (3)$$

where $x = I_\lambda[\mathrm{OIII}]\,5007\text{Å}/I_\lambda(\mathrm{H}_\beta)$. The large number of significant digits in these expressions is due to the approximation using polynomials with coefficients with alternating signs. Then

$$\frac{\mathrm{He}}{\mathrm{H}} = \mathrm{ICF(He)}\frac{\mathrm{He}^+ + \mathrm{He}^{++}}{\mathrm{H}^+}. \quad (4)$$

Figure 3a shows that there is no clear division of the dependences for different metallicities. This enables us to omit any interpolation between metallicities when choosing the approximation expression. Thus, the process of calculating the ICF(He) values was reduced to the following. We determined the O/H abundance using the expressions for ICF(O/H) we found in [2]. We then chose the appropriate expression from (1)–(3) and calculated ICF(He).

The values of He/H and their standard deviations Δ(He/H) we found for the HII regions studied are listed in columns 3–6 of Table 2. Values of He/H were not determined for objects in whose spectra the HeII λ4686 line was absent or which were not considered by Izotov *et al.* [4].

Figures 4a, 4b compare our values, $(He/H)_{our}$, with those of Izotov *et al.* [4], $(He/H)_{ITL97}$. Figure 4a shows the values of He/H calculated using the He^+/H^+ and He^{++}/H^+ values from [4] and the recombination coefficients of Brocklehurst [8] (B72); Fig. 4b shows the same data derived using the coefficients of Smits [9] for HeI (Sm96). In both cases, we used the collision rates of Kingdon and Ferland [10]. Straight lines corresponding to identical values of He/H are also drawn. We can see that, for most of the objects, the $(He/H)_{ITL97}$ values are, on average, higher than our values $(He/H)_{our}$, though they are consistent within the errors. Exceptions are the objects 0930+554 (IZw18) and 0749+568, in which the $(He/H)_{ITL97}$ abundance is obviously overestimated.

3.2. Determination of the Heavy-Element Abundances

To determine the heavy-element abundances A/H in the HII regions, we used the new ICF expressions presented in [2] and adopted the observed forbidden-line intensity ratios of [3–6]. The values of n_e, T_e, and the ion abundances A^{+i}/H^+ were determined using the usual diagnostic methods for nebular gas [7]. We used the DIAGN code for this purpose, but with atomic data for different ions identical to those in the Cloudy 94 code [13, 14], which was used in the calculation of the PhM grid.

Table 3 presents the abundances A/H for 41 HII regions in BCD galaxies, calculated with our expressions for the ICFs. The standard deviations of the

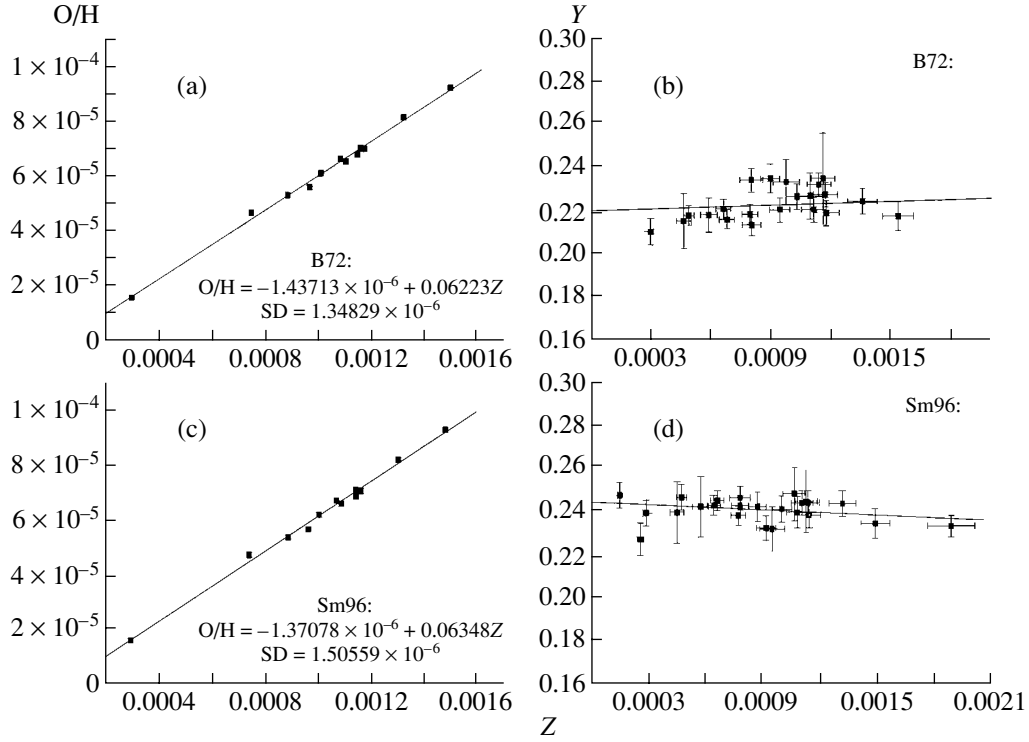

Fig. 6. (a, c) O/H $-Z$ and (b, d) $Y-Z$ dependences for the recombination coefficients of B72 and Sm96. The corresponding approximation expressions and standard deviations are also shown.

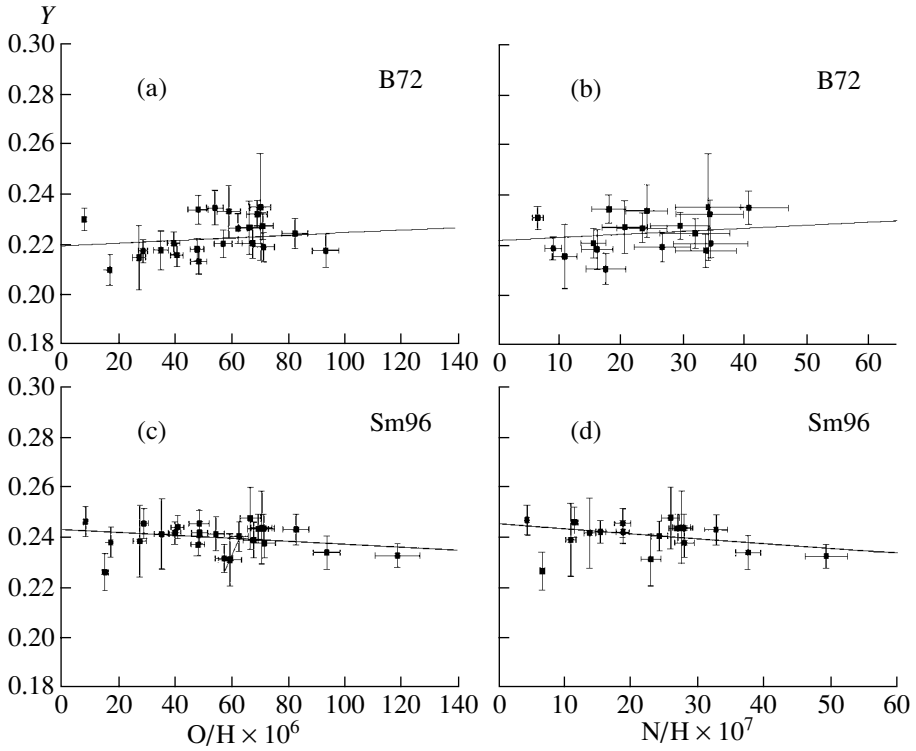

Fig. 7. $Y-$O/H and $Y-$N/H dependences for the recombination coefficients of (a, b) B72 and (c, d) Sm96.

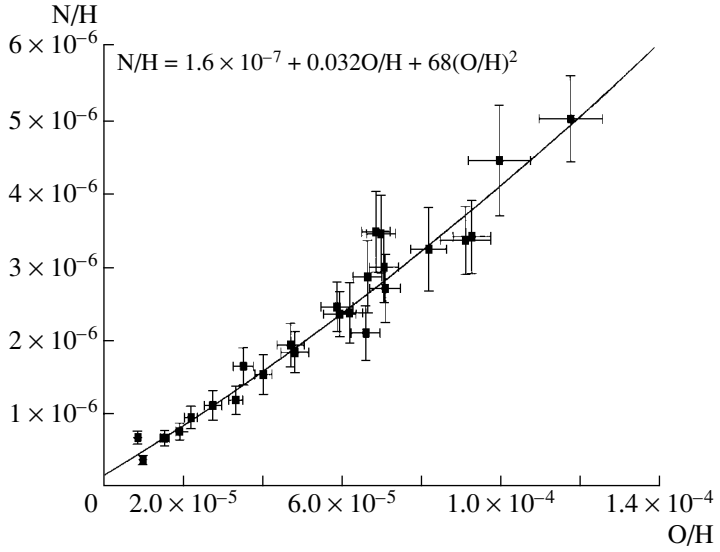

Fig. 8. N/H abundance as a function of O/H for HII regions in BCD galaxies.

abundances obtained by averaging the results found with the ICF expressions of [2] are also given. We used expressions G3, G4, and G26 from Table 3 in [2] to find O/H, G6 to find N/H, G29 to find Ne/H, G30 to find S/H, and G10, G11, G31, and G32 to find Ar/H. These expressions were chosen based on the requirement that the corresponding standard deviations for the A/H values not exceed 10%. Figure 5 compares our values $(A/\mathrm{H})_{\mathrm{our}}$ with the values $(A/\mathrm{H})_{\mathrm{ITL}}$ of Izotov *et al.* [3–5], Izotov and Thuan [6]. A straight line corresponding to identical values of A/H is also drawn. We can see that, for most of the objects, $(A/\mathrm{H})_{\mathrm{ITL}}$ is greater than $(A/\mathrm{H})_{\mathrm{our}}$.

3.3. Determination of the Primordial Helium Abundance Y_p and its Enrichment dY/dZ

We used a Y–Z relationship to determine Y_p and dY/dZ, where Y and Z are the atomic-mass abundances of helium and heavy elements, respectively:

$$Y = \frac{4\mathrm{He/H}(1-Z)}{1+4\mathrm{He/H}}, \quad Z = \frac{Za}{1+4\mathrm{He/H}+Za}, \quad (5)$$

$$Za = 14\mathrm{N/H} + 16\mathrm{O/H} + 20\mathrm{Ne/H} + 32\mathrm{S/H} + 40\mathrm{Ar/H}.$$

Extrapolating this dependence to $Z = 0$ yields Y_p, and its slope corresponds to the value of dY/dZ. When finding Z, we have not used other elements whose lines are absent from the spectra of HII regions in BCD galaxies. We have likewise not used Fe ions, since they are not included in the DIAGN code. When finding Y_p, we eliminated objects whose spectra do not display the HeII λ4686 line, since uncertainty in its intensity introduces uncertainty in the magnitude of the jump at λ228 Å [2]. For the same reasons as in [4], we have also omitted the low-metallicity object IZw18NW, which has an anomalously low helium abundance. The abundances of all these heavy elements are required to find Z [formula (5)]; therefore, the number of HII regions for the analysis decreases, since we could not find Ne/H and S/H abundances for some of them (Table 3). To take these into account, we plotted $\mathrm{O/H} - Z$ dependences using the appropriate data found with the B72 and Sm96 recombination coefficients (Figs. 6a, 6c). We see that this dependence is linear, and Z for the BCD galaxy HII regions can be estimated using the following approximations:

for B72:

$$Z = 16.07\mathrm{O/H} + 2.30 \times 10^{-5} \quad (6)$$

and for Sm96:

$$Z = 15.75\mathrm{O/H} + 2.16 \times 10^{-5}. \quad (7)$$

As a result, we derived Y–Z dependences for the B72 and Sm96 cases (Figs. 6b and 6d, respectively). Linear approximations to these dependences weighted with the corresponding standard deviations are

for B72:

$$Y = 0.2201(\pm 0.0040) + 2.90(\pm 4.48)Z, \quad (8)$$

for Sm96:

$$Y = 0.2440(\pm 0.0024) - 4.02(\pm 2.46)Z. \quad (9)$$

Some authors have used O/H and/or N/H instead of Z to determine Y_p. Using formulas (5)–(9), we constructed Y–O/H and Y–N/H relationships for

the B72 and Sm96 cases (Figs. 7a–7d). For the Y–O/H dependence, the value of Y was found by replacing Z in (5) with the corresponding relationships (6) and (7), and, for the Y–N/H dependence, the value of Y was determined using (5). The following approximation relationship between N/H and O/H is suitable for most of the objects (Fig. 7):

$$\mathrm{N/H} = 1.6 \times 10^{-7} + 0.032\mathrm{O/H} + 68(\mathrm{O/H})^2, \quad (10)$$

which we also used to determination the N/H abundance.

In this case, as we can see from Figs. 7a, 7b for the recombination coefficients from B72, the value of Y_p and the slopes $dY/d(\mathrm{O/H})$ and $dY/d(\mathrm{N/H})$ found via a linear approximation with weighting based on the corresponding standard deviations can be written as

$$Y = 0.2202(\pm 0.0039) + 50(\pm 72)\mathrm{O/H}, \quad (11)$$
$$Y = 0.2218(\pm 0.0041) + 1227(\pm 1616)\mathrm{N/H}.$$

Similarly, using the data from Sm96, we obtain (Figs. 7c, 7d):

$$Y = 0.2439(\pm 0.0024) - 60(\pm 39)\mathrm{O/H}, \quad (12)$$
$$Y = 0.2454(\pm 0.0027) - 1902(\pm 1007)\mathrm{N/H}.$$

Our analysis suggests that the Y_p values coincide with the corresponding values from Izotov *et al.* [4, 15] within the 1σ errors. However, note that the values of dY/dZ, $dY/d(\mathrm{O/H})$, and $dY/d(\mathrm{N/H})$ become negative if the recombination coefficients of Smits [9] are used. The same conclusion follows from the work of Ballantyne *et al.* [12].

4. CONCLUSIONS

We used new approximation relationships for the ionization correction factors ICF based on the calculated grid of PhM for HII regions in BCD galaxies of [2] to derive the chemical compositions of the nebular gas. When deriving the ICF expressions, we used for the first time ion abundances found by performing diagnostics of the integrated spectra for the PE model, $(A^{+i}/\mathrm{H}^+)_{\mathrm{diagn}}$, instead of the total ion abundances $(A^{+i}/\mathrm{H}^+)_{\mathrm{mod}}$. We tested the ICF expressions for their ability to reproduce the composition $(A/\mathrm{H})_{\mathrm{mod}}$ specified in the PhM. In most cases, these ICFs are able to reproduce the compositions of the nebular gas with accuracies $\leq 10\%$. We tested the ICF expressions used by a number of authors to find the compositions of nebular gas in similar ways (Table 1). For some expressions (e.g., T1, T3–T6, T11–T13, T15), the resulting composition accuracies are $\geq 10\%$; i.e., they are considerably poorer than for the new expressions [2]. Accordingly, we used the chemical compositions found using our own ICFs to derive Y–Z, Y–O/H, and Y–N/H relationships (Table 3). We analyzed these relationships in detail to determine Y_p and dY/dZ for the B72 and Sm96 recombination coefficients. We performed calculations using these recombination coefficients in order to verify the systematic difference of the Y_p values obtained in [4]. The corresponding helium ion abundances were taken from [4], but we used our ICF(He) expressions (1)–(3) to calculate the total helium abundance. This yielded for the B72 coefficients $Y_p = 0.2201 \pm 0.0040$ and $dY/dZ = 2.90 \pm 4.48$, and for the Sm96 coefficients $Y_p = 0.2440 \pm 0.0024$ and $dY/dZ = -4.02 \pm 2.46$. Hence, we obtained in both cases Y_p values close to those of [4]. However, the slope of dY/dZ for the Sm96 case is negative. This contradicts the standard theory of chemical evolution, according to which the helium abundance must increase with time. The origin of this negative slope is unclear. We plan in the future to improve these chemical compositions using an alternative and more accurate method—the construction of optimized PhM for HII regions in BCD galaxies. The results of this study should enable us to refine the primordial helium content Y_p, and to reject or confirm the negative dY/dZ slope.

ACKNOWLEDGMENTS

The authors are grateful to Yu.I. Izotov (Main Astronomical Observatory, National Academy of Sciences, Ukraine) for useful discussions.

REFERENCES

1. G. Stasińska, Astron. Astrophys., Suppl. Ser. **83**, 501 (1990).
2. V. V. Holovatyy and B. Ya. Melekh, Kinematika Fiz. Nebesnykh Tel (2002) (in press).
3. Yu. I. Izotov, T. X. Thuan, and V. A. Lipovetsky, Astrophys. J. **435**, 647 (1994).
4. Yu. I. Izotov, T. X. Thuan, and V. A. Lipovetsky, Astrophys. J., Suppl. Ser. **108**, 1 (1997).
5. Yu. I. Izotov, T. X. Thuan, and V. A. Lipovetsky, Astrophys. J. **445**, 108 (1995).
6. Yu. I. Izotov and T. X. Thuan, Astrophys. J. **497**, 227 (1998).
7. V. V. Holovatyy, R. E. Gershberg, Yu. F. Mal'kov, *et al.*, Izv. Krym. Astrofiz. Obs. **96**, 1 (1999).
8. M. Brocklehurst, Mon. Not. R. Astron. Soc. **153**, 211 (1972).
9. D. P. Smits, Mon. Not. R. Astron. Soc. **278**, 683 (1996).
10. J. Kingdon and G. J. Ferland, Astrophys. J. **442**, 714 (1995).
11. M.-H. Armour, D. R. Ballantyne, G. J. Ferland, *et al.*, Publ. Astron. Soc. Pac. **111**, 1251 (1999).

12. D. R. Ballantyne, G. J. Ferland, and P. G. Martin, Astrophys. J. **536**, 773 (2000).
13. G. J. Ferland, *Hazy: a Brief Introduction to Cloudy 94* (University of Kentucky, Physics Department Internal Report, 1999).
14. `http://www.pa.uky.edu/~gary/cloudy`.
15. Yu. I. Izotov and T. X. Thuan, Astrophys. J. **511**, 639 (1999).
16. M. Seaton, Mon. Not. R. Astron. Soc. **139**, 129 (1968).
17. P. Leisy and M. Donnefeld, Astron. Astrophys., Suppl. Ser. **116**, 95 (1996).
18. J. Lequeux, M. Peimbert, J. F. Rayo, *et al.*, Astron. Astrophys. **90**, 155 (1979).
19. B. Wilkes, G. Ferland, D. Hanest, *et al.*, Mon. Not. R. Astron. Soc. **197**, 1 (1981).
20. R. L. Kinsburg and M. J. Barlow, Mon. Not. R. Astron. Soc. **271**, 257 (1994).
21. M. Peimbert, V. Luridiana, and S. Torres-Peimbert, Rev. Mex. Astron. Astrofis. **31**, 147 (1995).
22. T. Barker, Astrophys. J. **267**, 630 (1983); **270**, 641 (1983).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 791–804. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 879–8938.
Original Russian Text Copyright © 2002 by Efremov.

The Nature of Peculiar Stellar Complexes

Yu. N. Efremov

Sternberg Astronomical Institute, Universitetskii pr. 13, Moscow, 119899 Russia

Received December 16, 2001; in final form, May 23, 2002

Abstract—The nature of stellar complexes with peculiar populations and morphologies is investigated. The existence in the LMC of complexes made up of isolated stars, on the one hand, and consisting exclusively of clusters, on the other hand, could be due to different turbulence patterns in the initial gaseous medium. Arc-shaped stellar complexes are unlikely to be the result of star formation in a gaseous shell swept up by a central source of pressure, and instead probably reflect the shape of a bow shock that develops when a sufficiently dense cloud is subject to dynamical pressure. A peculiar arc-shaped complex in NGC 6946, which contains a young, massive cluster, may be the result of an oblique infall of a high-velocity cloud onto a region of the gaseous disk of the Galaxy with a strong, regular magnetic field; the properties of this complex can be explained as the result of a collision of the resulting shocks. The arc-shaped complexes in the LMC were also probably produced by high-velocity clouds moving obliquely through the more tenuous gas of the LMC disk. A similar complex in NGC 300 may owe its origin to the effect produced on a dense cloud by the shock from an extremely powerful external explosion, whose stellar remnant may have survived as an X-ray source now located along the line of symmetry of the arc of the complex. The rareness of such structures can be explained by the narrow range of conditions under which they can develop. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

Stellar complexes constitute the highest elements in the hierarchical structuring of young stellar groupings and contain about 90% of all stellar associations and young clusters [1–3]. Most stellar complexes owe their origin to common and inevitable processes occurring in the gaseous disks of galaxies on all scales, first and foremost, to large-scale gravitational instability, which produces most rapidly large clouds that give birth to stellar complexes [4]. These are often called star-forming regions (or complexes), although star formation usually continues only in a small volume of the complex, and the age of its oldest stars can reach 100 Myr. In only 1% of all complexes does star formation occur throughout their entire volume (at least in Local Group galaxies); these are known as superassociations. The origins of the synchronous star formation in such complexes may differ from object to object. In any case, superassociations are peculiar complexes and do not represent an early stage in the development of ordinary complexes. See [5] for a brief analysis of this problem.

Complexes in irregular galaxies are distributed chaotically, whereas, in spiral galaxies, they are arranged in chains along the spiral arms and correspond to the long-known stellar clouds (which, however, were considered before our work to be random conglomerates of stars and clouds [6] and not physical groupings united by a common origin for their constituent objects). Complexes in spiral arms form under the action of gravitational instability that develops along the arm [7] or due to the Parker instability [8]. Both theories can explain the regular arrangement—with a spacing of 1–3 kpc—of star–gas complexes along the spiral arms observed in many galaxies, including our Milky Way [9]. The small amount of available data on the character of magnetic fields in galaxies with different spiral-arm structures suggest that the Parker instability plays the most important role.

The origin of isolated complexes lying outside spiral arms or at the peripheries of irregular galaxies may be associated with local effects. Data for complexes with characteristic arc-like shapes are presented in our previous paper [10], hereafter Paper I. Some complexes are distinguished by peculiarities in their compositions—they contain either clusters or stars exclusively. The current paper is concerned with possible mechanisms for the formation of complexes with peculiar shapes or compositions. Unfortunately, the observational data required by this type of study are currently available for only a small number of such complexes.

2. COMPLEXES OF CLUSTERS AND OF STARS

Our Galaxy and the LMC contain complexes in which there are unusually small or large numbers

1063-7729/02/4610-0791$22.00 © 2002 MAIK "Nauka/Interperiodica"

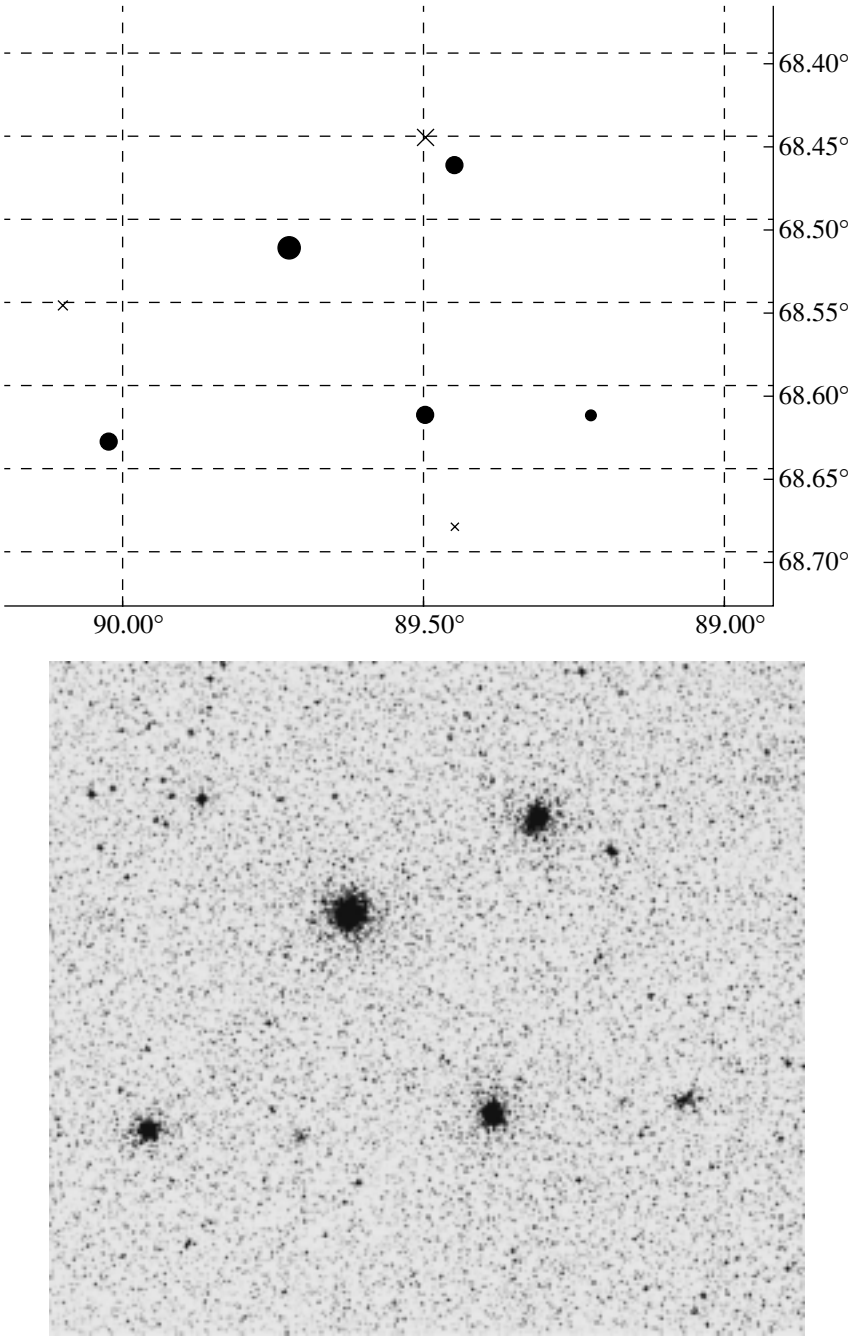

Fig. 1. A group of massive clusters around NGC 2164 in the LMC. Cepheids are shown by crosses. 1° corresponds to 900 pc.

of clusters relative to the number of isolated stars of the same age [1]. An especially striking example is a group of coeval clusters around NGC 2164 in the LMC that contains only three Cepheids (plus a few more inside clusters), although the age of this complex is optimal for it to contain Cepheids and the density of the region is not crowded (Fig. 1).

An objective comparison of the distributions of Cepheids and clusters of the same age in the LMC yielded three more groups of clusters; only one of the four groups of clusters coincided with a concentration of Cepheids [11]. The results of the OGLE program show that this group, which is located near the eastern tip of the bar around the massive clusters NGC 2058 and NGC 2064, contains about twenty smaller clusters and about 150 Cepheids, of which 20 are cluster members. Immediately to the southeast of this complex is a dense group of ~180 Cepheids of about the same size (about 200 × 300 pc), which, however, contains no conspicuous clusters (Fig. 2).

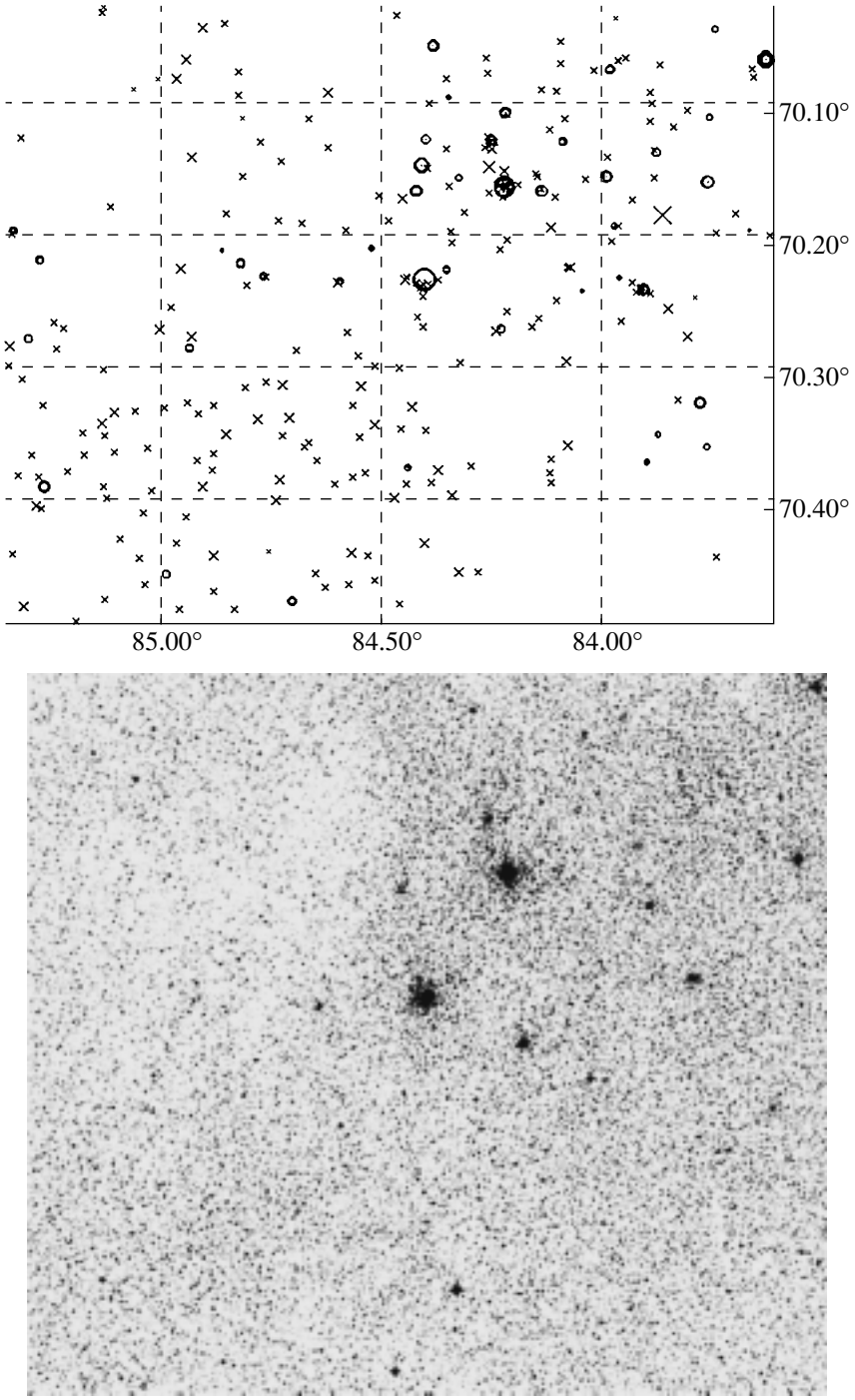

Fig. 2. Groups of clusters and Cepheids in the LMC. At the top right of the chart is a complex of Cepheids and clusters (circles) around NGC 2058 and NGC 2064, and at the bottom left is a complex consisting exclusively of Cepheids (crosses). The map is based on the results of the OGLE program.

The central ~ 0.1 kpc of this group of Cepheids is devoid of clusters, whereas the density of Cepheids is about 900 kpc^{-2}; i.e., two orders of magnitude higher than in the solar neighborhood. The periods (and, consequently, ages) of most of these Cepheids are contained within a narrow interval (3–5 days). It follows that this complex to the southeast of NGC 2058 is a relic of an unusual burst of star formation that failed to yield large, gravitationally bound clusters. A detailed analysis of this entire region of the LMC is critically important for our understanding of the factors determining whether both clusters and isolated

stars or exclusively isolated stars form. Note that the region located almost symmetrically with respect to this group of Cepheids relative to the bar axis contains the region of ongoing star formation 30 Dor, showing that this location is favorable for intense star formation.

Young, compact clusters resembling the rich, young clusters in the LMC and, in particular, the NGC 2164 group, are now known in many irregular and spiral galaxies. The discovery of such clusters in the latter galaxies is due, first and foremost, to the systematic search performed by Larsen and Richtler [12] in 21 spiral galaxies. The number of such clusters normalized to the luminosity of the galaxy varies over a wide range but is correlated with the star-formation rate of the galaxy. Larsen and Richtler [13] concluded that the formation of numerous compact clusters in interacting and starburst galaxies can be explained if the same mechanisms occurring in normal galaxies are operating, but under extreme conditions. Whitmore [14] supported this conclusion, and found that the relation between the star-formation rate and the number of young, massive clusters found in [13] can be extended to these galaxies.

However, this is not always the case, at least locally, as indicated by the existence of complexes of clusters that do not coincide with high densities of isolated stars. More substantial deviations from this relation are also known. For example, the cluster formation rate in the irregular Local Group galaxy IC 1613 normalized to the star-formation rate is a factor of 600 lower than in the LMC, which is a galaxy of the same type [15]. Similar discrepancies are found, not only for different locations, but also for different times in a single galaxy. The prolonged break in the formation of (at least massive) clusters in the LMC 4–14 Gyr ago was not accompanied by a decrease in the star-formation rate [16].

It appears that this situation can be understood in terms of the theory of star formation in turbulent media, which, in contrast to standard theories of star formation, ascribes the main role to the balance between gravity and turbulence and considers the role of magnetic fields to be negligible. In this theory, star formation is a rapid process that occurs over a time comparable to one crossing time of the initial molecular cloud, as is suggested by a number of observational facts. In particular, the age range inside a star-forming region increases as the square root of its size, pointing to turbulence as the determining factor in star formation [17–19, 20].

According to this theory, the rapid free collapse of a molecular cloud is prevented not by the magnetic field, but by turbulence, whose characteristics determine whether clusters or isolated stars predominantly form. The presence of gravitation alone results in the rapid formation of clusters. This can also occur in the presence of decaying or long-wavelength turbulence, whereas only isolated field stars form under the action of short-wavelength turbulence [20, 21].

This last circumstance is very important, since it may provide the best explanation for the existence of stellar complexes without clusters (without, however, explaining their high density), as well as the existence of isolated, young, massive stars in general. This has been conclusively demonstrated by Massey *et al.* [22, 23], who investigated high-luminosity stars in the field and associations of the LMC and found that half of these stars were field objects. Given their young age, this cannot be explained by their escape to great distances from their probable birth places in associations. The number of such stars also appears to be too large for them to have been ejected as a result of dynamical interactions of stars in the cores of dense young clusters. It is important that the mass function for field stars has a significantly steeper slope (indicative of a larger fraction of lower-mass stars) than the mass function for stars in associations [22]. A steep slope for massive stars was also obtained for the case of short-wavelength turbulence, which corresponds to isolated star formation [24].

Thus, the question arises of why the turbulence characteristics differ in different places and at different times. This may become a key problem in star formation. If our hypothesis is correct, the data on stellar complexes in the LMC imply that regions with different turbulence parameters have characteristic scales of several hundred parsecs. This is close to the sizes of gaseous filaments, voids, and clumps observed in the LMC [25]. The distribution of high-luminosity stars in the LMC does indeed resemble the network of gaseous filaments observed in HI, especially if we ignore star clusters.

The conclusions of Klessen *et al.* [21] and Massey *et al.* [22] obviously imply that, in the absence of support from turbulence, the collapse of a gaseous protocluster and the star formation in it are so rapid that the cluster remains gravitationally bound even if it contains a large number of massive stars. Klessen *et al.* [21] find that, at high densities, gas collapses into dense cores over several free-fall times, and the star-formation efficiency exceeds 50%, so that differences in the magnitude and pattern of turbulence may be sufficient to explain the preferential formation of either isolated stars or stars in clusters.

It is quite possible that currently observed old, classical globular clusters also formed in the absence of support due to turbulence in the initial cloud. Phinney [26] concluded that the large number of neutron stars and massive white dwarfs (i.e., remnants of massive stars) in globular clusters indicates that most of the gas in the protocluster produced essentially all

the massive stars very rapidly—over a single crossing time—before even one percent of the massive stars began to affect the surrounding gas as supernovas or O stars. This means that the formation time for massive, compact clusters is, indeed, short, close to the free-fall time, which is equal to 1 Myr for half the globular-cluster radius. Arguments in favor of the rapid formation of stars in clusters and a small role for magnetic fields can also be found in [27].

It is thus possible that not only the high external pressure but also the very rapid collapse of a protocluster in the absence of turbulence can explain the formation of massive, gravitationally bound clusters. This implies that turbulence can sometimes be absent from fairly large regions; 300×300 pc for the cluster of clusters around NGC 2164. The theory of rapid star formation determined by the turbulence conditions does not require the regeneration of turbulence to prevent the collapse of molecular clouds, which can be transient formations [19].

There is apparently no need to look for any special origins of the existence of regions without turbulent flows; the existence of alternating patterns of cells with laminar and turbulent flows is a well-known phenomenon, and Chernin [28] recently suggested that this stochastic phenomenon may explain the small dispersion of galaxy velocities in the Hubble expansion of the Universe. It is important to try to elucidate the scale lengths and time scales expected for this phenomenon in the gaseous disks of galaxies. We conclude that the stellar complexes outside spiral arms may correspond to the largest cells of such turbulence patterns in the interstellar gas (see also [17]).

3. THE FORMATION OF ARCLIKE COMPLEXES FROM GASEOUS SUPERSHELLS

We now turn from complexes with peculiar compositions to those with peculiar shapes, considered in Paper I. The characteristic shape of arclike stellar complexes immediately suggested that they were formed by the action of some central source of pressure that produced an expanding shell of swept-up gas. Many authors (see, e.g., [29]) have investigated the possibility that a supershell of gas swept up by a central source of pressure from a fairly dense medium fragmented into clouds that gave birth to star clusters. This possibility was considered by Vader and Chaboyer [30] and Efremov and Elmegreen [31] for a giant arc of clusters in NGC 1620 and two arcs in the northeastern part of the LMC, respectively. Thus, the formation of giant arcs of stars and clusters appeared to be a special case of the origin of HI supershells. Precisely this approach was adopted by Tenorio-Tagle [32].

The energy of a central source of pressure capable of producing a supershell of swept-up gas with a size of about 1 kpc is greater than the energy of a single supernova by a factor of several tens or even hundreds. One proposed source of such energies is successive supernova explosions in a sufficiently rich cluster. However, in most cases, the supershell contains no central cluster. In some cases, this can be explained by the fact that the supershell is located in a region with a low degree of differentiality of the galactic rotation and/or a large thickness of the gaseous disk. In this case, the supershell has a long lifetime, and its age may be so large that the old central progenitor cluster has become inconspicuous [31, 33]. However, it is difficult to determine the parameters of a possible progenitor cluster from the age and size of the supershell in order to test this "standard" model for the formation of supershells.

In fact, such a verification was performed just recently. Rhode *et al.* [34] carried out a careful search for clusters inside HI supershells in the irregular galaxy Ho II. They found that only 6 of 44 supershells contained clusters with numbers of stars and ages consistent with the hypothesis that they, at one time, contained massive stars in numbers sufficient to produce the supershells. There are no clusters inside the largest supershells in Ho II, which, in addition (like the supershells in our Galaxy), are located at the periphery of the galaxy, where young massive stars are absent or very rare.

In the absence of central clusters, it has been proposed that expanding supershells formed as a result of the impact of a high-velocity gaseous cloud into the galactic plane. This hypothesis was initially suggested to explain stellar "super rings" in the LMC [32]. However, it also faces certain problems; for example, many galaxies with HI supershells have no such clouds in their vicinity. Other possibilities have also been considered. One idea that is worth investigating is that a supershell with subsequent triggered star formation could form around the region of intersection of a gaseous galactic plane and a sufficiently massive (globular), high-velocity cluster [35].

One has the impression that, near the center of a stellar arc, there is either a very massive cluster or no clusters at all. It is possible that only in very massive (of the order of $10^6\ M_\odot$) clusters are supernova explosions frequent enough for supershells to form. Efremov *et al.* [36] point out that the mean rate of energy input from 1000 supernovas over 2×10^7 yr implies a fairly low rate of heating of the interstellar medium, close to the rate at which it is cooled at normal pressures, so that the formation of a supershell, indeed, requires very massive clusters.

Such massive, and rather young, clusters are indeed observed. The molecular hydrogen supershell found near the center of M82 has a radius of 450 pc, kinetic energy of about 10^{55} erg, and age of $(1-2) \times 10^6$ yr. Its formation may have required 10^3-10^4 supernova explosions over this time, implying a central cluster mass of $2 \times 10^6\ M_\odot$, in agreement with observational estimates of the number of red supergiants in this cluster [37]. Several supershells around giant young clusters in NGC 5253 were studied in [38]. An older large cluster can be found in a giant arc of young clusters (3 kpc in diameter) in NGC 1620 [30]; no neutral hydrogen data are available for this galaxy. An HI supershell is strikingly absent around a cluster of mass $10^6\ M_\odot$ with an age of 12 Myr in NGC 6946 (see Section 5).

The largest HI void in the irregular galaxy IC 2574 coincides with a complex made up of about ten young clusters, with an older cluster with a mass of $1.4 \times 10^5\ M_\odot$ located near its center [33]. This is the only known case where the available observations, including those suggesting the presence of hot gas inside the supershell, agree quite well with the theory that the formation of the HI supershell and then of new clusters was driven by the energy of stars of an older cluster. The ages of the central cluster and of the supershell are also in agreement. However, the energy required to create the swept-up shell (2.6×10^{53} erg) is an order of magnitude higher than that provided by the supernovas and stellar wind of the central cluster, leading Stewart and Walter [33] to conclude that this model overestimates the energy required for the formation of the supershell. Indeed, the model of Ehlerova and Palous [40] yields an energy an order of magnitude lower for this shell.

However, it is possible that the energy provided by stars was, in fact, higher than follows directly from estimates of the mass and age of the cluster assuming a normal luminosity function. Rich clusters should also contain extremely massive stars, which can explode as hypernovas, releasing energies an order of magnitude higher than provided by ordinary supernovas. Very massive and rapidly rotating stars can also form as a result of mergers of stars in dense cores during a short but recurrent stage of gravothermal core collapse, even in clusters that are not very massive (see [41] and references therein).

Finally, if a galaxy moves through a sufficiently dense intergalactic medium, initially small cavities in its gaseous disk could grow as a result of ram pressure. This idea was suggested to explain the existence of numerous cavities in the galaxy Ho II, whose outer density contours are bounded on one side by a regular circular arc, indicating that the galaxy is moving through the intergalactic gas of the M81 group [42].

If supershells form due to the action of a large number of supernovas and hot O stars on the interstellar medium, why are no supershells observed aound many clusters that have undoubtedly contained such objects, judging by the number of stars now observed in these clusters and the ages of these stars? This could be explained by the presence of a very high density of neutral and/or molecular hydrogen around such clusters, which is unlikely in most cases. It is most likely that the energy required to produce a supershell in a medium of normal (and especially of high) density can be provided only by supernova explosions that are so frequent and numerous that they could occur only in the most massive clusters. However, hypernovas (which may be associated with gamma-ray bursts) can occasionally appear in almost any cluster. It is very important also to address the inverse problem—to search not for clusters inside supershells, but for supershells around clusters and, based on the masses and ages of the clusters and the parameters of the ambient medium, to try to explain the absence of HI voids around many clusters that could have produced them in the framework of the standard model.

The problem of the origin of supershells has a long history. Heiles [43], who was the first to discover about a dozen supershells in our Galaxy, pointed out that they could have been produced by a large number of type-II supernovas exploding in OB associations, though "their [the supershells'] lack of association with extreme Population I objects argues strongly against this possibility." He even assumed that "...it is extremely improbable that we have ever directly observed the agent responsible for their [the supershells'] existence. The agent may be itself a new unknown kind of astronomical object" [43, p. 544]. Heiles may have well been right. Gamma-ray bursts were suggested to be such objects, capable of providing the interstellar medium with energies sufficient to produce supershells.

4. STELLAR ARCS IN THE LMC AND GAMMA-RAY BURSTS

The very existence of a multiple system of giant stellar arcs in the LMC (Fig. 3) means that they could not have formed in a supershell swept up by supernovas and O stars that existed in central clusters [31]. We can suppose that these original clusters are now barely discernible, but why are all arcs in the LMC located within 1 kpc of each other? The possibility that these arcs were formed as a result of events associated with gamma-ray bursts whose progenitors had a common origin somewhere nearby cannot be ruled out. This supposes that gamma-ray bursts are produced during mergers of binary systems consisting of

compact objects, which form primarily as a result of dynamical interactions between compact objects and stars in the dense cores of star clusters (see [47] and references therein).

The number of binary systems made up of compact objects formed during stellar encounters in dense clusters may be much higher than previously believed. Taking into account the unequal masses of stars in clusters and the presence of initial binary systems, it turns out that the dynamical interaction of stars in clusters is much more efficient than thought earlier. It is due to this incorrect notion of the low probability of this mechanism that it has largely been neglected when estimating the time between the formation of a pair and the subsequent merger of its components, as well as the rate of these events. On the contrary, the preservation of an initial binary after both the primary and secondary components explode as supernovas, resulting in the formation of a pair of compact objects, is a fairly artificial situation, requiring special assumptions about the character of the explosions. The known spatial velocities for three of the four binary neutron stars in the Galaxy are significantly lower than the velocities of single pulsars, which are believed to have acquired relatively high velocities as a result of asymmetric explosions of their progenitor supernovas (see references in [47]).

Recent results of numerical computations support the old analytical conclusion that compact-object pairs formed in a cluster become closer as a result of subsequent stellar encounters, acquiring increasingly higher velocities in the process, and are eventually ejected from the cluster. Such systems escape from the cluster when they are already quite close, and can merge during a gamma-ray burst much more rapidly than can initial systems of compact objects; this merging time is only a few million years for black-hole binaries escaping from the most massive clusters (of the order of $10^6\ M_\odot$) (see references in [47]).

A pair of compact objects that have escaped from a cluster can have a wide range of velocities, depending on the parameters for its last encounter, but are mostly close to the minimum escape velocity (i.e., they exceed the velocity dispersion V in the cluster by a factor of a few). Dokuchaev *et al.* [48] derived for the velocity V_{ej} for the ejection of a superelastic binary from a cluster the formula $V_{ej} = 724(V/100)$ km/s, which, for globular clusters, yields velocities of about 20–70 km/s. These velocities and merging times mean that a gamma-ray burst accompanied by the release of energy exceeding that characteristic of supernovas by one to three orders of magnitude can be observed several kiloparsecs from the cluster from which its parent pair of compact objects escaped.

This is precisely how we explained the presence in the LMC of several giant stellar arcs at distances of up to 1 kpc from the massive cluster NGC 1978 but not around it ([47] and references therein). The fact that objects related to the progenitors of gamma-ray bursts are concentrated in this cluster provides further evidence suggesting that this very massive cluster could had been a source of gamma-ray burst progenitors. The only soft gamma-ray repeater identified with an object outside the Galaxy (the supernova remnant N49)—SGR 0526–66—is located eighteen arcmin from the cluster; this is also a region where there is a concentration of X-ray binaries, which could be relatives of the progenitors of gamma-ray bursts [49] or the stellar remnants of powerful supernovas.

The shapes and sizes of stellar arcs may be due to the protracted action on the interstellar medium of multiple precessing jets (as for the HI shell surrounding SS 433) or a bow shock from narrow jets whose working surfaces have been increased by various instabilities [41, 47, 50].

If giant stellar arcs are produced by energy releases associated with gamma-ray bursts, this has important implications concerning the nature of these bursts. Data suggesting that these bursts occur in star-forming regions are believed to be incompatible with the hypothesis that the progenitors of gamma-ray bursts are systems of compact objects, since some time must pass before the initial pairs merge, during which they must move several tens of parsecs away from their birth places [51].

However, this problem does not arise if the pairs of compact objects that produce gamma-ray bursts are born during close encounters of stars in the dense cores of clusters. Recurrent gamma-ray bursts can occur in the vicinity of the parent cluster, and are capable of producing star-forming regions near which subsequent bursts occur. Recently, some observational evidence supporting this scenario has been found. The nearest gamma-ray burst, GRB 980425, occurred in a star cluster inside a star-forming region and near which there is an arc-shaped feature whose center of curvature is situated near the cluster [52]. This is the only gamma-ray burst close enough to the Earth to enable such features to be detected. Similarly, stellar arcs are observed nearby SGR 0526–66—the only known extragalactic object of this class—in the same region of the LMC.

Recently, Tsvetkov *et al.* [53] found that, contrary to the conclusions of Bloom *et al.* [51], the distribution of galactocentric distances of gamma-ray bursts does not resemble the distributions of either star-forming regions or supernova remnants. In our view, it resembles the distribution of classical, old globular clusters, and exhibits a well-defined concentration

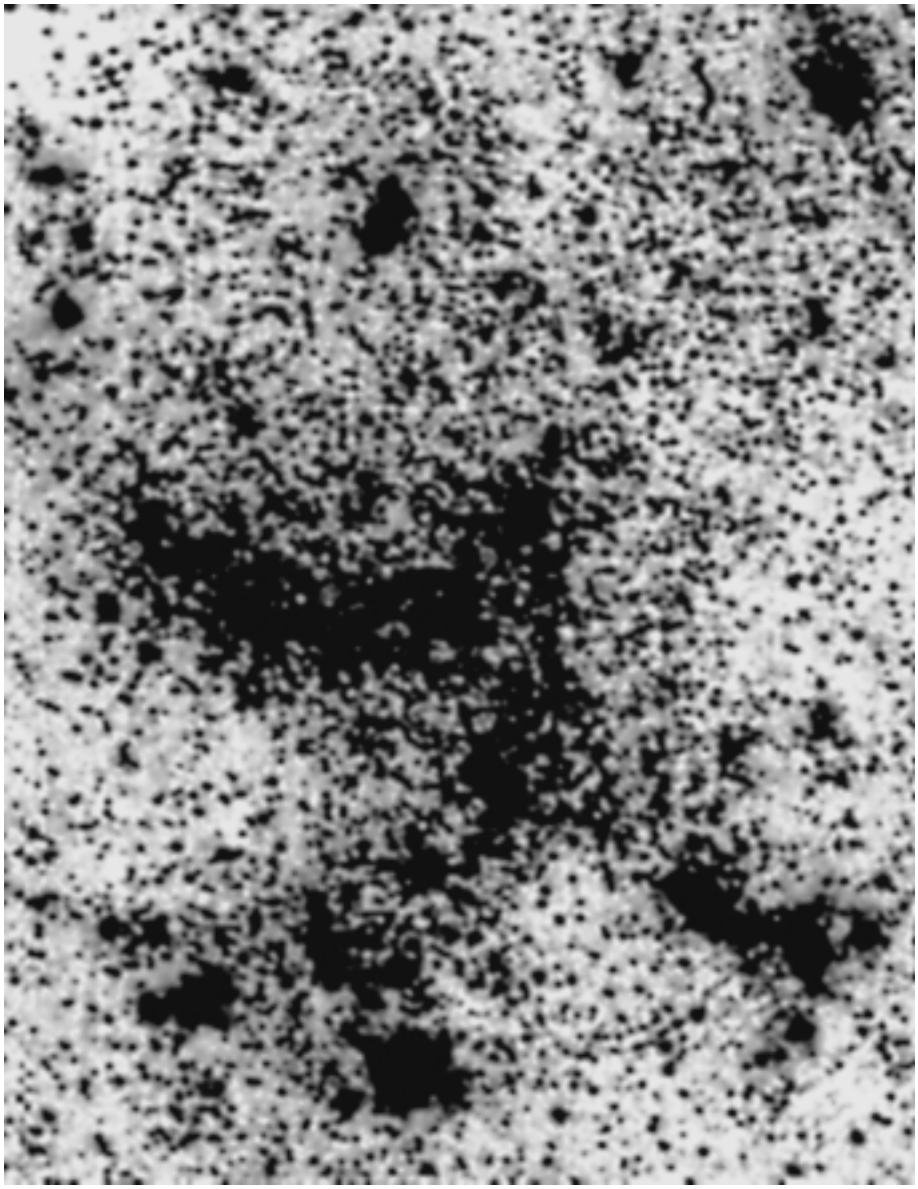

Fig. 3. Giant arcs of young stars and clusters in the LMC. The Quadrant (at the center) and Sextant (bottom right) have radii of curvature of about 300 pc and 200 pc, respectively.

toward the center of the "combined" galaxy [41]. It is also consistent with the birth of gamma-ray burst progenitors in dense galactic nuclei [48], if these nuclei eject pairs of compact objects that are already close. However, gamma-ray bursts are also observed in irregular galaxies without nuclei, implying that this scenario is by no means the only one possible.

Curiously, the distribution of gamma-ray burst redshifts indicates that most observed bursts occurred 8–12 billion years ago, when classical globular clusters had ages close to that of NGC 1978, which is about 2 Gyr old. This may imply that massive clusters become efficient sources of the objects that produce gamma-ray bursts at about this age [41]. This stage is not, however, very long, since present-day classical globular clusters (with ages of 12–14 Gyr) are not obviously associated with gamma-ray bursts.

5. HODGE'S COMPLEX IN NGC 6946

The history of the discovery of Hodge's complex and its main properties are described in Paper I. HST observations revealed this complex to have a high density of high-luminosity stars and to contain about 20 rich young clusters, suggesting it should be considered a relic region of violent star formation [46]. The unique feature of this complex is its sharp western boundary, which spans a 130-degree arc of a wonderfully regular circle (Fig. 4). The complex also contains

a giant cluster with an age of 12–15 Myr, which has a mass of $10^6 M_\odot$ and is gravitationally bound (if we assume a normal luminosity function) [54]. The galaxy, and especially its western half, contains many high-velocity clouds and HI voids. Such objects can also be found in the vicinity of Hodge's complex, but the failure of the complex to coincide with any of the HI voids is striking. NGC 6946 is surrounded by a group of eight late-type dwarf galaxies, essentially all of which were detected in the HI line [39], so that they could be sources of gaseous clouds if there are appreciable tidal interactions with the main galaxy. This already suggests the impact of a gaseous cloud into the plane of the galaxy as a likely origin for the complex.

Long-slit Hα spectroscopy performed at the 6-m telescope of the Special Astrophysical Observatory and the Keck-I telescope showed that the radial velocity of the main cluster of the complex is 150 km/s, which is 20–30 km/s higher than the local velocity of rotation measured from HII observations [55]. Since the plane of the galaxy makes an angle of 34° to the plane of the sky, this difference appears to imply for the cluster a velocity deviation from the circular rotation velocity that is unacceptably large for young objects. This may suggest that the cluster and surrounding complex are bodies foreign to NGC 6946.

The cross sections of the complex and its surroundings made at three slit positions reveal velocity-field perturbations that are especially large east of the complex. Some of these perturbations can be interpreted as rapidly expanding shells of ionized gas (see Fig. 7 in [55]). The data also exhibit some evidence for a velocity gradient, consistent with the possible rotation of the complex about an axis in the plane of the galaxy.

The galaxy NGC 6946 possesses a magnetic field that is regular outside the bright spiral arms (see [56] and references therein). Santillan *et al.* [57] simulated the impact of a high-velocity cloud with a galaxy possessing such a field. According to this study, at certain angles between the cloud trajectory, the galactic plane, and the force lines of the magnetic field, the field prevents the falling cloud from penetrating into the galactic disk, possibly explaining the absence of (at least a large) HI void. An oblique impact produces a complex pattern of magnetohydrodynamical waves, resulting in the development of a peculiar structure for the complex. Chernin *et al.* [58] showed that the interaction between shocks can produce a gas and dust lane at the collision front, a collimated wave propagating in the direction of decreasing density, and the presence of two generations of stars whose ages are separated by 20–30 Myr. All these features are, indeed, observed in Hodge's complex (see [46], Fig. 4).

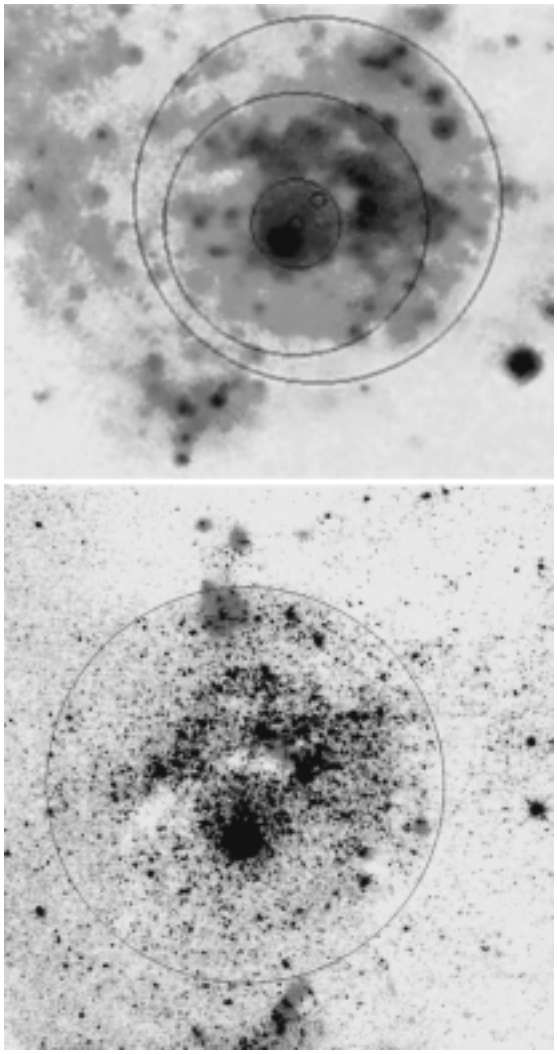

Fig. 4. The peculiar stellar complex in the galaxy NGC 6946 discovered by Hodge in 1967. Its diameter is about 600 pc. The top panel shows an image obtained by S. Larsen at the NOT telescope. North is to the top and west to the right. The centers of the two concentric circles coincide with the centers of the concentration of older stars and of the arc of young stars (see Fig. 9 in [46]). The outer circle outlines the western boundary of the complex, and its center (shown by the small circle) is shifted to the northwest. The arc of young clusters and HII regions located outside the complex and to the southeast may be part of the same structure. The lower panel shows an HST image. North is to the upper left and east to the bottom right. The complex is parted in the middle by two extended dark clouds. The bright object right of the border of the figure is a Milky Way star or a very compact cluster in NGC 6946. There is also a weak X-ray source nearby.

Moreover, the oblique-cloud impacts computed by Bloom *et al.* [51] predict the development of a bow shock and vortex motions in the oscillating tail of the cloud, followed by the development of the Parker

instability. The sharp arc-shaped western edge of the complex (Figs. 4, 5) leads us to conclude that the cloud moved along a slanted trajectory from east to west, and that this might be why the most conspicuous radial-velocity perturbations are observed to the east of the complex. The Parker instability could be responsible for the formation of the young, giant, gravitationally bound cluster. However, this cluster could also be another result of collisions between shock waves.

Note that Chernin *et al*. [58] analyzed only a collision of two shocks, whereas the impact of a cloud into a magnetized gaseous disk obviously should produce a much more complicated pattern. A more detailed analysis of the spatial, temporal, and kinematic structure of the complex should make it possible to reconstruct the characteristics of the parent event.

6. WHY DO STELLAR SUPERARCS HAVE THE SHAPES OF REGULAR CIRCULAR FRAGMENTS?

The most striking feature of the arc-shaped stellar complexes in the LMC and NGC 6946 is that they themselves or large portions of their boundaries form fragments of regular circles (Fig. 5). We believe this to be the key property for understanding the origin of these structures. The most simple explanation would be that the circular shapes of these objects seen in projection reflect the fact that these structures are actually segments of spherical layers (in the case of the Quadrant and Sextant in the LMC) or a segment of a filled sphere (in the case of Hodge's complex in NGC 6946) seen from the side. Figure 3 in Paper I shows how this model fits the arc of the Quadrant. The inclinations to the plane of the sky of both the LMC and NGC 6946 are close to about 30°, so that circles lying in the planes of these galaxies would be projected into ellipses easily distinguishable from the observed circular arcs.

In a model with an initial expanding swept-up gaseous shell, the spherical boundary of the resulting stellar complex obviously implies that the star formation had mostly ended before the shell broke out of the gaseous disk of the galaxy (i.e., when its radius was less than half its thickness), after which the stellar sphere continued to expand with the velocity it had at the time of the star formation. Assuming equal expansion velocities for the stellar arcs of the Quadrant and Sextant (this assumption has recently been theoretically justified for star formation in a swept-up shell [40]), we found this velocity to be 12 km/s (Paper I), implying that the stellar shell 300 pc in radius and 15 Myr old (the Quadrant) had a radius of 120 pc at the time of star formation (130 pc for the Sextant). This value is close to the half-thickness of the gaseous disk of the LMC (see Paper I).

One way or another, only two clear cases are known where there is an older cluster near the center of a complex of clusters forming an HI supershell that could have been responsible for the formation of the entire structure. It is also important that the young clusters in the complex IC 2574 are arranged in an irregular way, and its shape corresponds to the elliptical shape of the HI void, which, unlike our stellar arcs, corresponds to a circle in the plane of the galactic disk, taking into account its inclination to the plane of the sky. The giant arc in NGC 1620, which, judging from images of the galaxy, may even be a spiral-arm fragment, also has an irregular shape. Both these features have little in common with the regular arcs of the Quadrant and Sextant in the LMC or the western edge of Hodge's complex in NGC 6946.

If the source of energy for the central pressure forming the swept-up shell was stars of a hypothetical central cluster or the impact of a dense cloud, only a very specific local density distribution would allow the resulting shell to acquire the shape of a regular circular segment. A segment of a spherical surface could also form if it was produced as a result of an explosion outside the plane, which suggests the explosion of an object associated with a gamma-ray burst; however, the orientation of the arcs of the Quadrant and Sextant is inconsistent with this picture [44].

The regular circular shapes of Hodge's complex and of the arc-shaped complexes in the LMC can also be understood if these complexes are flat circular disks inclined to the plane of the galaxy, with the inclination of the complex exactly compensating the inclination of the galaxy to the plane of the sky, so that the projection of the complex is transformed from an ellipse back into a circle. This seems artificial; however, the plane of the Gould belt is inclined to the plane of the Galaxy (by 18°), and the possible trajectories of the inclined impacts of the clouds into the disks of NGC 6946 and the LMC are consistent with this suggestion, if the planes of these complexes are rotated about an axis whose orientation is close to that of the cloud's trajectory. In this case, the filled interior of Hodge's complex may a result of the magnetic field in the disk of NGC 6946.

7. DYNAMICAL PRESSURE AND STAR FORMATION

There is another, apparently more likely, explanation for the regular and circular projected boundaries of these complexes. It is well known that the gaseous corona of a galaxy moving through a sufficiently dense medium is subject to dynamical pressure and acquires

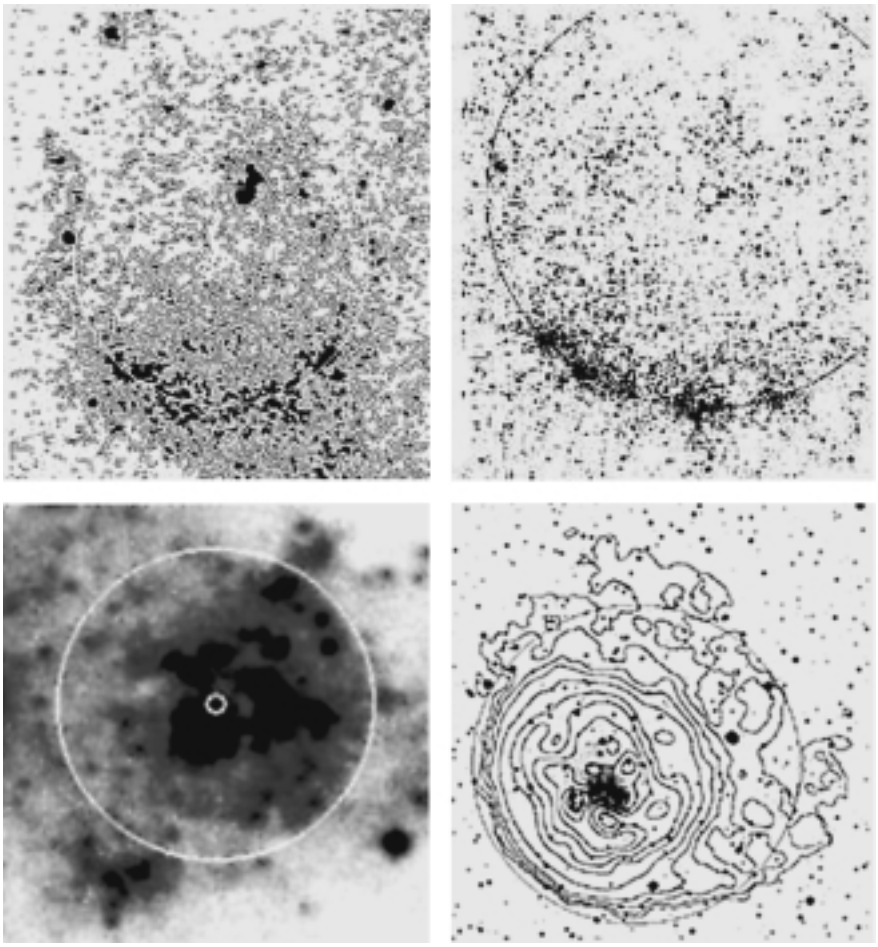

Fig. 5. Images of the Quadrant (top, left), the Sextant (top, right), Hodge's complex in NGC 6946 (bottom, left), and galaxy Ho II with HI-density contours [42] superimposed (bottom, right). Also shown in all images are circular shells. The centers of curvature for Hodge's complex and the Sextant are shown by small circles.

a characteristic comet-like shape with a sharp half-circular boundary at its leading side. For example, the outer HI-density contours of the galaxies Ho II [43] and NGC 7421 [46] have a parabolic shape on one side, which coincides over $\sim 100°$ with a circular arc (Fig. 5), in what appears to be an evident signature of a bow shock that has developed due to the galaxy's motion through the dense intergalactic gas of the corresponding group of galaxies.

The regular circular arc that bounds the stellar disk on one side of DDO 165 (Paper I) obviously owes its existence to the fact that the star formation in this galaxy was to a significant degree determined by the dynamical pressure of the intergalactic gas. The sharp, bright, half-circular western edges of the stellar disks of the galaxies NGC 7421 and NGC 2276 are also clearly due to bow-shock triggered star formation. In the case of the NGC 2276, this is further confirmed by the characteristic size of its disk, which is outlined by ionized hydrogen [59]. Moreover, the western edge of NGC 7421 is outlined by three straight intervals inscribed into a half circle. The same appears to be true of Hodge's complex, although with less certainty. This morphology is clear evidence for the action of a shock wave [60].

Thus, as we already suspected [61], the direct observational data lead us to conclude that arc-shaped or circular-arc-bounded stellar complexes may be products of star formation triggered by the action of a one-sided external pressure on the initial gaseous cloud. The hypothesis of dynamical pressure acting on a cloud arising when the cloud moves through a lower-density medium can be used to explain both the Quadrant and Sextant arcs, as well as the western boundary of Hodge's complex. In all three cases, the initial cloud moves through the gaseous disk of the galaxy at a small angle to the disk plane, and we view the resulting complex from the side. The concentration of stellar arcs in the northeastern part of the LMC may reflect the fact that this is the leading side of the galaxy in terms of its orbital motion. The orientation of the Quadrant and Sextant arcs is consistent with

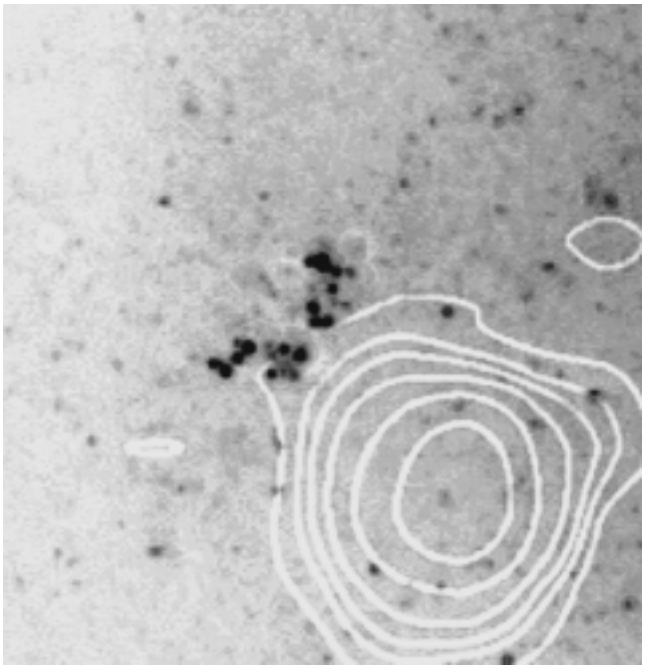

Fig. 6. The arc-shaped complex AS 102 in NGC 300 (NOT image taken by S. Larsen) and X-ray flux contours from [64] indicating the position of the P42 X-ray binary.

this possibility [61]. Note also that, at least in the case of the Quadrant, the dense leading arc is followed by a tail of lower stellar density, and the entire structure is bounded by a parabola (see Fig. 1 in Paper I), as should be the case for structures formed under the action of dynamical pressure.

Note also that star formation in a bow shock can explain the regular shapes of arc-like complexes, whereas a complex of clusters formed in a swept-up gaseous shell will have an irregular shape, as observed in IC 2574. Moreover, only the formation of arcs as a result of external pressure on a dense cloud can explain why all such complexes have approximately the same opening angle (about 100°; see Fig. 5). The available theoretical data show that only a narrow range of conditions leads to triggered star formation in a collision of clouds [62], especially if we require that the resulting complex should maintain the shape of the bow shock, naturally explaining the rarity of arc-shaped stellar complexes. The presence of a magnetic field may constitute one such condition [63], as is clearly relevant for the complex in NGC 6946.

A physically similar situation arises when a sufficiently dense cloud is subjected to a shock produced by a nearby and powerful explosion propagating in the more tenuous gas of the galactic disk. Obviously, the convex side of the interaction front between the shock and cloud must be oriented toward the explosion, and the resulting arc-shaped stellar complex must be symmetric with respect to the direction toward the source of the shock. In one case, this interpretation appears to be natural. In NGC 300, an arc of bright stars with a chord size of about $45''$ (~ 400 pc) is located near the most intense X-ray point source in that galaxy, with its convex side pointed toward this source (Fig. 6). This is the object P42 = H13, which is classified as an X-ray binary with a black hole (with a mass of about 5 $M_\odot$) and is the only such object in NGC 300 [64]. It may well be the stellar remnant of a hypernova.

This arc is listed in a catalog of OB associations and complexes in NGC 300 as AS102 [65]; it is classified as a stellar complex, because, as is evident from Fig. 6, it consists of four subgroups. The age of this complex is about 5 Myr [66], and it is immersed in a bright HII region, like the Sextant arc. Over this time, the optical or radio gaseous remnant of the supernova has disappeared, but the stellar remnant—black hole accreting matter from the secondary—could, under certain conditions, be detected as an X-ray source.

Figure 6, which presents a NOT image obtained by S. Larsen with an X-ray map [64] superimposed, shows that this unique X-ray source in NGC 300 is located not only near the complex, but exactly on its symmetry axis. Unfortunately, the available HI data for NGC 300 have a resolution too low to enable identification of HI features in this region. If our hypothesis is correct, there should be no gas between the complex and the object P42. The rareness of such

structures may be due to the fact that their formation requires a rather large and dense cloud, as well as the appropriate orientation.

8. CONCLUSIONS

The available data on stellar complexes with peculiar compositions agree with the idea that turbulence plays the decisive role in the process of star formation. It is possible that the complexes containing only star clusters formed under conditions of no turbulence or disruptive turbulence and that the sizes of stellar complexes identified outside the spiral arms of galaxies are determined by the characteristic sizes of cells with the same turbulence pattern.

We have considered regular circular arcs bounding several stellar complexes and analyzed evidence for the past existence of a source of pressure at their centers, whose presence would suggest that these clusters formed from a gaseous shell swept up by this pressure. However, they are more likely to have formed as a result of oblique impacts of intergalactic clouds, in which case they reflect the shape of the bow shock that develops due to the dynamical pressure as the clouds move through the more tenuous medium of the gaseous disk of the galaxy.

It appears that only a rare combination of conditions (velocity and angle of incidence, pressure ratio, the presence of a magnetic field) leads to star formation in the bow shock near the leading surface of the moving cloud. This is precisely what the theory predicts [62, 63], and this may also explain the rarity of such structures. The large and similar opening angles at the centers of all these arcs also support the idea that they formed as a result of a bow shock that developed in the presence of speeds and densities close to those observed in gaseous clouds and galactic disks. Hodge's complex in NGC 6946 is distinguished by a high density of both individual stars and clusters, as well as its complex structure, with signatures of several shells of stars and gas. This structure is probably the result of a collision of shocks produced by the oblique impact of a gaseous cloud onto the magnetized gas of the galactic disk. The interaction of shocks may also have produced high pressures in the gas, favoring the formation of a massive, gravitationally bound cluster.

A physically similar situation may also develop when a dense cloud is subject to a shock produced by a fairly nearby and powerful explosion, whose stellar remnant is still visible in the case of the arc-shaped complex in NGC 300. However, we cannot rule out the possibility that the Quadrant and Sextant arcs in the LMC formed as a result of explosions of objects associated with gamma-ray bursts, which would not necessarily have been located at the center of curvature of these arcs. Studies of peculiarly shaped stellar complexes and of complexes containing either clusters or stars exclusively appears to be a very efficient tool for comparing theories of the formation of clusters and isolated stars with observations.

9. ACKNOWLEDGMENTS

I am grateful to A.D. Chernin, E. Alfaro, B. Elmegreen, F. Combes, S. Larsen, and M.E. Prokhorov for useful discussion of the problems considered in this paper, and to S. Larsen for sharing the images taken with NOT. I also thank the referee, M.A. Shchekinov, for his valuable comments. The charts of stellar complexes were drawn using the convenient AstroView program developed by E.Yu. Efremov. Bibliographic data contained in the NASA Astrophysics Data System facilitated the work to a great degree. I also acknowledge with gratitude the use of OGLE data and the images obtained with the HST NASA/EAS as part of a program in which the author was a participant. This work was supported by the Russian Foundation for Basic Research (project no. 00-02-17804) and the Council for the Support of Leading Research Schools (grant no. 00-15-9662).

REFERENCES

1. Yu. N. Efremov, *Sites of Star Formation* [in Russian] (Nauka, Moscow, 1989).
2. Yu. N. Efremov, Astron. J. **110**, 2757 (1995).
3. B. Elmegreen, Yu. Efremov, R. Pudritz, and H. Zinnecker, in *Protostars and Protoplanets IV,* Ed. by V. Mannings, A. P. Boss, and S. S. Russell (Univ. of Arizona Press, Tucson, 2000), p. 179.
4. B. Elmegreen, Astrophys. J. **312**, 626 (1987).
5. Yu. N. Efremov, *Violent Star Formation,* Ed. by G. Tenorio-Tagle (Cambridge Univ. Press, Cambridge, 1994), p. 61.
6. P. N. Kholopov, *Star Clusters* [in Russian] (Nauka, Moscow, 1981), p. 421.
7. B. G. Elmegreen, Astrophys. J. **433**, 39 (1994).
8. J. Franco, J. Kim, E. Alfaro, and S. S. Hong, astro-ph/0111406 (2001).
9. Yu. N. Efremov, Astron. Astrophys. Trans. **15**, 3 (1998).
10. Yu. N. Efremov, Astron. Zh. **78**, 887 (2001) [Astron. Rep. **45**, 769 (2001)].
11. P. Battinelli and Yu. N. Efremov, Astron. Astrophys. **346**, 778 (1999).
12. S. Larsen and T. Richtler, Astron. Astrophys. **345**, 59 (1999).
13. S. Larsen and T. Richtler, Astron. Astrophys. **354**, 836 (2000).
14. B. C. Whitmore, astro-ph/0012546 (2000).
15. T. K. Wider, P. W. Hodge, and A. Cole, Publ. Astron. Soc. Pac. **112**, 594 (2000).

16. S. van den Bergh, in *New Views of the Magellanic Clouds (IAU Symposium 190),* Ed. by You-Hua Chu, N. Suntzeff, J. Hesser, and D. Bohlender, p. 573.
17. B. G. Elmegreen and Yu. N. Efremov, Astrophys. J. **466**, 802 (1996).
18. Yu. N. Efremov and B. G. Elmegreen, Mon. Not. R. Astron. Soc. **299**, 588 (1998).
19. B. Elmegreen, Astrophys. J. **530**, 277 (2000).
20. M.-M. MacLow, astro-ph/0012499 (2000).
21. R. S. Klessen, F. Hetsch, and M.-M. MacLow, Astrophys. J. **535**, 887 (2000).
22. P. Massey, C. C. Lang, K. De-Gioia-Eastwood, and C. D. Garmany, Astrophys. J. **438**, 188 (1995).
23. P. Massey, in *New Views of the Magellanic Clouds (IAU Symposium 190)*, Ed. by You-Hua Chu, N. Suntzeff, J. Hesser, and D. Bohlender, p. 173.
24. R. S. Klessen, Astrophys. J. **556**, 837 (2001).
25. S. Kim, M. A. Dopita, L. Stavely-Smith, and M. S. Bessel, Astron. J. **118**, 2797 (1999).
26. E. S. Phinney, in *The Origins, Evolution, and Destinies of Binary Stars in Clusters,* Ed. by E. F. Milone and J.-C. Mermilliod (Astronomical Society of the Pacific, San Francisco, 1996), Astron. Soc. Pac. Conf. Ser. **90**, 185 (1995).
27. L. Hartmann, Astron. J. **121**, 1030 (2001).
28. A. D. Chernin, Usp. Fiz. Nauk **171**, 1153 (2001).
29. B. G. Elmegreen, Astrophys. J. **427**, 384 (1994).
30. P. Vader and B. Chaboyer, Astrophys. J. **445**, 691 (1995).
31. Yu. N. Efremov and B. G. Elmegreen, Mon. Not. R. Astron. Soc. **299**, 643 (1998).
32. G. Tenorio-Tagle, Astron. Astrophys. **88**, 61 (1980).
33. S. G. Stewart and F. Walter, Astron. J. **120**, 1794 (2000).
34. K. Rhode, J. J. Salzer, D. J. Westpfahl, and L. A. Radice, Astron. J. **118**, 323 (1999).
35. J. F. Wallin, J. L. Higdon, and L. Staveley-Smith, Astrophys. J. **459**, 555 (1996).
36. Yu. N. Efremov, B. G. Elmegreen, and P. W. Hodge, Astrophys. J. **501**, L163 (1998).
37. S. Matsushita, R. Kawaba, H. Matsumoto, *et al.*, Astrophys. J. **545**, L107 (2000).
38. D. K. Strickland and I. R. Stevens, astro-ph/9902188 (2000).
39. I. D. Karachentsev, M. E. Sharina, and W. K. Hutchtmeier, Astron. Astrophys. **362**, 544 (2000).
40. S. Ehlerova and J. Palous, Mon. Not. R. Astron. Soc. **330**, 1022 (2002); astro-ph/0111495 (2001).
41. Yu. N. Efremov, in *Gamma-ray Bursts in the Afterglow Era,* Ed. by E. Costa, F. Frontera, and J. Hjorth (Springer, Berlin, 2001), p. 243; astro-ph/0102161 (2001).
42. M. Bureau and C. Carignan, astro-ph/0104117 (2001).
43. C. Heiles, Astrophys. J. **229**, 533 (1979).
44. Yu. N. Efremov, S. Ehlerova, and J. Palous, Astron. Astrophys. **350**, 457 (1999).
45. A. Loeb and R. Perna, Astrophys. J. **503**, L35 (1998).
46. S. S. Larsen, Yu. N. Efremov, B. G. Elmegreen, *et al.*, Astrophys. J. **567**, 896 (2002).
47. Yu. N. Efremov, Pis'ma Astron. Zh. **26**, 649 (2000) [Astron. Lett. **26**, 558 (2000)].
48. V. I. Dokuchaev, Yu. N. Eroshenko, and L. M. Ozernoy, Astrophys. J. **502**, 192 (1998).
49. H. C. Spruit, Astron. Astrophys. **341**, L1 (1999).
50. Yu. N. Efremov, Usp. Fiz. Nauk **170**, 899 (2000).
51. J. S. Bloom, S. R. Kulkarni, S. G. Djorgovski, *et al.*, astro-ph/0010176 (2000).
52. J. U. Fynbo, S. Holland, M. I. Andersen, *et al.*, astro-ph/0009014 (2000).
53. D. Yu. Tsvetkov, S. I. Blinnikov, and N. N. Pavlyuk, astro-ph/0101362 (2001).
54. S. S. Larsen, J. P. Brodie, B. G. Elmegreen, *et al.*, Astrophys. J. **556**, 801 (2001).
55. Yu. N. Efremov, S. A. Pustilnik, A. Y. Kniazev, *et al.*, Astron. Astrophys. **389**, 855 (2002).
56. Ch. Fendt, R. Beck, and N. Neininger, Astron. Astrophys. **335**, 123 (1998).
57. A. Santillan, J. Franco, M. Martos, and J. Kim, Astrophys. J. **515**, 657 (1999).
58. A. D. Chernin, Yu. N. Efremov, and P. A. Voinovich, Mon. Not. R. Astron. Soc. **275**, 313 (1995).
59. R. A. Gruendl, S. M. Vogel, D. S. Davis, and J. S. Mulchaev, Astrophys. J. **413**, L81 (1993).
60. Yu. N. Efremov and A. D. Chernin, submitted to Usp. Fiz. Nauk.
61. Yu. N. Efremov, in *Extragalactic Star Clusters (IAU Symposium 207),* Ed. by E. K. Grebel, D. Geisler, and D. Minniti (in press).
62. R. I. Klein, T. Woods, and C. F. McKee, in *American Astronomical Society Meeting No. 198, 2001,* #87.04.
63. M.-M. MacLow, C. F. McKee, R. I. Klein, *et al.*, Astrophys. J. **433**, 757 (1994).
64. A. M. Read and W. Pietsch, Astron. Astrophys. **373**, 473 (2001).
65. G. Pietrzynski, W. Gieren, P. Fouque, and F. Pont, astro-ph/0103374 (2001).
66. S. G. Kim, H. Jung, and M. G. Lee, astro-ph/0203032.

Translated by A. Dambis

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 805–818. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 894–908.
Original Russian Text Copyright © 2002 by Shenavrin, Fernie, Weigelt, Yudin.

Photometric Observations of R Coronae Borealis in the Optical and Infrared

V. I. Shenavrin[1], J. D. Fernie[2], G. Weigelt[3], and B. F. Yudin[1]

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

[2] *David Dunlap Observatory, University of Toronto, Box 360, Richmond Hill, L4C 4Y6 Ontario, Canada*

[3] *Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

Received November 20, 2001; in final form, February 1, 2002

Abstract—The results of long-term photometric observations of R CrB in the $UBVJHKLM$ bands are presented. The temporal and color characteristics of the emission of the star itself and of its extended dust envelope are analyzed in detail. No stable harmonic has been found in the semiregular variations of the optical brightness of R CrB. Two harmonics with periods $P \approx 3.3$ and 11.3 yrs have been detected in the brightness variations of the dust envelope; the minima of these variations coincided in 1999, resulting in a record decrease in the LM brightness of the envelope. This by chance coincided in time with a deep minimum of the visual brightness of the star, resulting in a unique decrease in the total brightness of the star and dust envelope. This enabled estimation of the bolometric flux of the hot dust clouds, which made up only a few per cent of the bolometric flux of the dust envelope. The brightness variations of the dust envelope are not accompanied by appreciable color changes and are associated with variations of its optical depth $\tau(V)$ in the range 0.2–0.4. The dust envelope forms at a large and fairly constant distance from the star ($r_{\rm in} \approx 110R_*$, $T_{\rm gr} \approx 860$ K), from material in its stellar wind, whose intensity ($\dot{M}_{\rm gas} \approx 2.1 \times 10^{-7} M_\odot$/year) obeys a Reimers law. No variations synchronous with those of the optical depth of the dust envelope, in particular, with the period $P \approx 3.3$ yrs, have been found in the optical emission of R CrB, suggesting that the stellar wind is not spherically symmetric. The dust envelope consists of small grains ($a_{\rm gr} \leq 0.01$ μm), while the clouds screening the star from the observer are made up of large grains ($a_{\rm gr} \approx 0.1$ μm). The activity of R CrB, whose nature is unclear, is reflected in variations of the stellar-wind intensity and the appearance of dust clouds in the line of sight: these variations are repeated by corresponding changes in the optical depth of the dust envelope with a delay of ~4 years (the time for a particle moving at $V_{\rm env} \approx 45$ km/s to move from the star to the boundary of the dust envelope). © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

R CrB is the prototype of a small group of peculiar supergiants with characteristic light curves and spectra (see [1] and references therein). The visual brightness of the star can suddenly, i.e., unexpectedly, decrease by several magnitudes over a few weeks, returning several months later to its previous level. The carbon lines in the spectrum are anomalously strong, while hydrogen lines are very weak or completely absent. The abrupt decreases in the visual brightness are associated with the condensation of dust clouds in the stellar envelope, which screen the star from the observer when they lie in the line of sight. The peculiar character of the spectrum is due to the unusual chemical composition of the stellar atmosphere, which is very strongly depleted in hydrogen and appreciably enriched in carbon.

In the deepest minima, the visual brightness of R CrB decreases by $\sim 8^m$. A statistical analysis of the time intervals between consecutive minima has shown that these intervals are purely random; i.e., there is no regularity in the appearance of deep minima in the light curve of the star (see [2] and references therein). In addition to the deep minima associated with the condensation of dust clouds, semiregular cyclic variations with an amplitude of $\leq 0\overset{m}{.}25$ can be noted in the visual light curve of R CrB, and are probably due to stellar pulsations [3].

A periodogram analysis of the light curve of R CrB revealed three harmonics with the periods ~52, ~44, and ~33 d, whose predominance varies from year to year (see [4] and references therein). At some epochs, a harmonic becomes obvious in the light curve, while at other epochs, no cyclic variations are visible. At the same time, analysis of the radial-velocity curve (149 estimates from 1949 to 1995) indicated only one dominant period, ~42.7 d [5], with the amplitude of the corresponding harmonic being ~4 km/s.

The very first observations of R CrB in the near infrared showed an IR excess for the stellar radiation

1063-7729/02/4610-0805$22.00 © 2002 MAIK "Nauka/Interperiodica"

with a color temperature of ∼900 K [6], which is formed by an extended ($\sim 100R_*$) dust envelope. The presence of such envelopes is characteristic of R CrB stars. IRAS observations of R CrB have shown that the object is extended at 60 and 100 μm [7]. This represents a relic dust envelope, whose size exceeds that of the dust envelope corresponding to the present stage of evolution of the star by a factor of $\sim 10^4$. From an evolutionary point of view, R CrB stars are born-again red giants (born-again AGB stars) such as FG Sge and V4334 Sgr (Sakurai's object), while the relic dust envelope represents a remnant of the planetary nebula produced when the star passed the postasymptotic (post-AGB) track for the first time.

It now seems reasonably clear that the deep minima are connected with the condensation of dust clouds, not of a dust envelope that entirely covers the star [1]. However, it remains unknown at what distance from the star they condense and to what extent they participate in the formation of the extended dust envelope. Most of the observational facts suggest that the dust-cloud condensation occurs near the star, and it has been supposed *a priori* that the extended dust envelope forms directly from these clouds [1]. However, both of these statements require observational confirmation.

Observations of R CrB in the L band have shown semiregular fluctuations in the brightness of the star with a period of ∼1260 d (see [8] and references therein). Since the main contribution to the radiation of R CrB at these wavelengths comes from its dust envelope, the L brightness variations imply that the dust envelope is not in a stationary state. If this is the case, we wish to elucidate what changes in the dust envelope lead to the observed variations of its brightness, within what limits they vary, and to what extent their changes are reflected in the visual brightness of the star. This information can help shed light on the structure of the extended dust envelope and its formation, and indicate the origin of the nonstationary behavior of this dust envelope.

In this paper, we present the first UBV observations of R CrB in 1994–1999 and $JHKLM$ observations in 1983–2001, as well as complete tables of our photometric observations of R CrB. We use these data to analyze in detail various parameters of the star's radiation in the optical and IR, estimate the parameters of the extended dust envelope at different brightness levels, study the character of dust formation in the envelope, and ultimately derive answers to the questions formulated above. The unique dip in the dust envelope's brightness, combined with the deep minimum of the visual brightness of R CrB in 1999, can be used to isolate the direct emission of the dust clouds.

2. OBSERVATIONS

The UBV photometric observations of R CrB in 1994–1999 were carried out on the 0.25-meter telescope of the Faiborn Observatory (Arizona, USA), which is used for automatic photoelectric observations [9]. These supplement similar observations of R CrB obtained in 1985–1993 (see [4] and references therein) and are presented together with these in the first electronic table (ET1) on the site http://infra.sai.msu.ru/ftp/rcrb. ET1 contains 1170 brightness estimates with accuracy no worse than $0^m.02$. HD 141352 ($V = 7.476$, $B - V = 0.439$, $U - B = -0.003$) was used as the standard star.

Photometric $JHKLM$ observations of R CrB were carried out on the 1.25-m telescope of the Crimean Observatory of the Sternberg Astronomical Institute in 1983–2001 using the photometer described in [10]. These supplement observations of R CrB obtained in 1976–1979 [11] and are presented together with these on the same site http://infra.sai.msu.ru/ftp/rcrb, in the second electronic table (ET2). ET2 contains 252 brightness estimates with accuracy no worse than $0^m.03$. The standard star was BS 5947 ($J = 2.09$, $H = 1.60$, $K = 1.30$, $L = 1.12$, $M = 1.35$).

Figure 1 shows the brightness variations of R CrB in the $VJKLM$ bands in 1983–2001. We have added to our observations V observations obtained by amateur astronomers of the Variable Star Observers League of Japan (VSOLJ, http://www.kusastro.kyoto-u.ac.jp/vsnet) and the J and L observations of Feast *et al.* [8]. During this time interval, several deep minima were observed in the visual light curve, with a temporary quiescence in their appearance in 1989–1995. This gives the impression of packet grouping of the minima, in other words, of a cyclicity in the rate of their appearance.

An unusually deep minimum was observed in the L and M light curves of R CrB in 1999, following a smooth decline of the brightness from its maximum in 1994. We can see from Fig. 1 that the duration of the smooth decline of the LM brightness approximately coincides with the duration of the temporary quiescence in the appearance of deep minima. To match these two intervals, the L and M light curves bands must be shifted back in time by ∼4 years.

3. PHOTOMETRIC CHARACTERISTICS OF R CrB IN ITS BRIGHT STATE

We define the bright state of R CrB by the two conditions $V \leq 6.05$ (ET1) and $J \leq 5.25$ (ET2). The maximum recorded brightnesses are $V \approx 5.69$ and $J \approx 4.89$. The mean magnitudes of R CrB in its bright

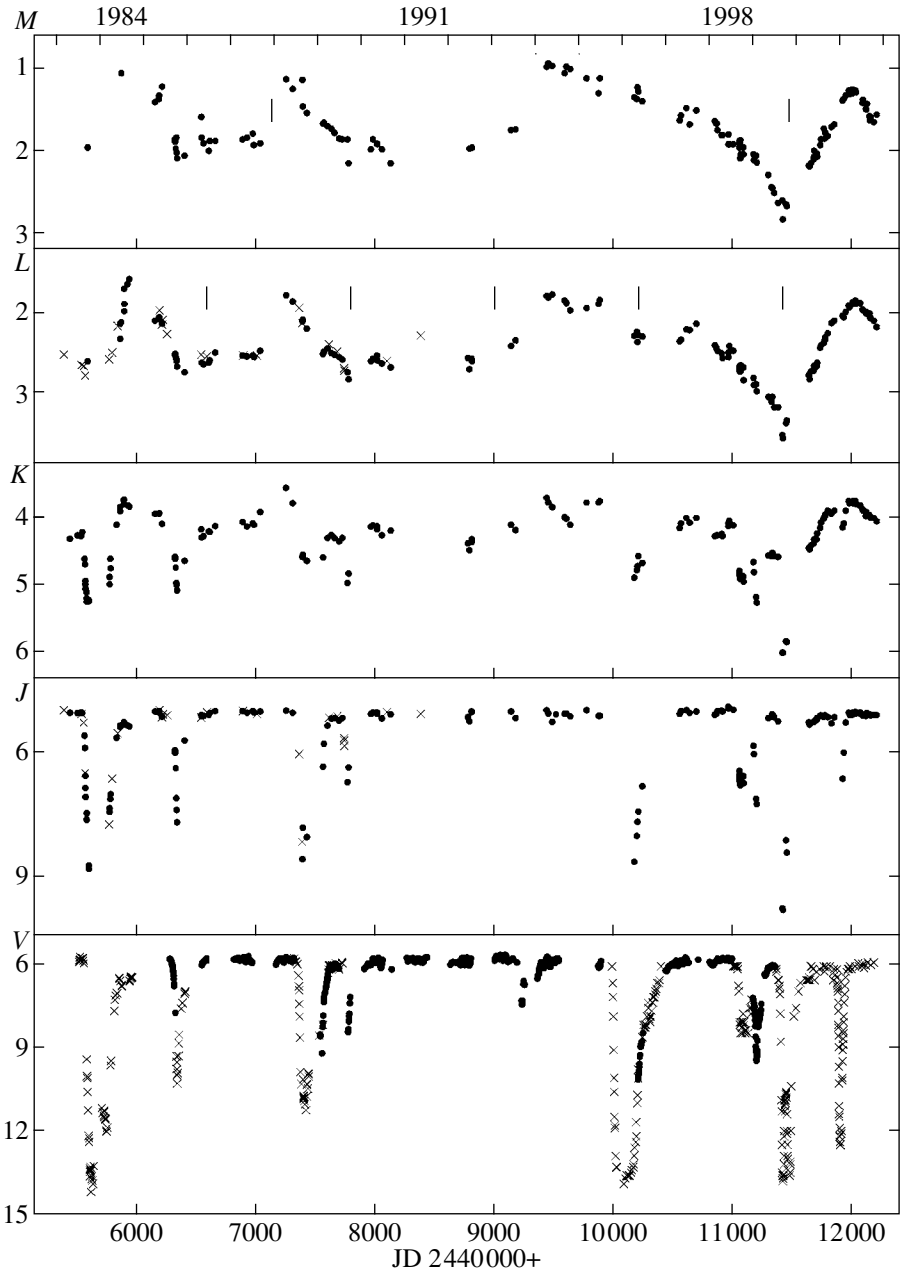

Fig. 1. Brightness variations of R CrB in the $VJKLM$ bands. The dots show our own observations, while the crosses show the VSOLJ observations (http://www.kusastro.kyoto-u.ac.jp/vsnet) in the (JD, V) plot and the data of Feast *et al.* [8] in the (JD, J) and (JD, L) plots. The vertical bars in the (JD, L) and (JD, M) plots mark the dates of LM brightness minima calculated using (15) and (16), respectively.

state (mean brightness of the star) are given in the table. These were calculated in the $UBVJHKLM$ and RI bands using the data of ET1, ET2, and [12, 13].

A periodogram analysis of the complete (1985–1999) visual light curve of R CrB in its bright state yields no pronounced harmonics in the spectrum of the stellar pulsations. Formally, the harmonic with $P \approx 39.8$ d has the greatest total amplitude $\sim 0^m.07$ (Fig. 2). The corresponding harmonic in the spectrum of the $B-V$ color-index curve has a total amplitude of $\sim 0^m.025$. Its minimum (the bluest color) leads the minimum of the visual light curve harmonic (maximum brightness) by ≤ 3 d. In our view, this provides evidence of in-phase variations of the brightness and color of R CrB; this same effect has been found for RY Sgr, whose pulsation period and amplitude are stable in time [1].

If the light curve is expanded in a Fourier series,

Mean magnitudes of R CrB in its bright state and maximum and minimum of the L brightness

Photometric band	Bright state	Maximum L	Minimum L
U	6.55	6.38	6.82
B	6.49	6.35	6.70
V	5.91	5.81	6.08
R	5.49	5.40	5.62
I	5.31	5.23	5.40
J	5.08	5.01	5.11
H	4.91	4.76	5.08
K	4.13	3.74	4.56
L	2.38	1.79	3.12
M	1.70	0.97	2.48

a large number of harmonics with total amplitudes $\sim 0^m.05$ appear in its spectrum, concentrated mainly in the range $P = 35$–40 d. At the same time, some segments of the light curve have an obviously cyclic appearance [4]. Note also the absence of a trend in the visual brightness of the star or its $(B-V)$ color.

The same can be said about the J light curve of R CrB. The three main harmonics with $P \approx 43.7$, 33.2, and 28.1 d have total amplitudes of $\sim 0^m.08$. The difference in the periods of the first harmonics for the visual and IR brightnesses of the star may be due to the non-simultaneity of the visual and IR observations, combined with the lack of stability in the brightness pulsations of R CrB over long time intervals. At the same time, Rao *et al.* [5] suggest that the radial-velocity curve (149 measurements from 1942 to 1995) has a cyclic character, with a pulsation period of ~42.7 d. There may be no contradiction here, since the brightness of the star can vary not only due to its own pulsations, but also in association with active phenomena on its surface, such as the formation of dark spots [1].

Figure 3 shows the spectral energy distribution (SED) of R CrB corresponding to its mean brightness, after correcting the observed magnitudes for interstellar reddening using the color excess $E(B-V) = 0.05$ [14]. The 0.23-μm and 0.25-μm fluxes were estimated from an IUE spectrum of R CrB obtained on May 31, 1991, when the optical and IR brightnesses of the star were close to their mean levels.

IRAS photometric observations of R CrB were carried out on September 12 and 13, 1983 [15], when our estimates indicate the L brightness to be $\approx 2^m.63$. The mean magnitude of R CrB in this filter at this time was $\sim 2^m.38$ (table), so that $\Delta L \approx 0.25$. It follows from [16] that the amplitude of the brightness variations at 11 μm is a factor of ~3 smaller than at 3.5 μm (the L band). Therefore, we increased the IRAS fluxes at 12, 25, 60, and 100 μm by 8% to calibrate them relative to the mean brightness level of the star. Note that the 60 and 100 μm fluxes do not include a contribution from the spatially extended relic envelope [7].

The bolometric flux of R CrB corresponding to its mean brightness is $F_{\rm bol,\,mean} \approx 1.55 \times 10^{-7}$ erg cm^{-2} s^{-1}. The wavelength ranges $\lambda \leq 0.36$ μm, $\lambda \geq 5$ μm, and $\lambda \geq 12$ μm contain ~4%, ~10%, and ~3% of the bolometric flux, respectively. In other words, ~85% of the bolometric flux is concentrated in the spectral range of our photometric observations, and possible inaccuracies (for example, in the correction of the IRAS fluxes in accordance with the L magnitude of R CrB), leave the bolometric-flux estimates virtually unaffected.

To estimate the IR excess (bolometric flux of the dust envelope), we must reconstruct the SED of the star after its passage through the circumstellar dust layer in the line of sight. We will denote the star's own radiation flux as $F_{*0}(\lambda)$, and the flux that emerges through the dust envelope (i.e., the observed flux) as $F_*(\lambda)$. We will call this simply the stellar radiation. Furthermore, for the sake of brevity, we will call the star + dust envelope combination the "object." The IR excess is $F_{\rm IR} = F_{\rm bol} - F_{*\rm bol}$.

We can derive an upper limit for the dust envelope's contribution to the J band when its L brightness is close to its mean value directly from the 1996 observations of R CrB (Fig. 1). In the deep minimum of 1996, the J brightness fell to $J \approx 8^m.62$ (ET2). If the J radiation at that time was due only to the dust envelope, the envelope's contribution to the object's radiation was ~4%. However, our photometric observations were not carried out at the very minimum of the optical brightness, but when the star was recovering from a minimum, and was already $\sim 0^m.6$ brighter. Moreover, it turns out that the contribution of the star to the J brightness remains appreciable in the brightness minimum (see below). As a result, we conclude that the contribution of the dust envelope to the object's radiation is $\leq$2%. Thus, the desired SED $F_*(\lambda)$ coincides with the observed energy distribution at wavelengths $\lambda \leq 1.25$ μm (i.e., downward from the J band).

Figure 4 presents $(U-B, V)$, $(B-V, V)$, $(V-R, V)$, $(V-I, V)$ and $(V-J, V)$ diagrams, which show the color variations of the star that occur as its brightness changes. The brightness estimates were taken from tables ET1, ET2 and from [12, 13]. The

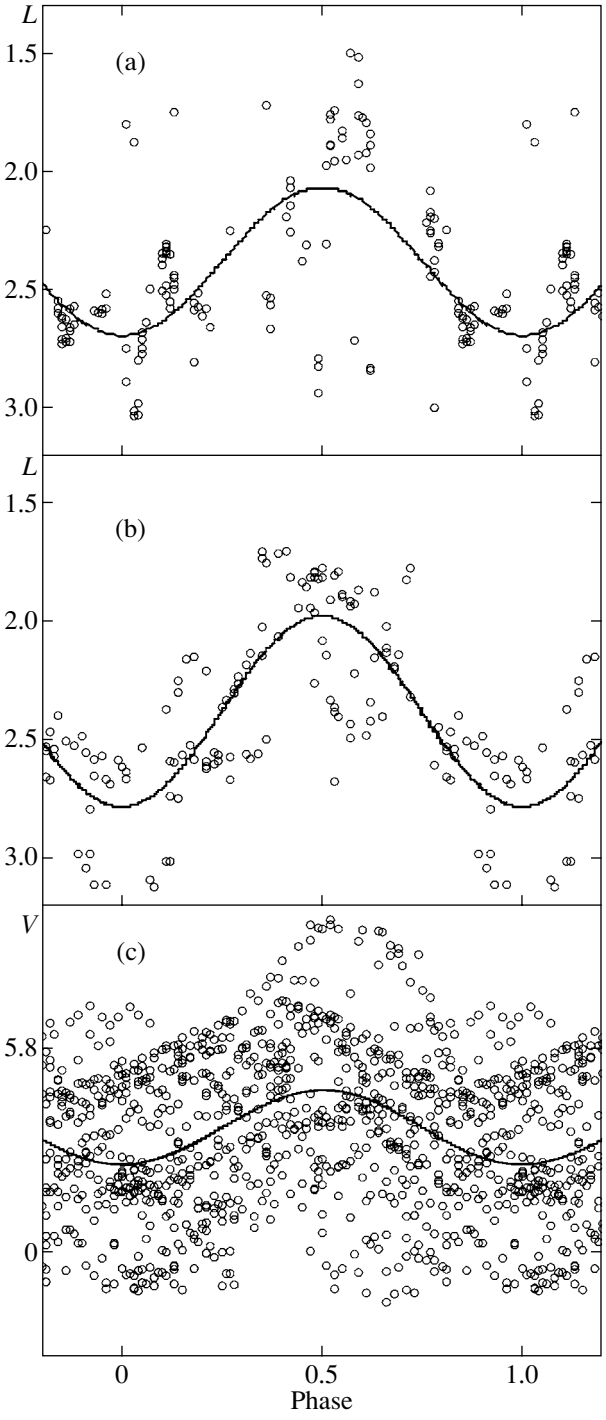

Fig. 2. Convolution of the (a) V light curve with the period 39.82 d and the L light curve with the periods (b) 1206 d and (c) 4342 d. The corresponding harmonics in the spectra are shown by the curves.

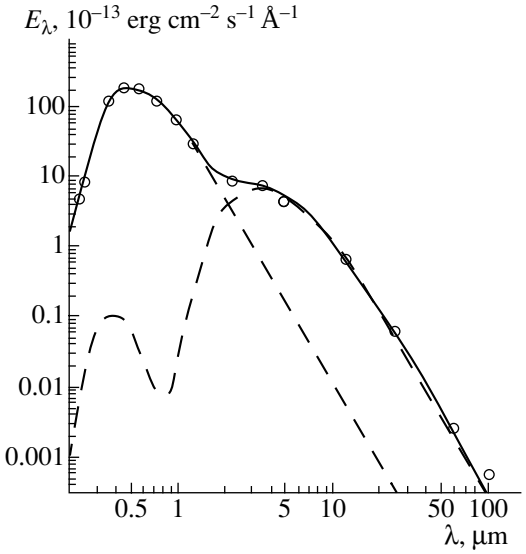

Fig. 3. Spectral energy distribution of R CrB for its mean brightness level (circles). The short- and long-dashed curves show the SEDs for the star and dust envelope, respectively. The solid curve shows the calculated SED.

$(V-J, J)$ diagram was plotted using data from these tables for observations carried out on the same night. Furthermore, we have used only data for dates when the L brightness of R CrB was below its mean level, and have taken into account the 2% contribution of the dust envelope's radiation in the J band. We derived the following linear approximations to these relationships:

$$U-B = 0.402V - 2.31, \quad (1)$$

$$B-V = 0.280V - 1.08, \quad (2)$$

$$V- = 0.301V - 1.36, \quad (3)$$

$$V-I = 0.430V - 1.94, \quad (4)$$

$$V-J = 0.418V - 1.62, \quad (5)$$

where $5.79 \le V \le 6.05$. We can see that the star reddens over the entire wavelength range considered as the star's brightness decreases.

During the minimum of 1999, the J and H brightnesses of R CrB fell to record low values: $J \approx 9.76$ and $H \approx 8.01$ (see Table ET2). Accordingly, the dust envelope's contribution to the object's radiation in these bands in the bright state did not exceed 1% and 6%. At that time, the L brightness of the star was minimum (Fig. 1). Plotting a (H, J) diagram for the case when R CrB was in its bright state and the L brightness was minimum ($L \ge 3.06$) using the data of ET2 and deriving a linear approximation for these data, we obtain the following relationship between the star's J and H brightnesses:

$$H = 0.90J + 0.44. \quad (6)$$

We have taken into account the 5% contribution of the dust envelope to the object's radiation in the second term. Using (5), we obtain the relationship

$$V-H = 0.474V - 1.95. \quad (7)$$

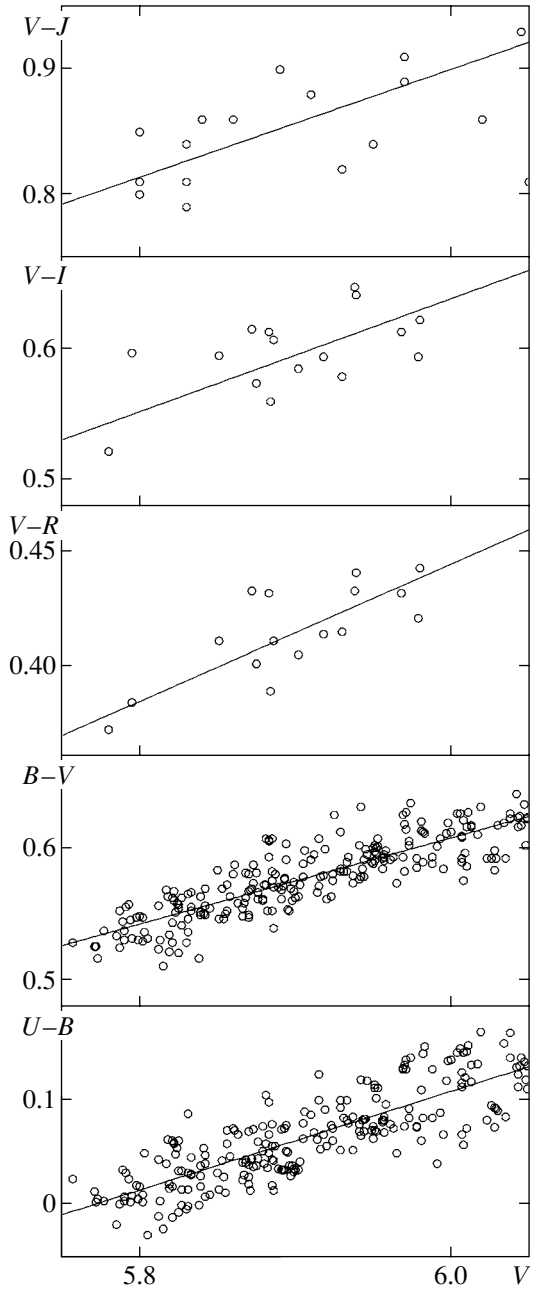

Fig. 4. $(U-B, V)$, $(B-V, V)$, $(V-R, V)$, $(V-I, V)$, and $(V-J, V)$ plots. The lines show linear approximations to these dependences [see (1)–(5)].

This can be used to estimate the star's H-band radiation.

To extend the SED to 1.62–4.8-μm (the $HKLM$ bands), we use calculated colors of supergiants whose parameters are typical of R CrB stars, including their atmospheric chemical compositions [14, 17]. These colors have been computed as functions of the effective temperature and surface gravity of the star. We adopt for R CrB the following values, which correspond to its mean brightness level: $T_{\rm eff} = 6750$ K, $L_{\rm bol} = 10^4 L_\odot$, $\log g = 0.5$ [14, 18]. Accordingly, $R_* \approx 73.4 R_\odot$, $M_* \approx 0.62 M_\odot$, $V_{\rm esc} \approx 57$ km/s, $D \approx 1.44$ kpc, $M_{\rm bol} \approx -5.3$. The star's intrinsic colors are $(H-K)_0 = 0.07$, $(K-L)_0 = 0.06$, $(L-M)_0 = 0.03$ [17]. At wavelengths $\lambda \geq 4.8$ μm, we set the SED equal to that of a blackbody with a temperature of 6750 K.

These intrinsic colors of the stars must also be corrected for interstellar and circumstellar reddening. However, since both corrections are very small (the dust envelope of R CrB is optically thin, $\tau(V) \approx 0.3$ and $\tau(K) \approx 0.05$), they will leave the the bolometric flux of the star virtually unchanged (the estimated bolometric flux increases by only $\sim$0.3%). Figure 3 shows the total SED $F_*(\lambda)$, derived as described above.

Thus, the bolometric flux of the star at its mean brightness level is $F_{*\rm bol} \approx 1.17 \times 10^{-7}$ erg cm^{-2} s^{-1}, and the IR excess is $F_{\rm IR,mean} \approx 3.8 \times 10^{-8}$ erg cm^{-2} s^{-1}. Thus, the dust envelope corresponding to the mean IR brightness of R CrB absorbs $\sim$24% of the star's radiation. If the envelope is spherically symmetric, its optical depth averaged over the spectral energy distribution is $\tau_{\rm eff,mean} \approx 0.28$. Note that, for a yellow star with an effective temperature of 6500–7000 K and amorphous carbon grains, $\tau_{\rm eff}$ differs from the optical depth at 0.55 μm ($\tau(V)$) by only a few per cent.

4. THE DUST ENVELOPE

4.1. Model of the Dust Envelope at the Mean Brightness of R CrB

Figure 3 shows the calculated SED of a star surrounded with a spherically symmetric dust envelope containing amorphous carbon grains with radius $a_{\rm gr} = 0.01$ μm. The index of refraction of the grains was taken from [19] (see sample cel1000 in [19]), and the efficiency factors for absorption and scattering were calculated in the framework of Mie theory for spherical particles. It is obvious that the optical brightness of the dust envelope cannot exceed the brightness of R CrB in the deepest minima, when the direct stellar radiation is blocked by a dust cloud. Accordingly, the albedo of the particles must be fairly small, and they should have fairly small sizes [20] not exceeding 0.01 μm in the case of R CrB.

The input parameters for the model calculations, which use the CSDUST3 code [21], are described in [22]. In particular, the input parameters include the SED of the intrinsic stellar radiation $F_{*0}(\lambda)$, which is not known. Therefore, after specifying the dust envelope's optical depth at 0.55 μm (input parameter), we calculated this distribution as $F_{*0}(\lambda) = F_*(\lambda) e^{\tau(\lambda)}$.

In this approach, the calculated SED will automatically match the observed SED down to the I band. However, this does not play a crucial role if we are interested in deriving the dust envelope's parameters from its radiation, which is predominantly in the IR.

The dust envelope is estimated to have optical depths at 0.55 and 2.2 μm $\tau(V) \approx 0.32$ and $\tau(K) \approx 0.05$, inner radius $r_{\rm in} \approx 111R_*$, grain temperature and density at the inner radius $T_{\rm gr} \approx 860$ K and $\rho_{\rm dust} \approx 1.1 \times 10^{-20}$ g/cm^3, radial density dependence $\rho(r) \propto r^{-2}$, and grain-production rate at the inner boundary $\dot{M}_{\rm dust} \approx 3.1 \times 10^{-9} M_\odot$/year. If we assume that all carbon is bound in grains, then, in accordance with the atmospheric chemical composition of R CrB [18], the gas density at the inner radius of the dust envelope will be $\rho_{\rm gas} \approx 7.3 \times 10^{-19}$g/cm^3. The mean mass-loss rate of R CrB is $\dot{M}_{\rm gas} \approx 2.1 \times 10^{-7} M_\odot$/year for an envelope expansion velocity of $V_{\rm env} \approx 45$ km/s (the choice of this value will be justified below). However, some of the carbon in the stellar envelope can also be bound in CO, which, by virtue of its chemical inertness, does not stick to grain surfaces. In this case, only some of the carbon participates in dust-grain production. If we assume that all the oxygen in the gas envelope of R CrB is bound in CO and take its abundance from [18], the estimated gas density and mass-loss rate should be increased by a factor of $\sim$2.5. For the adopted expansion velocity, the time for the gas lost by the star to reach $r_{\rm in}$ is $\sim$4 years.

The visual brightness of the dust envelope is $V \approx 14.6$, and its color indices are $U{-}B \approx -0.65$, $B{-}V \approx -0.28$, $V{-}R \approx -0.41$, and $R{-}I \approx 1.47$. The SED of the envelope is presented in Fig. 3. Since the grain size is much smaller than optical wavelengths, the grains scatter optical light in accordance with a Rayleigh law ($\sigma_{\rm sca} \propto \lambda^{-4}$), which gives a blue color to the envelope in the $UBVR$ bands. In the I band, the thermal radiation of the grains already dominates over scattering, and the envelope's color reddens.

In the deepest minima, the visual magnitude of R CrB reaches $\sim 14^m$. The model dust envelope made up of grains with sizes $\sim$0.01 μm has $V \approx 14.6$. If we increase the grain radius to, e.g., $a_{\rm gr} = 0.025$ μm, the visual brightness of the dust envelope will increase by a factor of $\sim$15 ($\Delta V \approx 3$) at the expense of a sharp increase in the grain albedo, and will appreciably exceed the brightness of R CrB in deep minima.

As a rule, the star becomes bluer in the final stages of the brightness decrease; i.e., the $U{-}B$ and $B{-}V$ color indices decrease, reaching their minima at the visual-brightness minimum. However, several days later, at the same visual brightness, this blue color disappears, and the star appreciably reddens. For example, in the minimum of 1999, the colors changed from blue ($B - V \approx 0.3$, $V \approx 13.7$) to red ($B - V \approx 1.1$, $V \approx 13.7$) in only two days [23] (see also ET1, ET2). Such rapid color variations rule out the dust envelope as the source of radiation responsible for this effect. Currently, the blue color is attributed to the chromospheric emission of atoms and molecules, which is most clearly visible when the dust cloud completely blocks the stellar disk but not the higher-lying chromospheric layers [24, 25].

Since the dust-formation rate at the inner boundary of the dust envelope varies with time (see below for details), a power-law distribution for the dust density in the model envelope is only an approximation. We have approximated the actual density distribution with the relationship $\rho(r) \sim r^{-2}$, as is quite natural. The density of every individual dust layer obeys this law as it expands. The fluctuation timescale and corresponding spatial scale are $\Delta t < 4$ years and $\Delta r < r_{\rm in}$.

As long as the envelope is optically thin in the IR, its color characteristics are virtually independent of the optical depth. Therefore, from the point of view of model calculations, we can replace the spherically symmetric dust envelope by an envelope in the form of a continuous cloud blanket with light and dark areas, while maintaining its spectral energy distribution. In addition, in order to preserve the brightness, the optical depth of the continuous cloud blanket averaged over a sphere should be equal to the optical depth of the spherically symmetric envelope.

The spherical symmetry can be replaced with axial symmetry by increasing the optical depth of a bipolar dust envelope in accordance with the part of the sphere it blocks, so that its bolometric flux is equal to the bolometric flux of the IR excess. In this model, it is possible that the dust envelope does not screen the star from the observer at all. In this case, $F_{*0}(\lambda) = F_*(\lambda)$. Note that preserving the IR excess also preserves the dust envelope's optical brightness.

However, such theoretically possible forms of the dust envelope's shape assume that its inner radius remains unchanged or varies only insignificantly. Considerable changes in the inner radius would result in appreciable variations of the color characteristics of the dust envelope. For example, if this radius is decreased by 10%, the $H{-}K$ and $K{-}L$ color indices of the envelope will increase by $\sim 0\overset{m}{.}1$. Thus, the extended dust envelope cannot be made up of dust clouds if these condense at much smaller distances from the star than its inner radius [1]. It is most likely that the clouds and envelope form separately from each other.

The extended dust envelope forms at a considerable and fairly constant distance from the star and, from this point of view, is analogous to the dust

envelopes of, say, red giants. Moreover, the optical depths of the dust envelopes in Miras vary with the brightness phase. In other words, the smooth decline of the dust-envelope density due to its expansion is overlaid with semiregular fluctuations, associated with changes in the dust-formation rate at its inner boundary. Of course, we cannot exclude the possibility that the dust clouds reach the inner boundary of the dust envelope without evaporating on the way. However, such clouds do not contribute appreciably to the IR excess.

The mass-loss rate of a yellow supergiant can be estimated using the well-known formula of Reimers [26]

$$\dot{M} = 4 \times 10^{-13} \eta LR/M\ M_\odot/\text{year}, \qquad (8)$$

where $0.3 \leq \eta \leq 3$, and the parameters of the star are expressed in solar units. If we substitute the values of L_*, R_*, and M_* we have adopted for R CrB, it turns out that, for $\eta \approx 0.5$, the mass-loss rate estimated by this formula coincides with that obtained from the model calculations of the dust-envelope structure. Thus, a superwind is not required to explain the dust envelope's formation.

Note also that the estimated stellar-wind density at a distance, for example, $r \approx 2R_*$, which is equal to $\rho_{\rm gas} \approx 2.6 \times 10^{-15}$ g/cm^3 ($\rho(r) \propto r^{-2}$), falls in the interval 6×10^{-16}–6×10^{-15} g/cm^3 in which the gas temperature behind a shock propagating at a velocity of $\geq$20 km/s can decrease to $\sim$1500 K [27]. This creates the conditions required for the formation of complex carbon molecules, from which dust grains could condense. However, this would also require that the lower-temperature region be appreciably screened from the direct stellar radiation, since, in contrast to complex molecules, whose temperature is determined by the gas kinetic temperature, the temperature of the dust grains is determined by the stellar radiation field. If the grains condense at a distance of $\sim 2R_*$, the direct radiation must be attenuated by two orders of magnitude.

It was shown in [3] that, in the atmosphere of a pulsating star, the gas density at a distance of $\sim 4R_*$ is $\sim 10^{-12}$ g/cm^3, which considerably exceeds the gas density at such distances obtained from our calculations. In this case, either only a small fraction of the extended atmosphere of R CrB is lost by the star or not all the carbon in the stellar wind condenses into grains.

4.2. The Dust Envelope at Its Maximum and Minimum Optical Depth

The table lists the stellar magnitudes of R CrB at its maximum and minimum L brightness in 1994 and 1999, respectively (Fig. 1), obtained by averaging optical and IR observations carried out on the same dates in April–May 1994 and May–June 1999 (ET1, ET2). The R and I magnitudes were estimated using (3) and (4). The fluxes at 12, 25, 60, and 100 μm were estimated from the L brightness of R CrB (as explained earlier). To estimate the fluxes at 0.23 and 0.25 μm based on their values for the mean brightness level, we assumed equal amplitudes for the brightness variations in the UV and in the U band. As mentioned above, the probable errors in the estimated fluxes in the UV, mid-IR and far-IR do not appreciably affect the estimated bolometric flux of the star.

The bolometric fluxes of R CrB at its maximum and minimum L brightness are $F_{\rm bol,max} \approx 1.90 \times 10^{-7}$ erg cm^{-2} s^{-1} and $F_{\rm bol,min} \approx 1.24 \times 10^{-7}$ erg cm^{-2} s^{-1}. Accordingly, the IR excesses and effective optical depths are $F_{\rm IR,max} \approx 6.1 \times 10^{-8}$ erg cm^{-2} s^{-1}, $\tau_{\rm eff,max} \approx 0.39$ and $F_{\rm IR,min} \approx 2.1 \times 10^{-8}$ erg cm^{-2} s^{-1}, $\tau_{\rm eff,min} \approx 0.19$. Thus, the range of optical-depth variations of the dust envelope of R CrB in the visual is $\Delta\tau_{\rm eff} \approx 0.2$. At the same time, the K–L color index of the IR excess associated with the dust envelope's radiation at its maximum and minimum optical depths turned out to be the same, $\sim 2\overset{m}{.}35$. Consequently, the inner radius of the dust envelope does not vary appreciably with changes in its optical depth. In other words, although the grain-condensation rate varies with time, the grain condensation takes place at approximately the same distance from the star.

In the case of a spherically symmetric dust envelope, the change in the observed bolometric flux (object's flux) is equal to the change in the bolometric luminosity of the star: $\Delta m_{\rm bol} = \Delta m_{*0,\rm bol}$. Correcting the estimated visual magnitude of R CrB for absorption in the circumstellar envelope at its maximum and minimum optical depths, we find $\Delta V_0 \approx \Delta m_{\rm bol} \approx 0.49$. Since $\Delta m_{*0,\rm bol} \approx \Delta V_0$ for yellow F supergiants, we find $\Delta m_{\rm bol} = \Delta m_{*0,\rm bol}$. Hence, the assumption that the dust envelope is spherical withstands this test. At the same time, we cannot reject a bipolar model for the dust envelope in which there is no dusty material in the line of sight based on such a test. In this case, $F_{*0}(\lambda) = F_*(\lambda)$; i.e., $\Delta m_{*0,\rm bol} = \Delta m_{*,\rm bol}$, $\Delta V_0 = \Delta V$, and we must have $\Delta V \approx \Delta m_{*,\rm bol}$. We can see from the table that $\Delta V \approx 0.27$, and the estimated value of $\Delta m_{*,\rm bol}$ is $0\overset{m}{.}24$.

4.3. Color Characteristics of the Dust Envelope and an Estimate of Its Optical Depth from the L Brightness of the Star

Figure 5 presents ($H_{\rm IR}$, $L_{\rm IR}$), ($K_{\rm IR}$, $L_{\rm IR}$), and ($M_{\rm IR}$, $L_{\rm IR}$) plots for variations in the dust envelope's

HKM magnitudes as a function of its L magnitudes. These were calculated individually for each date of IR observations of R CrB in its bright state assuming that the J radiation of R CrB is due entirely to the star. We estimated the brightness of the star in the $HKLM$ bands in accordance with the rules described above. We then subtracted the stellar radiation from the total radiation of R CrB and found the magnitudes corresponding to the IR excess. Linear approximations of the above plots yield the relationships

$$H_{\rm IR} = 1.22L_{\rm IR} + 4.22, \quad (9)$$

$$K_{\rm IR} = 0.97L_{\rm IR} + 2.35, \quad (10)$$

$$M_{\rm IR} = 0.98L_{\rm IR} - 0.66. \quad (11)$$

An approximation using a quadratic polynomial virtually coincides with the linear approximation. We can see from the above relationships that, on average, changes in the dust envelope's brightness due to changes in its optical depth take place without appreciable changes in its color characteristics. Thus, we again come to the conclusion that the inner radius of the extended dust envelope is constant, in contrast to the dust-formation rate. The mean L brightness of the dust envelope is $L_{\rm IR,mean} \approx 2.49$, and its colors are $(H-L)_{\rm IR,mean} = 4.76 \pm 0.4$, $(K-L)_{\rm IR,mean} = 2.28 \pm 0.15$, and $(L-M)_{\rm IR,mean} = 0.70 \pm 0.12$.

Knowing the mean H magnitudes of the dust envelope and object, we can estimate the H brightness of the star itself. Using (6), we estimate the J brightness of the star and then of the dust envelope, which is ≈ 10.16; the corresponding $J-H$ color index is ≈ 2.95. The calculated values of these parameters in the model for the star at its mean brightness (Fig. 3) nearly coincide with these values. However, since the contribution of the dust envelope's radiation in the J band is only a few per cent, and the actual dust grains may differ from the model grains, this coincidence should not be assigned too much importance. The color temperature corresponding to the above $J-H$ color index is ~870 K.

Since the color characteristics of the dust envelope are fairly stable, its bolometric magnitude ($m_{\rm IR,bol}$) should vary in direct proportion to, e.g., its L magnitude ($L_{\rm IR}$). Such a relation is, indeed, observed between the estimates of these parameters at the mean brightness of R CrB ($F_{\rm IR,mean}$), as well as at its maximum and minimum L brightnesses ($F_{\rm IR,max}$ and $F_{\rm IR,min}$):

$$m_{IR,\rm bol} \approx L_{\rm IR} + 4.55. \quad (12)$$

At the mean brightness, the L magnitude of the dust envelope is $L_{\rm IR} \approx 2.49$.

Taking into account the fact that $F_{\rm IR} = F_{\rm bol}(1 - \exp(-\tau_{\rm eff}))$, we can calculate the dependence $\tau_{\rm eff} = f(m_{\rm IR,bol})$, and, accordingly, $\tau_{\rm eff} = f(L_{\rm IR})$. Approximating the relationship $L_{\rm IR} = f(L)$ derived for all the dates of the IR observations with a quadratic polynomial and substituting it into (12) and the function $\tau_{\rm eff} = f(L_{\rm IR})$, we can estimate the bolometric magnitude of the dust envelope and its optical depth from the L magnitude of R CrB when the star is in its bright state:

$$m_{\rm IR,bol} \approx 0.046L^2 + 0.88L + 4.68, \quad (13)$$

$$\tau_{\rm eff} \approx 0.21L^2 - 1.34L + 2.26. \quad (14)$$

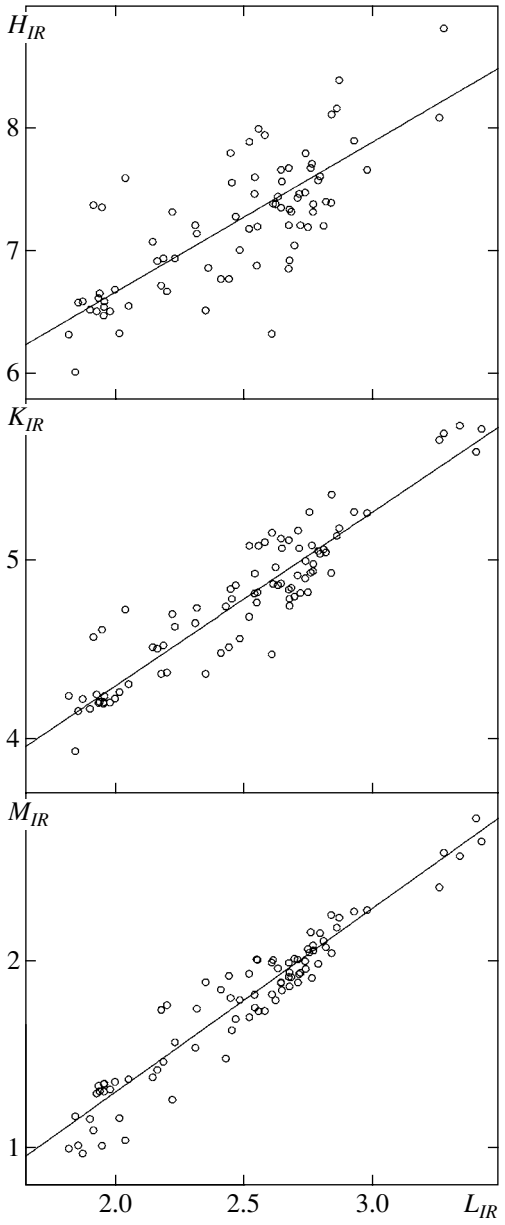

Fig. 5. ($H_{\rm IR}$, $L_{\rm IR}$), ($K_{\rm IR}$, $L_{\rm IR}$), and ($M_{\rm IR}$, $L_{\rm IR}$) plots for the radiation of the dust envelope of R CrB. The lines show linear approximations to these dependences [see (9)–(11)].

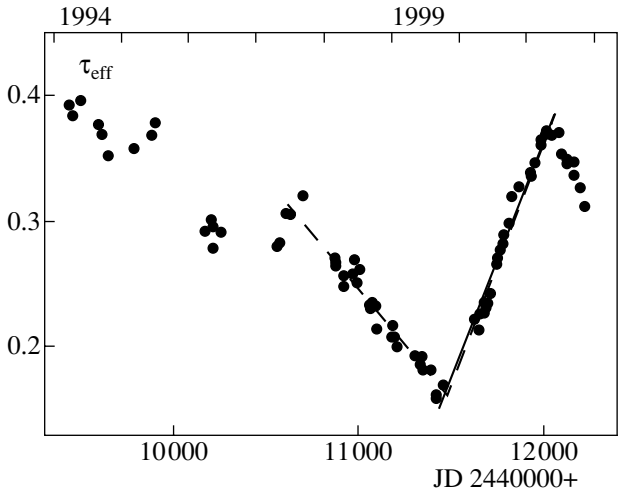

Fig. 6. Variation of the optical depth of the dust envelope near its deepest brightness minimum in 1999. The dashed lines show linear approximations to the (τ_{eff}, JD) relation. The solid line shows the model curve for this relation.

4.4. Characteristics of the Dust-Envelope Variability

Figure 6 shows the time dependence of the optical depth of the dust envelope near the deepest minimum in September 1999 calculated using (14). The sharp transition from the decrease to the increase of the optical depth and the nearly linear growth to its maximum value in 2001 cannot be reproduced in a model in which the dust-formation rate at the inner boundary of the dust envelope varies smoothly, e.g., according to a power law. This requires that the dust-formation rate initially increase stepwise near the minimum of the optical depth.

The solid line in Fig. 6 shows the calculated dependence for the growth of the optical depth of the dust envelope when the dust-formation rate at its minimum increases stepwise by a factor of 4.3 and then increases according to a power law $\propto t^{1.3}$ over ~560 days, almost until the maximum of the optical depth. The grain-condensation rate depends directly on the gas density, i.e., on the stellar-wind intensity. However, the changes of the stellar-wind intensity had begun four years before the corresponding changes in the grain-condensation rate. Recall that four years is the time of flight of a grain from the star to the inner boundary of the dust envelope at the stellar wind velocity of ~45 km/s; i.e., this is the time delay between such phenomena as changes of the stellar-wind intensity and of the dust-formation rate.

For such a delay, the stellar-wind enhancement coincides in time with the beginning of a new phase of activity of R CrB in the production of dust clouds, initiated at the end of September 1995 by a very deep and prolonged minimum in the visual light curve (Fig. 1). The same time delay is observed between the termination of the previous activity phase in 1989 (a successive packet of deep minima in the visual light curve) and the maximum of the LM brightness in 1994 (Fig. 1).

In other words, if we shift the L and M light curves of R CrB back by ~4 years, the LM minimum of 1999 will coincide with the termination of the phase of quiescence in the production of dust clouds in 1995. At the same time, the maximum of the LM brightness will coincide with the termination of the phase of activity in the production of dust clouds in 1989, and the interval of temporary quiescence in their production in 1989–1994 will coincide with the period of the smooth decrease of the LM brightness of the star.

In fact, we initially estimated the time interval needed to superpose these events to be ~4 years. Estimating the radius of the dust envelope from the model calculations, we determined the stellar-wind velocity to be ~45 km/s. Note that the nebulae detected around the stars FG Sge and V4334 Sgr, which probably also have the evolutionary status of born-again red giants, are expanding with approximately the same velocity.

Based on all these considerations, we conclude that the active phenomena on the surface of R CrB that stimulate the formation of clouds also stimulate an enhancement of the stellar wind, which is reflected in an increase in the optical depth of the extended dust envelope ~4 years later. In turn, the delay we have found between these phenomena confirms that the dust clouds form near the star.

We supplemented our L light curve of R CrB with brightness estimates obtained by Strecker [28] in 1968–1974 (35 points) and Feast *et al.* [8] in 1983–1991 (24 points), eliminated times of reduced visual brightness of the star, and carried out a periodogram analysis using the software of Yu.K. Kolpakov (http//infra.sai.msu.ru/prog/kolpakov). The first harmonic of the semiregular cyclic changes in the brightness of the dust envelope has a period of ~1206.6 d (~3.3 years) and a total amplitude for the brightness variations $\Delta L \approx 0.8$, while the second harmonic has ~4342 d (~11.9 years) and $\Delta L \approx 0.6$. Figure 2 shows convolutions of the L light curve of R CrB with these periods together with the corresponding harmonics. The dates of their minima are

$$\mathrm{Min}L = 2442986 + 1206\overset{\mathrm{d}}{.}6E, \qquad (15)$$

$$\mathrm{Min}L = 2442860 + 4342^{\mathrm{d}}E. \qquad (16)$$

In Fig. 1, the dates of minima are marked by vertical bars above the L and M light curves. In 1999, the minima of both harmonics coincided, and a very deep minimum was observed in the L and M light

curves of the star. Note that the estimated periods of the harmonics depend fairly substantially on the size of the series of observations being processed. In particular, if we restrict the L brightness curve to our own observations, the fundamental period becomes $\sim 1166^{\rm d}$, and the calculated minimum of this harmonic in 1999 occurs $\sim$2 months earlier than for the harmonic with $P \approx 1206$ d.

At the same time, if we process the entire series of observations without omitting times of reduced visual brightness of R CrB, the period of the first harmonic decreases by only $\sim$2.6 d ($\sim$0.2%), becoming $\sim$1204 d. The period of the 11-year harmonic becomes $\sim$4251 d; i.e., it diminishes by only $\sim$2%. Note that the average time interval between deep dips in the visual brightness of R CrB is $\sim$1100 d [29]. Furthermore, the characteristic time interval between groups of deep minima—i.e., between intervals of frequent and infrequent appearances of such minima—is $\sim$4400 d [30]. Note also that the colors of the first harmonic in the $JHKLM$ bands approximately coincide with the average colors of the dust envelope.

In the optical, where the dust envelope is manifest through its absorption of the star's radiation, the object's brightness should vary in antiphase with variations in the dust envelope's brightness associated with changes in its optical depth. Using (15) and substituting the average L brightness of R CrB, we find that $\Delta\tau_{\rm eff} \approx -0.34\Delta L$. As already noted, $\tau_{\rm eff}$ differs from $\tau(V)$ by only a few per cent for stars with effective temperatures of 6500–7000 K and amorphous carbon grains, with $\tau_{\rm eff} \leq \tau(V)$. In addition, for the small grains that make up the extended envelope, $\tau(U) \approx 1.8\tau(V)$.

Thus, if the dust envelope is spherically symmetric, there should be a harmonic in the V and U brightness variations of R CrB with a period of $P \approx 1206$ d and total amplitudes of $\Delta V \approx 0.3$, $\Delta U \approx 0.5$, which should be in antiphase with the corresponding harmonic of the dust envelope's brightness variations, whose amplitude is $\Delta L \approx 0.8$. However, a spectral analysis of the V and U light curves of R CrB does not reveal such a harmonic. As usual, times of reduced visual brightness of R CrB were omitted from the analyzed series of observations. This result was rather unexpected for us. It turns out that the stellar wind of R CrB is not spherically symmetric, and is less intense along the line of sight.

Thus, the activity of R CrB, whose nature remains unclear, is reflected in variations of the stellar-wind intensity and, with a time delay, of the optical depth of the dust envelope. The formation of dust clouds is apparently also connected with this activity. However, in order for a deep minimum to appear in the visual light curve, a cloud must not only condense, but condense in the line of sight; therefore, the sequence of deep minima does not display obvious signs of cyclicity. Nevertheless, during each activity period, there appears, on average, one cloud in the line of sight. This coincidence probably reflects the spatial structure of active phenomena on the surface of the star.

5. DUST CLOUDS

The appearance of dust clouds in the line of sight results in a strong dip of the visual brightness of the star. At the same time, the thermal radiation of the clouds has not been detected [8, 11]. There is no pronounced component with a color temperature that appreciably exceeds the color temperature of the extended dust envelope in the IR excess. Of course, there should not be any such component if the dust clouds condense far from the star.

It is obvious that the "hot" component must be searched for when the background radiation of the star and extended envelope is much decreased. Precisely such a unique case occurred in 1999, during a deep minimum of the visual brightness of the star superimposed on a record deep minimum of the dust-envelope emission (Fig. 1). It is also obvious that signs of the hot component should be searched for in the J and H bands.

We have $BVRI$ [23] and $JHKLM$ (ET2) observations of R CrB for two dates during the 1999 minimum (JD 2451423 and 2451435). The estimated brightness decreases of the star on these dates are $\Delta R \approx 8.03$ and 4.73, $\Delta I \approx 7.08$ and 4.12, $\Delta J \approx 4.71$ and 3.33, and $\Delta H \approx 2.99$ and 2.49, respectively. The ratio A(JD 2451423)/A(JD 2451435) in the $RIJH$ bands is $\sim$1.70, $\sim$1.72, $\sim$1.41, and $\sim$1.20, respectively.

It is natural to relate the drop in this ratio in the JH bands to the background radiation of the dust material, which begins to be significant in these bands during the brightness minimum. To preserve the ratio in the J band, the radiation of the dusty material (dust envelope + dust clouds) must make up $\sim$76% of the object's radiation at the time of minimum brightness of R CrB (JD 2451423): $J_{\rm total} \approx 9.8$, $J_{\rm dust} \approx 10.1$, and $J_* \approx 11.3$. All estimates of the hot component's parameters given below refer precisely to this time.

Using the estimate of the mean J brightness of the dust envelope obtained earlier and the fact that the dust-envelope colors do not change significantly with changes in its optical depth, we can estimate the J brightness of the dust envelope, which corresponds to its observed L brightness at the time of minimum: $J \approx 11.35$. We obtain the same value if we assume that the dust envelope contributes $\sim$90% to the object's flux in the K band at the time of minimum

brightness of R CrB, and that its colors are the same as for the mean brightness.

Accordingly, the J magnitude of the hot component of the IR excess is $J_{\rm hot} \approx 10.5$. In our view, this is the maximum brightness that we can assign to the dust clouds. In the J band, the clouds are the brightest component at the time of minimum brightness of the star and dust envelope, and contribute $\sim$70% of the object's radiation.

Performing the same analysis with the H brightness of the star, we obtain $H_{\rm tot} \approx 8.0$, $H_{\rm dust} \approx 8.1$, $H_* \approx 8.10$, $H_{\rm hot} \approx 9.5$, and, accordingly, $(J - H)_{\rm hot} \approx 1.0$. The color temperature corresponding to this color index is $\sim$2000 K. In the K band, the hot component contributes only $\sim$8% of the object's radiation.

If the cloud is optically thick and, accordingly, radiates as a blackbody with the above temperature, its bolometric flux is $F_{\rm bol,hot} \approx 4.4 \times 10^{-14}$ erg cm^{-2} s^{-1} Å, which comprises $\sim$1% of the mean bolometric flux of the dust envelope and $\sim$0.3% of the mean flux of R CrB. The size of the cloud is $R_{\rm hot} \approx 0.5R_*$. If it is behind the star and its heated side is observed, it is a distance $r \approx 10R_*$ from the star.

Recall that we are considering the deep minimum of the optical brightness of R CrB in 1999. This means that at least one dust cloud was certainly present in the dust envelope. However, because of its considerable optical depth, its outer side was most likely cool, so that its contribution to the total radiation of the dust material was negligible. Of course, estimates of parameters of the hot component should be treated with a healthy scepticism, since it is difficult to isolate the properties of this component owing to its low brightness. We can be sure only that its bolometric flux does not exceed a few per cent of the bolometric flux of the dust envelope.

Isolation of the hot component in the IR excess of R CrB is hindered by the considerable background radiation of the always present extended dust envelope. At the same time, the appearance of IR excesses in the spectral energy distributions of the stars FG Sge and V4334 Sgr—i.e., the appearance of dusty material in their envelopes—has been directly observed in 1992 and 1997. Without going into details, it is reasonable to say that the color temperature of the IR excess in the emission of both objects has been $\sim$1000 K since the time of its appearance [22, 31]. In other words, the dust envelopes of these stars formed from dust clouds that had condensed near the surface of the star, since we have found no obvious signs of a hot component in the IR excess from the very onset of the dust formation.

Figure 3 clearly illustrates that the 0.7–0.9-μm wavelength range is optimal for analyses of the optical behavior of grains in dust clouds blocking the emission of the star, since the contribution of the dust envelope to the emission of the object is minimum there. The contribution from emission lines of the gas envelope is also minimum. In other words, the emission of the star attenuated by the dust cloud dominates. This is confirmed by the absence of appreciable changes in the $R - I$ color index during the star's transition from a blue to a redder color; i.e., during appreciable changes of the B–V color index, which, as a rule, accompany deep dips in the brightness of R CrB [23, 25]. This dominance will be maximum when the optical depth of the dust envelope is minimum, i.e., when its brightness is minimum. Therefore, the minimum of 1999 provided optimal conditions, not only for searches for the clouds via their direct emission, but also for analyses of their optical properties.

We used the $BVRI$ observations of R CrB in this minimum [23] to plot a $(I, R - I)$ diagram. A quadratic approximation to these data yields the relationship $I \approx -10.6(R - I)^2 + 28.3(R - I) - 6.34$, and, accordingly, $A_I/(A_R - A_I) \approx -21.2(R - I) + 28.3$. At the minimum brightness, $R - I = 1.11 \pm 0.07$, and $A_I/(A_R - A_I) \approx 5$. For amorphous carbon grains with radius $\leq$0.1 μm, $A_I/(A_R - A_I) < 4$ ($\sim$2.5 for $a_{\rm gr} \approx 0.01$ μm), independent of whether this ratio is calculated for the extinction or absorption coefficient.

For grains ~ 0.2 μm in size, absorption and scattering at 0.7–1-μm already becomes neutral, so that the value of $A_I/(A_R$-$A_I)$ tends to infinity. Thus, in contrast to the small grains constituting the dust envelope, the grains in the dust cloud must be large, in order to provide sufficiently neutral absorption of the stellar radiation up to the near infrared. We arrived at the same conclusion from an analysis of the first deep minimum in the light curve of Sakurai's object in 1998, if this minimum was associated with the formation of a dust cloud in the line of sight [32].

6. CONCLUSION

Stellar winds are an essential characteristic of supergiants, and create extended gaseous, and, if the wind is sufficiently intense, gas–dust envelopes. However, this term is rarely used for R CrB stars, and it is not used at all when referring to dust envelopes. Most attention has been concentrated on the formation of dust clouds, which, once they have formed, make up the dust envelope of the star. At least we are not aware of any papers in which dust clouds and the extended dust envelope of R CrB are treated as independent structural elements of the circumstellar envelope from the point of view of their formation.

Nevertheless, it seems that they are, in fact, independent structural elements. The extended envelope

of R CrB forms from the usual supergiant stellar wind, at a considerable ($\sim 100R_*$) and fairly constant distance from the star, so that its color characteristics do not change appreciably. In other words, it is a full analog of the envelopes of supergiants, perhaps with the one exception that the grain-condensation rate at the inner boundary of the dust envelope can vary appreciably. In turn, the clouds may form closer to the star, possibly from gas clouds. Existing models estimate this distance to be from $\sim 2R_*$ [1] to $\sim 20R_*$ [33].

The variations of the grain-condensation rate are obviously connected to variations in the stellar-wind intensity. In contrast to Miras, which display considerable radial pulsations, the changes in the stellar-wind intensity of R CrB must be related to activity on the stellar surface, which may resemble phenomena taking place on the solar surface. As is known, the solar activity has an 11-year cycle; this is probably a coincidence, but the same cycle is detected for the stellar-wind variations of R CrB.

Solar activity is manifest in variations of the number of dark spots and flares. The structure of dark magnetic spots in the atmospheres of R CrB stars was investigated in [34]. There is no doubt of the presence of spots on the solar surface; they are visible with the unaided eye. We do not know whether spots exist on R CrB stars. Soker and Clayton [34] studied the structures of such spots supposing they do exist, but their existence has not been proved. There has been no dicussion of flares.

The cyclic variations in the stellar-wind intensity of R CrB may provide indirect evidence for surface activity, possibly somewhat different from that of the Sun. These phenomena could initiate the formation of dust clouds. Such phenomena could include outflows of material during flares, as occur at the solar surface above sunspot groups. The ejected clump of material could be optically thick at visible wavelengths, so that it blocks the direct stellar radiation, creating conditions for grain condensation near the star. On the other hand, such gas clouds could be the centers of dust formation when they have moved to considerable distances from the star ($r \geq 10R_*$), where grain condensation becomes thermodynamically possible without screening of the direct stellar light.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 02-02-16235.

REFERENCES

1. G. C. Clayton, Publ. Astron. Soc. Pac. **108**, 225 (1996).
2. G. C. Clayton, B. A. Whitney, and J. A. Mattei, Publ. Astron. Soc. Pac. **105**, 832 (1993).
3. Yu. A. Fadeev, Astron. Zh. **75**, 750 (1998) [Astron. Rep. **42**, 662 (1998)].
4. J. D. Fernie and S. Seager, Publ. Astron. Soc. Pac. **106**, 1138 (1994).
5. N. K. Rao, D. L. Lambert, M. T. Adams, *et al.*, Mon. Not. R. Astron. Soc. **310**, 717 (1999).
6. W. A. Stein, J. E. Gaustad, F. C. Gillett, and R. F. Knacke, Astrophys. J. **155**, L3 (1969).
7. F. C. Gillett, D. E. Backman, C. Beichman, and G. Neugebauer, Astrophys. J. **310**, 842 (1986).
8. M. W. Feast, B. S. Carter, G. Roberts, *et al.*, Mon. Not. R. Astron. Soc. **285**, 317 (1997).
9. R. M. Genet, L. J. Boyd, K. E. Kissell, *et al.*, Publ. Astron. Soc. Pac. **99**, 660 (1987).
10. A. E. Nadzhip, V. I. Shenavrin, and I. G. Tikhonov, Tr. Gos. Astron. Inst. im. P. K. Shternberga **58**, 119 (1986).
11. V. I. Shenavrin, O. G. Taranova, V. I. Moroz, and A. V. Grigor'ev, Astron. Zh. **56**, 1007 (1979) [Sov. Astron. **23**, 567 (1979)].
12. J. D. Fernie, Publ. Astron. Soc. Pac. **94**, 172 (1982).
13. J. D. Fernie, J. R. Percy, and G. Richer, Publ. Astron. Soc. Pac. **98**, 605 (1986).
14. M. Asplund, B. Gustafsson, D. Kiselman, and K. Eriksson, Astron. Astrophys. **318**, 521 (1997).
15. N. K. Rao and K. Nandy, Mon. Not. R. Astron. Soc. **222**, 357 (1986).
16. W. J. Forrest, F. C. Gillett, and W. A. Stein, Astrophys. J. **178**, L129 (1972).
17. M. Asplund, private communication.
18. M. Asplund, B. Gustafsson, and D. A. Lambert, Astron. Astrophys. **353**, 287 (2000).
19. C. Jager, H. Mutschke, and T. Henning, Astron. Astrophys. **332**, 291 (1998).
20. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).
21. M. P. Egan, C. M. Leung, and G. F. Spagna, Comput. Phys. Commun. **48**, 271 (1988).
22. A. M. Tatarnikov, V. I. Shenavrin, and B. F. Yudin, Astron. Zh. **75**, 428 (1998) [Astron. Rep. **42**, 377 (1998)].
23. V. M. Larionov, private communication.
24. J. B. Alexander, P. J. Andrews, R. M. Catchpole, *et al.*, Mon. Not. R. Astron. Soc. **158**, 305 (1972).
25. P. L. Cottrell, W. A. Lawson, and M. Buchhorn, Mon. Not. R. Astron. Soc. **244**, 149 (1990).

26. D. Reimers, in *Problems in Stellar Atmospheres and Envelopes,* Ed. by B. Baschek, W. H. Kegel, and G. Traving (Springer-Verlag, Berlin, 1975), p. 229.
27. P. Woitke, A. Goeres, and E. Sedlmayr, Astron. Astrophys. **313**, 217 (1996).
28. D. W. Strecker, Astron. J. **80**, 451 (1975).
29. J. Jurcsik, Acta Astron. **46**, 325 (1996).
30. A. E. Rosenbush, Inf. Bull. Var. Stars, No. 5025, 1 (2001).
31. A. M. Tatarnikov, V. I. Shenavrin, B. F. Yudin, *et al.*, Pis'ma Astron. Zh. **26**, 587 (2000) [Astron. Lett. **26**, 506 (2000)].
32. V. I. Shenavrin and B. F. Yudin, Astron. Zh. **78**, 729 (2001) [Astron. Rep. **45**, 631 (2001)].
33. Y. A. Fadeyev, Mon. Not. R. Astron. Soc. **233**, 65 (1988).
34. N. Soker and G. C. Clayton, Mon. Not. R. Astron. Soc. **307**, 993 (1999).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 819–829. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 909–920.
Original Russian Text Copyright © 2002 by A. Boyarchuk, Pakhomov, Antipova, M Boyarchuk.

Analysis of Chemical Abundances in the Atmospheres of Moderate Barium Stars

A. A. Boyarchuk , Yu. V. Pakhomov, L. I. Antipova, and M. E. Boyarchuk

Institute of Astronomy, ul. Pyatnitskaya 48, Moscow, 119017 Russia

Received February 1, 2002

Abstract—We used high-resolution spectra to compute model atmospheres to derive the atmospheric abundances of moderate barium stars. Comparing our results with analogous data for normal red giants, we find that the moderate barium stars appear to not differ systematically from normal red giants. Their chemical abundance anomalies show the same patterns and can be interpreted in terms of evolutionary effects: the evolutionary stage, mass, luminosity, and metallicity of the objects. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Red giants represent a heterogeneous group of stars. Stars that have reached the giant stage in their evolution generally move along the giant branch of the Hertzsprung–Russell diagram, forming loops in the diagram due to the activation of different nuclear sources. The result is that, in contrast to the main sequence, where each point region corresponds to a concentration of stars with similar masses and ages, a region of the giant branch can be populated by stars with rather different masses and ages and, consequently, with different chemical abundances in their atmospheres.

In our paper [1], we studied the chemical abundances of red giants in the young Hyades cluster, and found the abundances of heavy elements (s-process elements) to be solar. On the other hand, similar studies of normal field red giants (with no spectral peculiarities), which are presumably much older than the Hyades, reveal appreciable excesses of s-process elements in their atmospheres compared to the Sun [2]. In the current paper, we study several field giants that differ from normal giants in the presence of stronger barium lines in their spectra.

Barium stars were first recognized as a special class by Bidelman and Keenan in 1951 [3] in their development of a two-dimensional spectral classification system. Their analysis of G–K stars using low-dispersion spectrograms (76 Å per mm in the photographic range) revealed objects with an abnormally strong (compared to other stars of the same spectral type) λ 4554 Å ionized barium line, whose intensity was comparable to that in the spectra of S stars. In addition, the $SrII$ ($\lambda\lambda$ 4077, 4215 Å) lines and CH, C_2, and CN molecular bands of these objects are strong and have intensities characteristic of carbon stars rather than G–K giants. Bidelman and Keenan [3] concluded that the spectral properties of these objects did not fit the sequence of spectral types and luminosity classes for normal stars and assigned them to a special group—BaII stars, or barium stars.

The first high-dispersion spectroscopic observations of barium stars provided interesting information on the abundances of various elements in their atmospheres. For example, it turned out [4, 5] that, in addition to strong resonance BaII and SrII lines, the spectra of barium stars also display lines of heavier rare-earth elements, which are generally weaker than BaII and SrII lines in the spectra of late-type stars, so that they could not be detected in low-dispersion spectrograms.

It was also noted that the lines of moderate-atomic-weight elements (from Ca to the iron-group elements) correspond to "normal" spectra that can be described in terms of the MK-system standard temperature sequence. The spectral classifications of 20 barium stars were presented in [6] based on the above two properties. Each star was assigned a spectral type according to the MK criteria using the line intensity ratios for elements with moderate atomic weights, and a "barium degree," i.e., the degree of enhancement of lines of BaII and heavier elements, with the Ba-line intensity given on a scale from one to five.

Later, analyses of low-dispersion ($\sim$ 80 Å per mm) spectrograms identified stars with moderately strong BaII lines, intermediate between those for normal giants and BaII stars. In the spectral classification of Morgan and Keenam [7], such objects are assigned the designation Ba0 stars and are called moderate

1063-7729/02/4610-0819$22.00 © 2002 MAIK "Nauka/Interperiodica"

barium stars, in contrast to the classical barium stars (BaII stars). For such stars, the degree of BaII enhancement is assigned using a scale from zero to one. Thus barium stars are considered to form two groups: classical and moderate barium stars.

Considerable attention has been paid to these objects, and numerous studies aimed at investigating their nature have provided many interesting results. The number of known barium stars has also increased: whereas only five such objects were presented in [3], a recent list of classical and moderate barium stars [8] contains about 400 objects. However, the number ratio of barium stars to normal cool giants remains low, only 0.5–1% [9].

Quantitative, high-resolution spectroscopic studies have shown (see, for instance, [10, 11]) that, compared to normal G and K giants, the atmospheres of classical barium stars demonstrate a moderate excess of carbon (by approximately 0.3 dex), approximately normal nitrogen and oxygen abundances, and anomalously high abundances of heavy elements (heavier than Fe) formed in the *s* process. Such abundance anomalies are expected for stars on the asymptotic giant branch (AGB), when hydrogen and helium burn in shell sources and the third deep dredge-up occurs. However, the problem is that the luminosities of classical barium stars are lower than expected for AGB stars during this phase of their evolution [14], so that the observed abundance anomalies cannot be associated with this stage.

The discovery of binarity in barium stars became key to understanding their nature. Many-year radial-velocity observations have shown that virtually all stars with barium degrees above Ba1 according to [6] display radial-velocity variations and form wide pairs [12–15]. According to current theories of stellar evolution, the primary stars in binaries on the AGB experience intense mass loss and contaminate their companion's atmosphere with the products of helium burning and the *s*-process; as a result, the companion acquires the properties of a barium star, while the evolved star loses its envelope and becomes a white dwarf. It follows from mass-function estimates that, if the mass of the new primary that has accreted matter of the ejected envelope is 1.5 $M_\odot$, its companion's mass should be 0.2–0.6 $M_\odot$ [8], so that the companions should indeed be a white dwarfs. These white dwarfs should be observable in the ultraviolet (UV), where they are brighter than the more massive new primaries. In fact, searches for UV radiation exceeding the contribution of the red giant have been successful, demonstrating the presence of white-dwarf companions associated with several barium stars [16–18].

Studies of atmospheric abundances for moderate barium stars are not as numerous as those for classical barium stars, but have resulted in several firm conclusions:

(1) in contrast to classical barium stars, which display overabundances of carbon, the C abundances for moderate barium stars do not differ from those for normal red giants [19, 20];

(2) the abundances of *s*-process elements are higher than those for normal giants but lower than those for classical barium stars [21, 22].

It is interesting that, according to [12], the issue of binarity is not as clear for moderate barium stars as it is for classical barium stars; many show no radial-velocity variations. It is also claimed that the percentage of binaries among stars showing no enhancement of *s*-process elements does not differ from the value for normal G and K giants.

We included moderate and classical barium stars in our studies of atmospheric abundances of red giants belonging to different subgroups with specific spectroscopic properties. These studies are based on homogeneous, high-quality spectroscopic material and the use of a single technique for determining the abundances. Our hope is that comparative analyses of data for different spectral subgroups will provide trustworthy information on the effects of stellar evolution in various phases. The current paper is devoted to our results for moderate barium stars.

2. OBSERVATIONS AND REDUCTION

A list of our program stars is presented in Table 1. We selected five moderate barium stars for our study, as well as two normal red giants (ϵ Vir and κ Oph) as comparison stars. The spectral types for each of the stars are taken from [23]. Our spectroscopic observations were acquired in 1996–2001 with the 2.6-m Shain telescope of the Crimean Astrophysical Observatory. The light detector was a CCD camera installed in the first compartment of the diffraction spectrograph, at the Coudé focus. The dispersion was 3 Å per mm, the spectral resolution was 0.05 Å per pixel, and all the spectrograms have signal-to-noise ratios S/N > 100. The data reduction was performed using software developed at the Crimean Astrophysical Observatory. The technique used to measure the equivalent widths and derive atmospheric parameters is discussed in detail in [2]. The measured equivalent widths can be found at our web site, www.inasan.rssi.ru.

Our spectral analysis employed stellar model-atmosphere grids computed with the ATLAS9 code [24]. The Sun was chosen as the comparison star; when computing relative chemical abundances, we

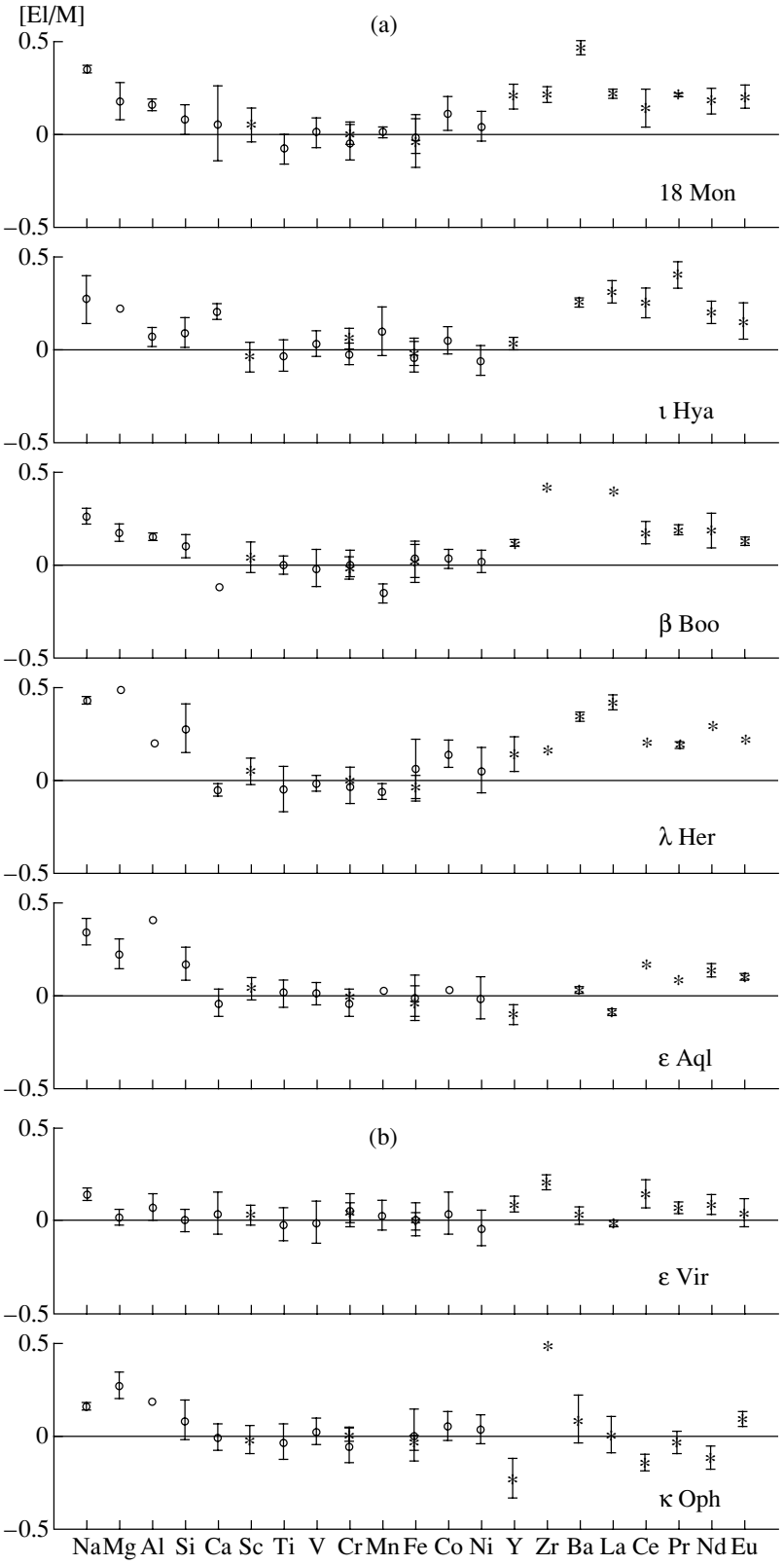

Fig. 1. Chemical abundances derived from spectra of (a) moderate barium stars and (b) normal red giants, relative to the solar abundances. The circles and asterisks show abundances derived from lines of neutral atoms and ions, respectively. Error bars derived from the dispersion of values for individual lines are shown. If there is no error bar, the corresponding abundance was derived from a single line.

Table 1. Program stars and their adopted atmospheric parameters

No.	HD	Star	Spectrum	$T_{\rm eff}$, K	$\log g$	ξ, km/s
			Moderate barium stars			
1	49293	18 Mon	K0+IIIa Ba0.2	4717	2.45	1.42
2	83618	ι Hya	K2.5III-IIIb Ba0.3	4355	1.95	1.47
3	133208	β Boo	G8IIIa Ba0.3 Fe-0.5	5075	2.68	1.54
4	158899	λ Her	K3.5III Ba0.1	4180	1.57	1.34
5	176411	ϵ Aql	K1-III CN0.5 Ba0.2	4691	2.35	1.39
			Normal red giants			
6	113226	ϵ Vir	G8IIIab	5187	3.20	1.41
7	153210	κ Oph	K2III	4593	2.52	1.14

used solar abundances for the same lines computed for a Kurucz model with $T_{\rm eff} = 5887$ K, $\log g = 4.57$, and $V_t = 0.80$ km/s, which was demonstrated in [25] to provide the best representation of the solar spectrum. The derived atmospheric parameters for the program stars are collected in Table 1, and our estimates of the abundances of 22 elements computed using the corresponding model atmospheres are presented in Table 2 and Fig. 1. Table 2 indicates that the program stars have different metallicities. Figure 1 shows the abundances of various elements relative to the metallicity, i.e., the [El/M] values.

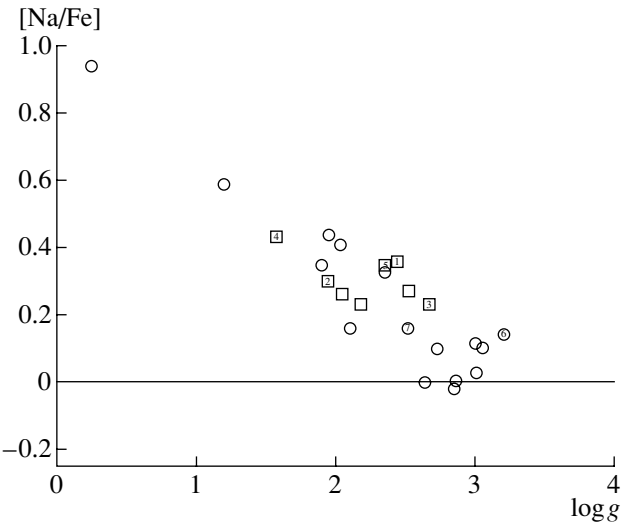

Fig. 2. Excess sodium-to-iron abundance ratios in the atmospheres of red giants compared to the solar values. The circles and squares show data for normal giants and moderate barium stars. The stars are numbered according to Table 1.

3. ANALYSIS OF THE RESULTS

First, note that our technique for deriving the stellar atmospheric parameters is based on the analysis of absorption lines of iron-group elements [2]. According to [6], the relative abundances of these elements in the spectra of barium stars do not differ from the solar values, so that our estimates of the atmospheric parameters should provide a good description of the stellar line spectra, making our further analysis of the atmospheric abundances reliable.

It follows from Table 2 and Fig. 1 that the program stars exhibit different overabundances of Na, Al, and Si. A similar result was earlier obtained for normal red giants, as discussed in [2]. It was demonstrated that the abundance excess for each of these elements increased with the luminosity of the star and that, in the case of Na, giants obeyed a relation similar to that derived earlier for F supergiants [26–28]. This suggested that the Na overabundances in supergiants and giants had the same origin. Figures 2–4 display the relationships between the overabundances [Na/Fe], [Al/Fe], and [Si/Fe] and the surface gravities from [2] for normal giants and supergiants (without spectral peculiarities). The abundances of these elements in the atmospheres of our moderate barium stars are also plotted (the numbers refer to the numbering of the stars in Table 1), along with values for three stars of this type studied by us earlier [2]. All three figures demonstrate that the moderate barium stars follow the same relations between overabundance and surface gravity as the normal red giants.

Figure 1 also demonstrates that the abundances of s-process elements in the atmospheres of moderate barium stars are, on average, somewhat higher than in the solar atmosphere, despite considerable star-to-star variations. In our earlier abundance analysis for s-process elements in red giants belonging to

Table 2. Elemental abundances in the atmospheres of the program stars relative to the solar abundances

Element	Moderate barium red giants										Normal red giants			
	18 Mon		ι Hya		βBoo		λ Her		ϵ Aql		ϵ Vir		κ Oph	
	n	[El/H]	n	[El/H]	n	[El/H]	n	[El/H]	n	[El/H]	n	[El/H]	n	[El/H]
Na I	2	0.44	2	0.24	2	0.30	2	0.37	2	0.31	2	0.38	2	0.29
Mg I	4	0.27	1	0.19	2	0.21	2	0.43	2	0.19	2	0.25	2	0.40
Al I	2	0.25	2	0.04	2	0.19	1	0.14	1	0.36	2	0.31	1	0.32
Si I	11	0.17	10	0.06	14	0.14	7	0.22	6	0.13	20	0.24	14	0.21
Ca I	4	0.15	3	0.17	1	−0.08	2	−0.11	6	−0.08	5	0.27	3	0.12
Sc II	5	0.14	6	−0.07	6	0.08	6	−0.01	5	0.00	9	0.26	8	0.11
Ti I	28	0.01	25	−0.06	16	0.04	15	−0.11	17	−0.03	36	0.22	29	0.10
V I	6	0.10	7	−0.00	13	0.02	4	−0.08	6	−0.03	24	0.23	7	0.15
Cr I	20	0.04	16	−0.05	11	0.04	8	−0.09	5	−0.08	32	0.29	16	0.08
Cr II	4	0.09	3	0.02	6	0.02	1	−0.06	3	−0.04	8	0.28	3	0.14
Mn I	3	0.10	3	0.07	2	−0.12	2	−0.12	1	−0.01	4	0.26	–	–
Fe I	67	0.08	69	−0.06	40	0.07	43	0.00	52	−0.05	111	0.24	125	0.15
Fe II	6	0.05	8	−0.05	6	0.05	2	−0.11	6	−0.07	12	0.23	6	0.09
Co I	17	0.20	20	0.02	6	0.07	6	0.08	1	−0.01	22	0.27	11	0.18
Ni I	18	0.13	13	−0.09	17	0.05	16	−0.01	16	−0.05	38	0.19	25	0.16
Y II	6	0.29	3	−0.00	3	0.15	3	0.08	2	−0.14	5	0.32	4	−0.11
Zr II	2	0.30	–	–	1	0.45	1	0.10	–	–	2	0.44	1	0.62?
Ba II	2	0.55	2	0.22	–	–	3	0.28	2	−0.01	2	0.26	3	0.21
La II	2	0.31	2	0.28	1	0.43	2	0.36	2	−0.13	2	0.23	3	0.13
Ce II	4	0.23	3	0.22	3	0.21	2	0.15	2	0.13	4	0.38	4	−0.02
Pr II	1	0.30	2	0.37	2	0.22	2	0.12	1	0.04	2	0.30	3	0.09
Nd II	7	0.27	5	0.17	4	0.22	4	0.23	3	0.09	7	0.32	9	0.01
EuII	2	0.29	2	0.12	2	0.16	1	0.17	2	0.06	2	0.27	2	0.21
[M/H]*	$+0.09 \pm 0.06$		-0.03 ± 0.04		$+0.04 \pm 0.02$		-0.06 ± 0.06		-0.04 ± 0.02		$+0.24 \pm 0.03$		$+0.13 \pm 0.03$	

* The mean metallicity is defined as the mean abundance of iron-group elements.

different subgroups [2], we found that slight anomalies of these elements were characteristic not only of moderate barium stars, but also of a number of normal field red giants. A comparison of the s-process element abundances for stars of the two subgroups demonstrated that these abundances were somewhat higher for barium stars than for normal red giants [2, Fig. 5]. Only three moderate barium stars were analyzed in [2]. Accordingly, we have carried out a similar comparison based on all the stars studied here and in [2]. In Fig. 5, similar to Fig. 5 in [2], we plot the average abundances of these elements for 12 normal red giants and 8 moderate barium stars. This figure shows that the abundances of s-process elements are somewhat higher for moderate barium stars than for normal giants, in agreement with the conclusions of [2]; note, however, that the difference between the values formally does not exceed the observational errors (0.10–0.13 dex).

Thus, it follows from Table 2 and Figs. 1–5 that the chemical abundances of the studied moderate barium stars show no significant differences from the abundances of normal red giants.

Let us consider the positions of the studied stars

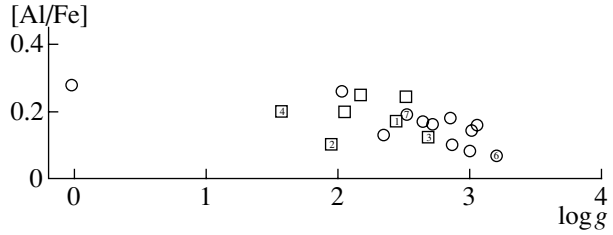

Fig. 3. Same as Fig. 2 for aluminum.

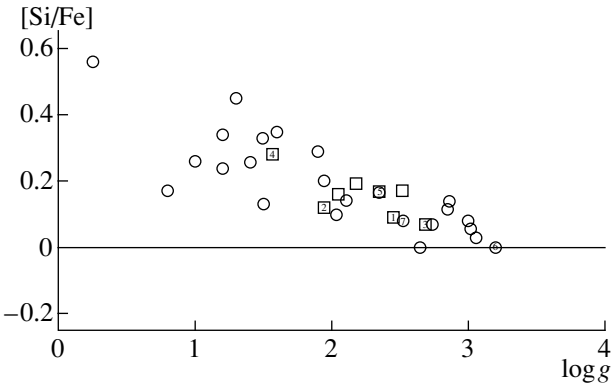

Fig. 4. Same as Fig. 2 for silicon.

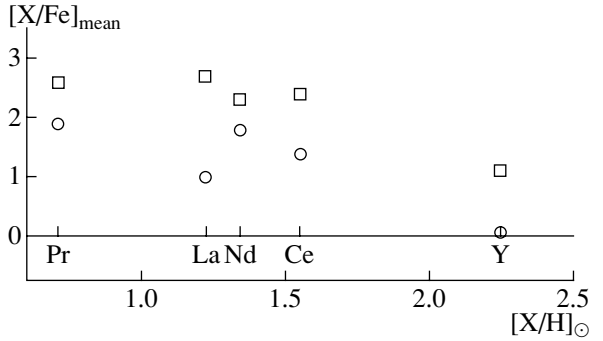

Fig. 5. Excesses of the ratios of the abundances of s-process elements to iron in red giants compared to the Sun. The circles and squares show the mean values for 12 normal giants and 8 moderate barium stars.

in the Hertzsprung–Russell diagram, as shown in Fig. 6. We have plotted evolutionary tracks in accordance with [29], and chosen $\log g$ and $\log T_{\rm eff}$ for the coordinates. For each star, the dots show the positions (i) corresponding to the $\log g$ and $T_{\rm eff}$ derived from our spectral analysis (Table 1) and (ii) corresponding to coordinates computed from photometric and parallax measurements (the intersections of the two error bars); in both cases, the size of the bars corresponds to the accuracy of the data. The two positions for each star are connected with a straight line, and the star's number in Table 1 is indicated.

Let us consider the computation of a star's coordinates in the Hertzsprung–Russell diagram using the second technique in more detail. We employed the parallax data obtained by the HIPPARCOS satellite [30], the apparent magnitudes from the Bright Star Catalogue [23], and the bolometric corrections from [31]. Each star's temperature was derived from the 13-color photometry of [32] as follows. We first plotted a number of color–temperature calibration curves. The calibration was based on data for field stars whose effective temperatures were derived accurately from observations of their infrared fluxes. To do this, we selected 42 stars from those listed in [31] with atmospheric parameters characteristic of the red giants studied, namely, $4000 < T_{\rm eff} < 6000$, $1.5 < \log g < 3.0$, and mean metallicity $[{\rm M}] = -0.07 \pm 0.17$. Using the values of $T_{\rm eff}$ from [31] and the color indices from [32], we plotted a total of 12 color–$T_{\rm eff}$ relations; Fig. 7 shows the C(52–99)–$T_{\rm eff}$ calibration relation as an example. Each of the relations was approximated with a third-order polynomial. The mean temperature error for the approximation was about 50 K. The polynomial coefficients are collected in Table 3. A star's effective temperature can be calculated using the formula

$$T_{\rm eff} = A_0 + A_1 C + A_2 C^2 + A_3 C^3, \qquad (1)$$

where C is the value of a color index from [32] and A_1 are the corresponding approximation coefficients from Table 3. We determined the effective temperatures of our program stars using all 12 calibration curves. The resulting temperatures for a given color index along with its average value and the scatter for each of the stars are given in Table 4.

Figure 6 shows that, as a rule, a star's position according to its spectroscopic coordinates is to the left of its photometric position. This means that the effective temperature derived spectroscopically is almost always somewhat higher than the photometric estimate. We came to a similar conclusion about the Sun in [33], where we also noted that a single effective temperature could not represent equally well the solar continuum energy distribution and solar line spectrum, possibly due to imperfections in the model atmospheres (some factors not taken into account in their computation). Thus, it appears that this conclusion is also relevant for red giants. Figure 6 also shows that the spectroscopic $\log g$ values are lower than the corresponding parallax values. This has been known for a long time, and has been extensively discussed in the literature (see, for example, [34] and references therein). The origin for this behavior could likewise be imperfections in the model atmospheres.

Note, however, that slight deviations of a star's position in the Hertzsprung–Russell diagram when

Table 3. Coefficients of the color index–temperature approximations based on the 13-color photometry of [32]

Color index	A_0	A_1	A_2	A_3
33–52	7005.9	−2369.4	810.59	−113.84
35–52	6752.9	−2368.9	922.06	−152.58
37–52	6997.1	−2550.4	965.23	−154.24
40–52	9198.8	−6404.2	3081.1	−561.11
45–52	19832	−76291	130760	−77990
52–58	12461	−34928	53440	−29022
52–63	12615	−21427	19698	−6531.8
52–72	13685	−19217	13948	−3588.9
52–80	13921	−15798	9135.5	−1872.5
52–86	13180	−13215	6950.9	−1305.7
52–99	12169	−10293	4798.5	−807.07
52–110	11310	−7835.3	3124.5	−457.97

Table 4. Temperatures derived using the color indices of [32] and Eq. (1)

Color index	18 Mon	ι Hya	β Boo	λ Her	ϵ Aql	ϵ Vir	κ Oph
33–52	4611	4338	4897	4139	4631	4944	4550
35–52	4607	4353	4894	4141	4627	4930	4527
37–52	4630	4342	4950	4140	4648	4948	4539
40–52	4596	4321	4919	4147	4670	4968	4551
45–52	4555	4330	4849	4220	4747	4975	4644
52–58	4679	4374	5000	4208	4691	4966	4553
52–63	4690	4298	4977	4167	4656	5007	4558
52–72	4673	4342	4990	4197	4726	4972	4550
52–80	4675	4333	4960	4208	4773	4991	4571
52–86	4685	4324	4956	4200	4750	5007	4567
52–99	4674	4325	4968	4193	4744	5026	4585
52 -110	4634	4288	5036	4188	4743	5024	4616
$T_{\rm eff}$, K	4642	4331	4950	4179	4700	4980	4567
$\sigma T_{\rm eff}$, K	41	21	49	29	50	30	31

using atmospheric parameters obtained using different techniques does not impede qualitative comparisons of star-to-star chemical abundances. Table 2 indicates that the largest excesses of s-process elements are found for ι *Hya* and λ Her (Stars 2, 4 in Table 1), which, according to Fig. 6, are in a later stage of evolution than the remaining the stars. A comparison of the abundances of β Boo, 18 Mon, and λ Her (stars 3, 1, and 4, respectively), which have approximately the same mass (about 3 $M_\odot$), also shows that the excess of s-process elements increases toward later stages of evolution. In [35], our comparison of the abundances of s-process elements for four stars with approximately equal temperatures

Table 5. Metallicities and abundances of CNO-cycle elements in normal red giants and moderate barium stars according to [36–46].

	[Fe/H]	$^{12}C/^{13}C$	[C/Fe]	C/N
Normal giants (53 stars)	-0.10 ± 0.21	17.2 ± 8	-0.19 ± 0.15	1.72 ± 0.8
Our 12 stars	-0.00 ± 0.13	17.0 ± 6	-0.22 ± 0.14	1.62 ± 0.6
Moderate barium stars (20 stars)	-0.09 ± 0.15	18.7 ± 9	-0.30 ± 0.20	1.10 ± 0.5
Our 7 stars	-0.02 ± 0.15	19.0 ± 4	-0.37 ± 0.23	0.91 ± 0.5

but different masses indicated that the overabundance of s-process elements increased with mass. The origin of this tendency is probably that the convection responsible for the dredge-up of products of nuclear reactions in the stellar interiors to the star's atmosphere is stronger in higher-mass stars, and can reach deeper layers. We reach a similar conclusion based on a comparison of the s-process abundances in the atmospheres of stars with similar temperatures (Table 1, Fig. 6): β Boo, whose mass exceeds that of ϵ Vir, also displays a higher excess of s-process elements.

Thus, we conclude that the overabundances of s-process elements in the studied stars are probably influenced by several parameters: the star's evolutionary stage (age), mass, and, possibly, metallicity [33].

We can derive information about the evolutionary stage of the stars from their atmospheric abundances of light elements: C, N, and O. For example, a star's ascent to the giant branch is characterized by the development of a convective envelope and the dredge-up of products of the CNO-cycle to the stellar surface, accompanied by a decrease of the carbon abundance and increase in the nitrogen abundance in the star's atmosphere (compared to the solar value). Thus, the C/N abundance ratio is sensitive to this process (it decreases). The later evolutionary stage of shell burning of helium and hydrogen is also accompanied by convective mixing and the dredge-up of the products of helium burning to the atmosphere, so that this phase is characterized by an increased abundance of carbon, compared to red giants undergoing the first stage of convective mixing. This is observed for M and S giants on the AGB and for classical barium stars: in the binary hypothesis, this is explained by the addition to the giants' atmospheres of matter from the envelope ejected by their companions, which have evolved into white dwarfs and are thus enriched in the products of helium burning. Different amounts of added material will result in different strengths of the "barium features" characterizing the barium subtype.

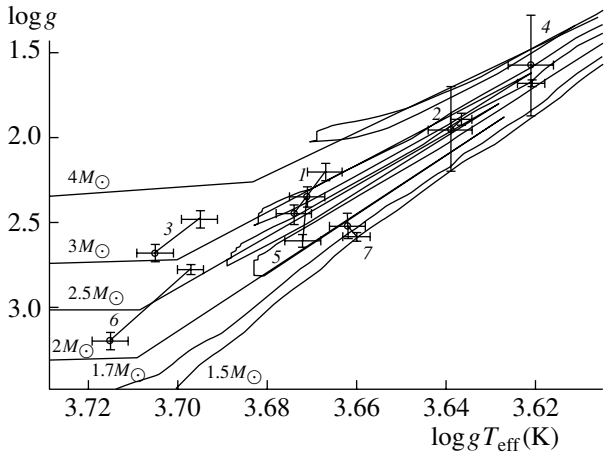

Fig. 6. Hertzsprung–Russell diagram for the studied stars (see text for details).

Determination of the C, N, and O abundances is beyond the scope of our present study. Therefore, we attempted to compare the abundances of these elements in moderate barium stars and normal red giants using data from the literature. We found such data for 53 normal giants (including the 12 stars studied by us previously) and 20 moderate barium stars (7 of them our program stars) in [19, 20, 36–44]. Table 5 presents mean abundance ratios and metallicities calculated using the data of [19, 20, 36–44], separately for all stars of a given subgroup and for the stars considered in this paper and in [2]. The similarity of the metallicities for all the stars testifies that all of them belong to the same Galactic population, namely the thin disk. It also follows from Table 5 that, within the errors, the abundances for our program stars coincide with those for all stars of the corresponding subgroup; i.e., the stars studied here are typical of their subgroups. In addition, the atmospheric carbon abundance in the moderate carbon stars is not enhanced compared to the value for normal giants, in contrast to the case of classical barium stars, which exhibit carbon excesses. This is in agreement with the results obtained for the same stars in [19, 20]. Note also that although the [C/Fe] abundances for the moderate barium stars and normal red giants coincide within the errors, this abundance ratio is still somewhat lower for the barium stars. We

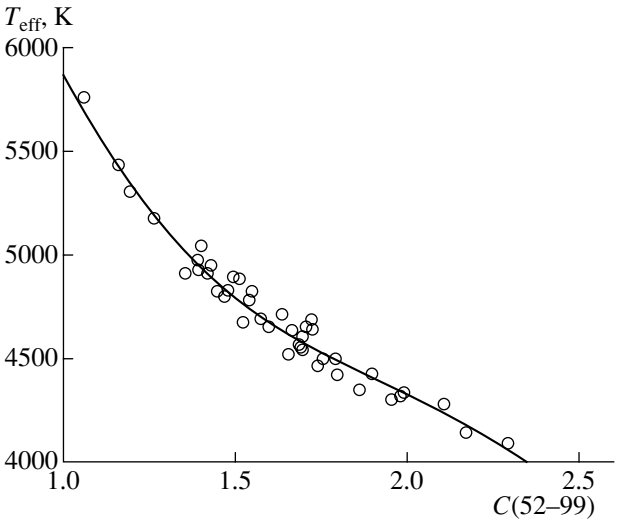

Fig. 7. Example of an approximate color–effective temperature relation using the C(52–99) color index of the 13-color photometric measurements of [32].

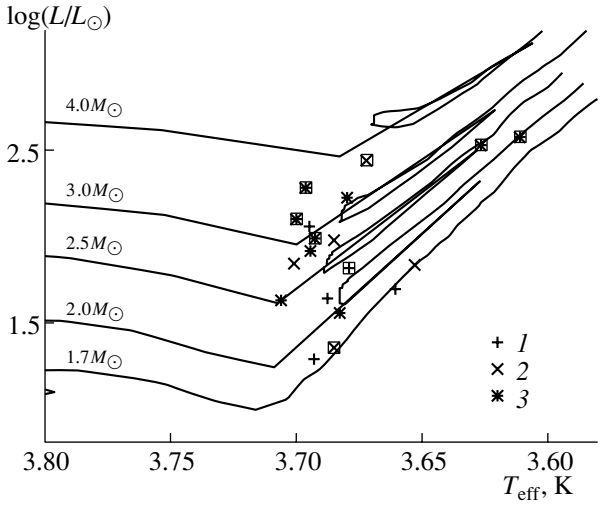

Fig. 8. Positions of red giants with different s-process anomalies in their atmospheres in the Hertzsprung–Russell diagram, according to this paper and [2]. (*1*) $[s/\mathrm{Fe}] < 0.12$ dex; (*2*) $0.12 < [s/\mathrm{Fe}] < 0.23$ dex; (*3*) $[s/\mathrm{Fe}] > 0.23$ dex. The squares show stars identified as moderate barium stars.

can explain this fact, as well as their simultaneously somewhat lower C/N ratios, if we assume that moderate barium stars experience deeper convective mixing and, consequently, a larger dredge-up of CNO-cycle products compared to normal red giants. The agreement of the carbon isotope ratios, $^{12}C/^{13}C$, for normal red giants and moderate barium stars is consistent with the hypothesis that the latter stars experience more developed convection. Detailed studies of the red giant phase show that this carbon isotope ratio decreases sharply as soon as the convective envelope reaches the layer enriched in ^{13}C and transports this material to the stellar surface but remains virtually unchanged during the remainder of the red giant phase, as the convection reaches deeper layers [45]. Thus, the abundances of CNO-cycle elements in the atmospheres of moderate barium stars provide evidence that this cycle occurs in the interiors of these stars, and do not support the binary hypothesis for moderate barium stars.

Finally, let us make the following remark. The classification of barium stars is based primarily on the intensity of the BaII λ 4554 Å line. Moderate barium stars are defined as those intermediate between normal giants and classical barium stars, and display only slightly enhanced intensities of this barium line. Note that the line originates from the ground level and is very strong, so that appreciable non-LTE effects are expected. In addition, since ion lines in the spectra of red giants are sensitive to the value of $\log g$, this line's intensity can differ somewhat for stars with different surface gravities. This is probably the main reason why a star's classification as a moderate barium star is not unambiguous, so that different authors have sometimes classified the same stars as normal red giants or moderate barium stars. For example, the list of barium stars in [8], which includes both classical and moderate barium stars, contains more than four hundred stars; however, it does not include about a dozen fairly bright stars in the catalog [23] described as moderate barium stars with barium degrees from 0.2 to 0.4. In this context, it is also interesting to compare the two versions of the Bright Star Catalog, [23] and [46]: some moderate barium stars in the earlier version are not indicated as having enhanced barium lines in the later version [46], whereas several stars are classified as moderate barium stars in [46], although they were considered normal red giants in [23]. We conclude that some red giants have been assigned to the class of moderate barium stars when the intensity of their BaII λ 4554 Å line was somewhat above the mean value, possibly because of their higher luminosity (or lower $\log g$) due to their more advanced evolutionary stage (associated with either their ages or a higher rate of evolution for more massive stars).

Since normal red giants and moderate barium stars both display overabundances of s-process elements, it is of interest to consider the positions in the Hertzsprung–Russell diagram of both groups of stars, whose atmospheres have different s-process element excesses. We carried out such a comparison for the stars studied here and in [2]; thus, we considered only stars whose atmospheric abundances were determined using the same technique and uniform spectroscopic material, so that the comparison should yield trustworthy results. Figure 8 shows

a Hertzsprung–Russell diagram plotted in accordance with the computations of [29]. The positions of stars are based on their parallaxes and photometric measurements (see above). We subdivided the stars into three groups based on the overabundances of *s*-process elements we derived (see the figure captions). The moderate barium stars are plotted as squares. This figure shows that normal red giants and moderate barium stars occupy the same region of the diagram. Stars with higher excesses of *s*-process elements tend to occupy positions near tracks corresponding to higher masses or to be located higher on the red giant branch, providing evidence for an evolutionary origin of the anomalous abundances of *s*-process elements in their atmospheres.

Thus we conclude that moderate barium stars with barium degrees from 0.1 to 0.8 probably do not differ systematically from normal red giants. The abundance anomalies for these two types of star follow the same relations, and can be explained in terms of evolutionary effects (the stage of evolution, mass, luminosity, and metallicity of the object in question), so that the binary model invoked for classical barium stars is not needed. Both the normal red giants and the moderate barium stars occupy the same region of the Hertsprung–Russell diagram. For some stars, the effects of evolution have made their *s*-process element abundances higher, and such stars have been described as barium stars with moderate "barium degrees" in various earlier works on spectral classification.

ACKNOWLEDGMENTS

This study was partially supported by the Russian Foundation for Basic Research (project no. 00-02-17660) and by a grant of the President of the Russian Federation (project no. 00-15-96722). Observations for several stars were acquired using an SDC-9000 Photometrics GmbH CCD chip installed on the Crimean Astrophysical Observatory's 2.6-m telescope thanks to support from the International Science Foundation (grant nos. R2Q000 and U1C000) and the ESO C&EE Program (grant no. A-05-067).

REFERENCES

1. A. A. Boyarchuk, L. I. Antipova, M. E. Boyarchuk, and I. S. Savanov, Astron. Zh. **77**, 96 (2000) [Astron. Rep. **44**, 76 (2000)].
2. A. A. Boyarchuk, L. I. Antipova, M. E. Boyarchuk, and I. S. Savanov, Astron. Zh. **78**, 349 (2001) [Astron. Rep. **45**, 301 (2001)].
3. W. P. Bidelman and P. C. Keenan, Astron. J. **114**, 473 (1951).
4. R. H. Garstang, Publ. Astron. Soc. Pac. **64**, 227 (1952).
5. E. M. Burbidge and G. R. Burbidge, Astrophys. J. **126**, 357 (1957).
6. B. Warner, Mon. Not. R. Astron. Soc. **129**, 263 (1965).
7. W. W. Morgan and P. C. Keenan, Ann. Rev. Astron. Astrophys. **11**, 29 (1973).
8. P. K. Lu, Astron. J. **101**, 2229 (1991).
9. W. P. Bidelman, in *Cool Stars with Excess of Heavy Elements,* Ed. by M. Jaschek and P. C. Keenan (Reidel, Dordrecht, 1985), p. 43.
10. V. V. Smith, Astron. Astrophys. **132**, 326 (1984).
11. C. Sneden, D. L. Lambert, and C. A. Pilachowski, Astrophys. J. **247**, 1052 (1981).
12. R. D. McClure, in *Cool Stars with Excess of Heavy Elements,* Ed. by M. Jaschek and P. C. Keenan (Reidel, Dordrecht, 1985), p. 315.
13. R. D. McClure, J. M. Fletcher, and J. M. Nemec, Astrophys. J. **238**, L35 (1980).
14. R. D. McClure, Astrophys. J. **268**, 264 (1983).
15. R. D. McClure and A. W. Woodsworth, Astrophys. J. **352**, 709 (1990).
16. J. F. Dominy and D. L. Lambert, Astrophys. J. **270**, 180 (1983).
17. E. Bohm-Vitense, J. Nemec, and Ch. Proffitt, Astrophys. J. **278**, 726 (1984).
18. E. Bohm-Vitense, K. Carpenter, R. Robinson, *et al.*, Astrophys. J. **533**, 969 (2000).
19. C. Sneden, C. A. Pilachowski, and D. L. Lambert, Astrophys. J. **247**, 1052 (1981).
20. S. V. Berdyugina, Pis'ma Astron. Zh. **19**, 933 (1993) [Astron. Lett. **19**, 378 (1993)].
21. C. A. Pilachowski, Astron. Astrophys. **54**, 465 (1977).
22. J. Tomkin and D. L. Lambert, Astrophys. J. **311**, 819 (1986).
23. D. Hoffleit and C. Jaschek, *The Bright Star Catalogue* (Yale University Observatory, New Haven, 1982, 4th ed.).
24. R. L. Kurucz, Rev. Mex. Astron. Astrofis. **23**, 181 (1992).
25. A. A. Boyarchuk, L. I. Antipova, M. E. Boyarchuk, and I. S. Savanov, Astron. Zh. **75** (4), 586 (1998) [Astron. Rep. **42**, 517 (1998)].
26. A. A. Boyarchuk and M. E. Boyarchuk, Izv. Krym. Astrofiz. Obs. **63**, 66 (1981).
27. A. A. Boyarchuk and L. S. Lyubimkov, Izv. Krym. Astrofiz. Obs. **64**, 3 (1981).
28. A. A. Boyarchuk and L. S. Lyubimkov, Izv. Krym. Astrofiz. Obs. **66**, 130 (1983).
29. G.Schaller, D. Schaerer, G. Meynet, and A. Maeder, Astron. Astrophys., Suppl. Ser. **96**, 269 (1992).
30. The HIPPARCOS Catalog ESA SP-1200 (1997).
31. M. L. Houdashelt, R. A. Bell, and A. Sweigart, Astron. J. **119**, 1448 (2000).
32. H. L. Johnson and R. I. Mitchell, Rev. Mex. Astron. Astrofis. **1**, 299 (1975).
33. A. A. Boyarchuk, L. I. Antipova, M. E. Boyarchuk, and I. S. Savanov, Astron. Zh. **75** (4), 586 (1998) [Astron. Rep. **42**, 517 (1998)].
34. R. E. Luck and G. G. Wepfer, Astron. J. **110**, 2425 (1995).

35. L. I. Antipova and A. A. Boyarchuk, Astron. Zh. **78**, 807 (2001) [Astron. Rep. **45**, 700 (2001)].
36. J. Tomkin and D. L. Lambert, Astrophys. J. **193**, 631 (1974).
37. J. Tomkin, D. L. Lambert, and R. E. Luck, Astrophys. J. **199**, 436 (1975).
38. R. W. Day, D. L. Lambert, and C. Sneden, Astrophys. J. **185**, 213 (1973).
39. R. G. Gratton, Astron. Astrophys. **148**, 105 (1985).
40. D. L. Lambert and L. M. Ries, Astrophys. J. **248**, 228 (1981).
41. P. Kjergaard, B. Gustafsson, G. A. H. Walker, and L. Hultqvist, Astron. Astrophys. **115**, 145 (1982).
42. V. V. Smith and N. B. Suntzeff, Astron. J. **97**, 1699 (1989).
43. D. L. Lambert, J. F. Doming, and S. Sivertsen, Astrophys. J. **235**, 114 (1980).
44. J. Bonnell and R. A. Bell, Mon. Not. R. Astron. Soc. **201**, 253 (1982).
45. A. V. Sweigart, L. Greggio, and A. Renzini, Astrophys. J. **364**, 527 (1990).
46. D. Hoffleit and W. H. Warren, Jr., *The Bright Star Catalogue* (Astronomical Data Center, 1991, 5th ed.), Preliminary Version.

Translated by N. Samus'

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 830–839. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 921–931.
Original Russian Text Copyright © 2002 by Baushev.

The Cyclotron Spectrum of Anisotropic Ultrarelativistic Electrons: Interpretation of X-ray Pulsar Spectra

A. N. Baushev

Space Research Institute, Russian Academy of Sciences, Profsoyuznaya ul. 84/32, Moscow, 117810 Russia

Received September 16, 2001; in final form, May 23, 2002

Abstract—The spectrum of cyclotron radiation produced by electrons with a strongly anisotropic velocity distribution is calculated taking into account higher harmonics. The motion of the electrons is assumed to be ultrarelativistic along the magnetic field and nonrelativistic across the field. One characteristic feature of the resulting spectrum is that harmonics of various orders are not equally spaced. The physical properties and observed spectra of four X-ray pulsars displaying higher cyclotron harmonics are analyzed. It was shown that at least in one of them, the cyclotron feature can apparently be only an emission line. Moreover, the observed harmonics are not equidistant, and display certain other properties characteristic of emission by strongly anisotropic ultrarelativistic electrons. In addition, there are indirect theoretical arguments that the electrons giving rise to cyclotron features in the spectra of X-ray pulsars are ultrarelativistic and characterized by strongly anisotropic distributions. As a result, estimates of the magnetic fields of X-ray pulsars (which are usually derived from the energies of cyclotron lines) and certain other physical parameters require substantial revision. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

A cyclotron line in the spectrum of an X-ray pulsar was first identified for the object Her X-1 in 1977 [1]. Reports of observations of higher cyclotron harmonics in the spectra of several X-ray pulsars in binary systems (X0115+63 [2], Vela X-1 [3, 4], 4U1907+09 [5, 6], and A0535+26 [7]) were published later.[1] Two cyclotron harmonics were also observed in the spectrum of 1E 2259+586 [8]; however, the nature of this object remains unknown and will not be discussed here. At least three harmonics have been detected in the spectrum of X0115+63 [9, 10]. The presence of higher cyclotron harmonics seems to be established most firmly for this source. Unfortunately, the situation for the other objects is less clear, since only weak second harmonics have been detected in their spectra. Nevertheless, even these first observations are very interesting, since they can be used to address many theoretical problems connected with the radiation of X-ray pulsars. The aim of the present paper is to consider the physical conditions for the formation of higher cyclotron harmonics in pulsar spectra. We find that a number of problems related to the structure of the emitting region of these sources that have remained poorly understood can be easily solved. Here is a list of such problems, which will be discussed in more detail below.

First, it remains unknown if the cyclotron features in the spectra of many X-ray pulsars are absorption or emission lines. As we show below, at least in one of them (in X0115+63), the lines are most likely emission lines.

Second, the velocity distribution of the electrons emitting the cyclotron lines is not entirely clear. We can reasonably assume that the velocity of their motion across the magnetic field is fairly low, i.e., weakly relativistic (or even nonrelativistic). Otherwise, the cyclotron line would have the characteristic synchrotron shape, with a large number of harmonics forming a quasi-continuum. In fact, we can usually see only one (the fundamental) harmonic. A few exceptions are listed above, but these sources also possess moderate transverse velocities. The electron velocities along the magnetic field are unclear. It is usually assumed that these velocity are small, considerably less than c. Nevertheless, it is possible that the electron-velocity distributions are strongly anisotropic, so that the electron velocities are weakly relativistic across the magnetic field and ultrarelativistic along the field. As was shown in the analysis of the spectrum of the X-ray pulsar Her X-1 in [11], there are strong arguments suggesting that the electrons forming the cyclotron line are ultrarelativistic, and that their velocities are characterized by

[1] Only references to the first publications for each source are given here.

1063-7729/02/4610-0830$22.00 © 2002 MAIK "Nauka/Interperiodica"

very strong anisotropy.[2] Unfortunately, if only one cyclotron harmonic is observed (as for Her X-1), full verification of this possibility purely through spectral analyses is not possible, so that some additional considerations (the dependence of the spectrum on the phase of the pulsar, and so on) must be used. The situation is completely different if several harmonics are observed. As will be shown below, cyclotron radiation by ultrarelativistic, anisotropic electrons displays very specific features. In particular, harmonics of various orders are not equally spaced. Therefore, by analyzing the spectrum, we can determine unambiguously if the motion of the emitting electrons along the field is ultrarelativistic or not.

We will use the following model for the emitting region ("hot spot") of an X-ray pulsar, put forward in [11]. When the accretion flow approaches the surface of the neutron star, a shock wave is formed at some distance from this surface, and a turbulent region with temperature T_e forms below the shock. The heated surface of the neutron star, characterized by the temperature T_s, is located below the turbulent region. The ultrarelativistic electrons emitting the cyclotron line are produced at the shock front; their optical depth is assumed to be negligible.

2. CALCULATION OF THE RADIATION SPECTRUM IN A COMOVING COORDINATE FRAME

Let us calculate the magnetobremsstrahlung radiation from electrons located in a constant magnetic field and possessing a strongly anisotropic distribution. We assume that the electron velocities along the magnetic-field direction (denoted V) are all equal and that the motion of the electrons along the magnetic field is ultrarelativistic. (The situation for the transverse distribution will be specified below.) Let us use a reference frame moving along the magnetic field with a constant velocity V. Although this coordinate system is not fully associated with the electrons, we will call it the comoving frame, while the initial fixed reference frame will be called the laboratory frame. The electrons have only a transverse velocity component v in the comoving frame. It is assumed in our model that $v \ll c$, and the velocity distribution is of the form[3]

$$dn = Nexp\left(-\frac{m_e v^2}{2T}\right)\, d\,\left(-\frac{m_e v^2}{2T}\right). \qquad (1)$$

In other words, this is a nonrelativistic, two-dimensional, Maxwellian distribution with temperature T, where $T \ll m_e c^2$. Let us first find the radiation field in the comoving frame. Further, we will be able to derive the distribution of radiation in the laboratory frame by performing a Lorentz transformation.

According to [12], the emission of each harmonic in the nonrelativistic case occurs at the fixed frequency

$$\omega_n = n\omega_H, \qquad (2)$$

where $\omega_H = \frac{eH}{m_e c}$ is the Larmor frequency of an electron and the intensity of the radiation at the nth harmonic emitted by a particle moving with velocity $v \ll c$ across the magnetic field H is given by the formula

$$F'_n = \frac{n^2 e^4 H^2}{2\pi m_e^2 c^2}\left(1-\frac{v^2}{c^2}\right) \qquad (3)$$
$$\times\left[\mathrm{ctan}^2\theta' J_n^2\left(n\frac{v}{c}\sin\theta'\right)+\frac{v^2}{c^2}\dot{J}_n^2\left(n\frac{v}{c}\sin\theta'\right)\right],$$

where θ' is the angle between the magnetic-field vector and the direction toward the observer, J and $\dot{J}$ are the corresponding Bessel function and its derivative, and a prime denotes quantities measured in the comoving frame. Here and below, the intensity is defined as the energy emitted by the system into unit solid angle per unit time.

The Bessel function can be expanded in a Maclaurin series in the vicinity of zero, retaining only the first term if the quantity $n\frac{v}{c}\cos\theta'$ is not large. Using the fact that $\left(1-\frac{v^2}{c^2}\right) \simeq 1$, we obtain after simple manipulation the asymptotic expansion

$$F'_n = \frac{3n^2\sigma_T c H^2}{16\pi^2}\frac{1}{n!}\left(\frac{nv}{2c}\sin\theta\right)^{2n}\left(\frac{2-\sin^2\theta'}{\sin^2\theta'}\right), \quad (4)$$

where the Thomson scattering cross section $\sigma_T = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2$ has been introduced. To obtain the total intensity of the radiation emitted by N particles distributed in accordance with (1), we must integrate (4) over the distribution:

$$I'_n = \int_0^\infty F'_n\, d\frac{m_e v^2}{2T}. \qquad (5)$$

After dividing this integral by N, we obtain the average intensity of the radiation emitted by one particle:

$$I'_n = \frac{3\sigma_T c H^2}{16\pi^2}\frac{n^{2n+2}}{n!}(\sin\theta')^{2n}\frac{(2-\sin^2\theta')}{\sin^2\theta'}\left(\frac{T}{2m_e c^2}\right)^n. \qquad (6)$$

[2] Here and below, a strongly anisotropic distribution will be defined as a velocity distribution for electrons whose motion along the magnetic field is ultrarelativistic and across the field is nearly nonrelativistic. This subject will be discussed in more detail below.

[3] We emphasize that, here and in (31), v is the total velocity in the comoving frame, not the velocity component perpendicular to the magnetic field in the laboratory frame.

We can see that $\lim_{n\to\infty} I'_n = \infty$ when $\theta' \neq \frac{\pi}{2}$. Obviously, this is related to the limited applicability of the asymptotic expansions of the Bessel function used in (4). In fact, expression (6) represents a zeroth approximation to the intensity. Taking into account the following terms of the expansion, we find that the condition for the applicability of (6) can be formulated

$$\frac{n(n+2)}{(n+1)}\frac{T}{m_e c^2} \ll 0.3. \tag{7}$$

Formula (6) yields substantially overestimated values for large transverse temperatures, and is not applicable at all when

$$\frac{n(n+2)}{(n+1)}\frac{T}{m_e c^2} \approx 1. \tag{8}$$

As noted above, the radiation intensities in the laboratory and comoving frames are related by an ordinary Lorentz transformation. The comoving frame moves with respect to the laboratory frame with the velocity V. Let us introduce the notation[4]

$$\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}, \qquad \mu = \frac{1}{1-\frac{V}{c}\cos\theta}, \tag{9}$$

$$\hat{\mu} = \frac{1}{1+\frac{V}{c}\cos\theta'}.$$

Using the results of [12], we can easily derive the relations

$$\omega = \omega'\frac{\sqrt{1-\frac{V^2}{c^2}}}{1-\frac{V}{c}\cos\theta} = \omega'\frac{\mu}{\gamma}, \tag{10}$$

$$\cos\theta = \frac{\cos\theta' + \frac{V}{c}}{1+\frac{V}{c}\cos\theta'} = \hat{\mu}\left(\cos\theta' + \frac{V}{c}\right). \tag{11}$$

We also emphasize the important relation

$$\mu\hat{\mu} = \gamma^2. \tag{12}$$

Uusing (9)–(12), we can easily obtain the relations

$$\sin\theta' = \frac{\gamma}{\hat{\mu}}\sin\theta = \frac{\mu}{\gamma}\sin\theta do' = \frac{\mu^2}{\gamma^2}do, \tag{13}$$

where do and do' are elements of solid angle in the laboratory and comoving frames.

Next, according to [12], the transformation of the intensity from the comoving to the laboratory frame takes the form

$$I = \frac{\mu^3}{\gamma^2}I'. \tag{14}$$

[4] As in the previous formulas, primed quantities refer to the comoving frame, and unprimed quantities refer to the laboratory frame or are frame independent.

Radiation at some specific nth harmonic in the comoving frame is monochromatic with frequency $n\omega_H$. In the laboratory frame, the frequency of the radiation emitted by one particle is an single-valued function of the observation angle θ (and, therefore, of θ'). Let us find this function in explicit form. According to (10), $\omega = \frac{\mu}{\gamma}\omega'$. Substituting $n\omega_H$ instead of θ' into this and using (12), we obtain

$$\omega = \frac{\mu}{\gamma}n\omega_H = \frac{\gamma}{\hat{\mu}}n\omega_H = \left(1+\frac{V}{c}\cos\theta'\right)\gamma n\omega_H. \tag{15}$$

If $V \simeq c$ and the angles θ' are not very close to π, we have [5]

$$\left(1+\frac{V}{c}\cos\theta'\right) \approx 1+\cos\theta'. \tag{16}$$

We obtain in the above approximation

$$\sin\theta' = 1-\left(\frac{\omega}{\gamma n\omega_H}-1\right)^2 = 2\frac{\omega}{\gamma n\omega_H} - \left(\frac{\omega}{\gamma n\omega_H}\right)^2. \tag{17}$$

We emphasize that (17) is valid only when $\tilde{\omega} \in [0;\ 2n\gamma\omega_H]$ (since $\theta' \in [0;\pi]$, so that $\sin\theta' \in [0;1]$).

Since the frequency of the radiation is a single-valued function of the angle, we can use this frequency, instead of the angle, as the variable specifying the direction. Substituting expression (6) for the intensity I'_n into (14), replacing $\sin\theta'$ in accordance with (17), and using the relation $\mu = \gamma\frac{\omega}{n\omega_H}$ [as follows from (10)], we finally obtain

$$\begin{aligned} I_n = &\frac{3\gamma\sigma_T c H^2}{16\pi^2}\frac{n^{2n+2}}{n!}\left(\frac{\omega}{n\omega_H}\right)^3 \\ &\times\left(2\frac{\omega}{\gamma n\omega_H} - \left(\frac{\omega}{\gamma n\omega_H}\right)^2\right)^{n-1} \\ &\times\left(1+\left(\frac{\omega}{\gamma n\omega_H}-1\right)^2\right)\left(\frac{T}{2m_e c^2}\right)^n. \end{aligned} \tag{18}$$

[5] The condition for smallness of the angle, when (16) is applicable, can be written

$$\frac{(1+\cos\theta') - (1+\frac{V}{c}\cos\theta')}{(1+\frac{V}{c}\cos\theta')} \ll 1.$$

If $\frac{V}{c} \simeq 1$, this condition is not valid only for angles θ' such that $\pi - \theta' \leq \frac{1}{\gamma}$. We can see that the solid angle corresponding to these plane angles is extremely small $\left(\sim \frac{\pi}{\gamma^2}\right)$. Since the intensities of all harmonics in the comoving frame have no features when $\theta' \simeq \pi$, the relative fraction of radiation corresponding to this solid angle is obviously also very small $\left(\sim \frac{1}{4\gamma^2}\right)$ and can be neglected.

This formula represents the intensity of the radiation emitted by one particle at the nth harmonic, where the directional dependence of the intensity is expressed in terms of the frequency emitted in that direction.

3. CALCULATION OF THE OBSERVEd CYCLOTRON SPECTRUM

Let us now find the radiation spectrum detected by an observer an infinite distance from the star. It is known (see, for example, [13]) that magnetobremsstrahlung radiation by ultrarelativistic particles is highly directional, and is nearly completely concentrated in a narrow cone (with an opening angle of about $\simeq \frac{1}{\gamma}$) along the direction of the particle's velocity (in the case under consideration, along the magnetic field). Therefore, the radiation detected by an observer at an infinite distance at any particular time is produced by a very small part of the emitting region, where the magnetic field is directed precisely toward the observer. The linear size of this region is on the order of $R\frac{1}{\gamma}$, where R is the star's radius.

Let us assume that the electrons are uniformly distributed over this part of the emitting region with surface density ρ, that their distribution function has the form (1), and that the magnetic field is constant and perpendicular to the stellar surface. Let this surface be separated into bands of width $Rd\varphi$, which are symmetric with respect to the line from the neutron-star center to the observer. Here, φ is the angle between rays drawn from the stellar center to the observer and a given point on the stellar surface (i.e., the latitude of this point). Let us consider one of these bands. Its area is $2\pi R^2 \sin\varphi d\varphi$, and the number of emitting electrons in this band is

$$\rho 2\pi R^2 \sin\varphi d\varphi. \tag{19}$$

Since we have assumed above that the magnetic field is everywhere perpendicular to the stellar surface, the angle between the normal to this surface and the direction toward an observer is equal to the angle between the magnetic field and the direction toward the observer. As follows from geometrical considerations,

$$\varphi = \theta \text{ and } d\varphi = d\theta. \tag{20}$$

The frequency of the radiation emitted by the band is given by (15), where θ can be replaced by φ:

$$\omega = \frac{\mu}{\gamma} n\omega_H = \frac{n\omega_H}{\gamma(1 - \frac{V}{c}\cos\varphi)}. \tag{21}$$

The angle φ is not absolutely constant within the band, and changes by an amount $d\varphi$. As a result, the radiation emitted by the band is not monochromatic, and covers some frequency interval $d\omega$. Let us calculate the width of this interval. Differentiating (21) with respect to φ, we obtain for $\frac{V}{c} \sim 1$

$$d\omega = n\omega_H \frac{\mu^2}{\gamma} \sin\varphi d\varphi. \tag{22}$$

Therefore, the band contributes to the total spectrum of the emitting region in the frequency interval $[\omega - d\omega; \omega]$, where ω and $d\omega$ are given by (21) and (22). Moreover, it can easily be deduced that the corresponding interval for the total spectrum is produced only by electrons in this band.

As follows from (22),

$$\sin\varphi d\varphi = d\omega = \frac{\mu^2}{\gamma}\frac{d\omega}{n\omega_H}. \tag{23}$$

Substituting (23) into (19), we obtain for the number of particles dq emitting in the band

$$dq = 2\pi\rho R^2 \frac{\gamma}{\mu^2}\frac{d\omega}{n\omega_H}. \tag{24}$$

Let do be the solid angle subtended by the observer at the neutron-star surface. Then, the total energy received by the observer from the band under consideration during a time dt will be

$$dE = I dq dt do,$$

where I is the intensity emitted by one particle. Substituting (24) for dq, using (10), and dividing both sides of the equality by $d\omega dodt$, we obtain an expression for the spectral energy density from the star P_n; i.e., for the amount of energy emitted by the star per unit frequency interval into unit solid angle per unit time:

$$P_n = I\frac{2\pi\rho R^2}{\gamma}\frac{n\omega_H}{\omega^2}. \tag{25}$$

Dividing this expression by $\hbar\omega$, we make the transformation from the spectral energy density P_n to the spectral particle-flux density Q_n. Let us introduce the dimensionless frequency

$$\tilde{\omega} = \frac{\omega}{\gamma\omega_H}. \tag{26}$$

Substituting formula (18) for the cyclotron intensity emitted by one particle into (25) and using the relation

$$\frac{3c\sigma_T H^2}{8\pi\hbar{\omega_H}^2} = \frac{e^2}{\hbar c} = \lambda \simeq \frac{1}{137}$$

(where λ is the fine-structure constant), we finally obtain

$$Q_n = \lambda\rho R^2 \frac{n^{2n}}{n!}\left(2\frac{\tilde{\omega}}{n} - \left(\frac{\tilde{\omega}}{n}\right)^2\right)^{n-1} \times \left(1 + \left(\frac{\tilde{\omega}}{n} - 1\right)^2\right)\left(\frac{T}{2m_e c^2}\right)^n. \tag{27}$$

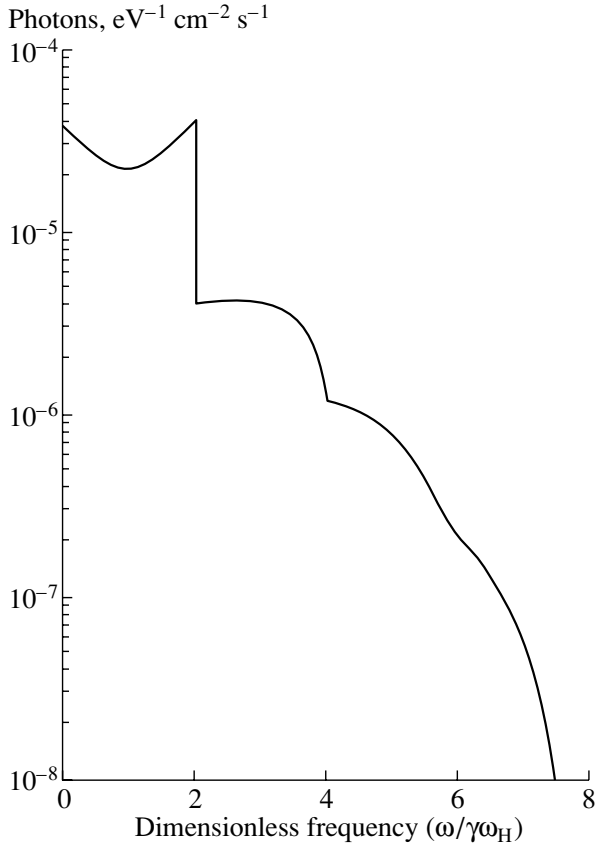

Fig. 1. Spectrum of cyclotron radiation emitted by an X-ray pulsar containing four harmonics, calculated using (27). The transverse temperature of the emitting electrons is $T = 20$ keV. The dimensionless frequency is plotted on a linear scale along the horizontal axis, and the number of particles on a logarithmic scale along the vertical axis.

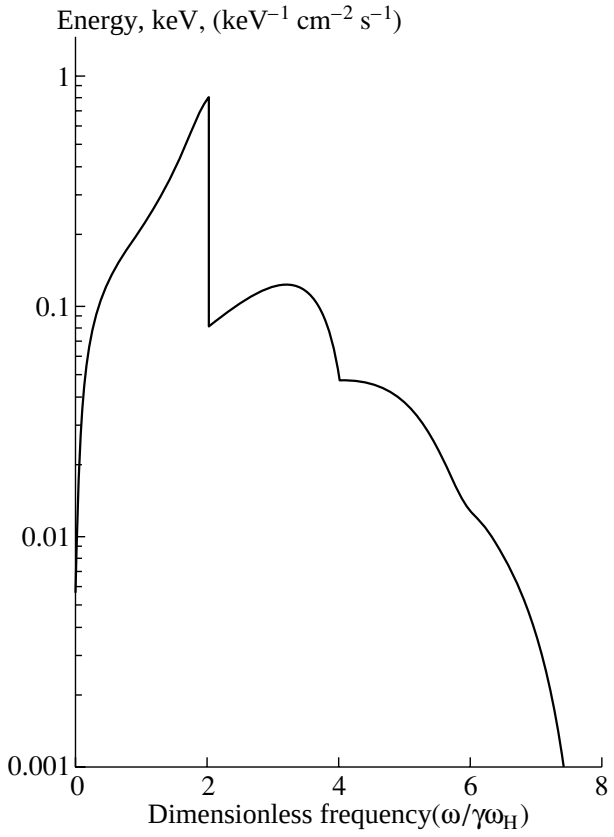

Fig. 2. Energy spectrum of cyclotron radiation by an X-ray pulsar containing four harmonics. The transverse temperature of the emitting electrons is $T = 20$ keV. The dimensionless frequency is plotted on a linear scale along the horizontal axis, and the emitted energy on a logarithmic scale along the vertical axis.

This formula determines the number of photons of the nth harmonic emitted at some time by the star per unit frequency interval into unit solid angle per unit time (i.e., the instantaneous observed spectrum of the pulsar). If the physical conditions are the same over the entire surface of the spot, this spectrum coincides with the time-averaged spectrum of the pulsar, with a coefficient corresponding to the fraction of time the hot spot is observed. Since the transformation (17) is valid for $\tilde{\omega} \in [0; 2n]$, formula (27) is valid when

$$0 \leq \tilde{\omega} \leq 2n. \tag{28}$$

Physically, this inequality reflects the fact that, according to (10), the photon frequency cannot increase by more than the factor 2γ in the transformation from the comoving to the laboratory frame, and all the photons have the frequency $\omega_n = n\omega_H$ in the comoving frame. Therefore, there are no photons with frequencies exceeding $2\gamma n\omega_H$ in the laboratory harmonic spectrum. Precisely this restriction is responsible for the sharp cut-off in the spectrum of the first harmonic.

We have neglected the gravitational redshift when deriving (27) and in subsequent relations. When this redshift is fairly small (the stellar radius R is far from the gravitational radius r_g), it can easily be taken into account by substituting the quantity

$$\Omega = \frac{\omega}{\gamma\omega_H}\sqrt{1 - \frac{r_g}{R}} \tag{29}$$

in place of the frequency given by (26). Nevertheless, we shall primarily use the previously derived formula (27).

Combining harmonics of several orders (beginning with the first) and using the constraint (28), we obtain the resulting cyclotron spectrum. Here and below, only the first four cyclotron harmonics will be used to construct the spectra, while lines of higher orders will be neglected. Let us assume that the distance to the pulsar is $r = 3.5$ kpc, the radius of the neutron star is $R = 10$ km (these correspond approximately to the parameters of X0115+63), and that the surface density of the radiating particles is

$\sim 10^{16}$ cm^{-2}. The value of $\hbar\gamma\omega_H$ will be taken to be 10 keV, which is also in reasonable agreement with the energy of the first harmonic observed in the spectrum of X0115+63 at ~ 20 keV.

The energy distribution of the photons that should be detected by an observer on the Earth for a transverse temperature of the emitting electrons $T =$ 20 keV is presented in Fig. 1. We can see that these harmonics are very broad and overlap each other. In addition, they are clearly not equally spaced; in particular, the maximum of the second harmonic nearly coincides with the maximum of the first harmonic. The first harmonic possesses a second maximum at $\omega \simeq 0$, which is somewhat weaker than the main maximum, since higher harmonics are also added to the main maximum.

It is usual to depict the frequency dependence of the total photon energy rather than of the number of photons. Figure 2 presents this dependence for the same transverse temperature $T = 20$ keV. The formula describing this spectrum can be obtained by multiplying (27) by $\hbar\omega$. As a result, the maxima of the harmonics will be shifted; in particular, the maximum of the second harmonic will not coincide with the maximum of the first harmonic and will appear distinct. In addition, the maximum at $\omega \simeq 0$ disappears. However, as before, the spectral features are not equidistant, and the harmonics are very broad and overlap each other.

These effects have a simple qualitative explanation. Let us consider the emission of photons in the comoving frame. First, we must answer the question of which photons will reach the distant observer. It is obvious that these will be photons moving away from the star in the laboratory frame, i.e., those for which $\theta \leq \frac{\pi}{2}$. According to (10), this corresponds to the angles $0 \leq \theta' \leq \pi - \frac{1}{\gamma}$ in the comoving frame.[6] Therefore, the observer will detect nearly all the photons, apart from those emitted precisely toward the star in the comoving frame.

As was noted above, the energy of all photons corresponding to a single harmonic is the same in the comoving frame. Nevertheless, the coefficient of transformation of the photon frequency from the comoving to the laboratory frame is not constant, and depends strongly on the angle θ'. Using (10), we can easily show that the frequency of photons emitted precisely toward the observer ($\theta' \simeq 0$) increases by the factor 2γ, the frequency of photons emitted perpendicular to the direction toward the observer ($\theta' \simeq \pi/2$ in the comoving frame) increases by the factor γ, and the frequency of photons emitted at the angle $\theta' = \pi - \frac{1}{\gamma}$ decreases by the factor γ. This obviously leads to considerable broadening of the lines. In addition, the equidistant character of the lines will be disrupted. This can be interpreted as follows. According to (6), most photons associated with the first harmonic are emitted at angles $\theta' \simeq 0$. Therefore, their frequency increases by the factor 2γ, and the maximum of the first harmonic occurs at $2\gamma\omega_H$. On the other hand, most photons associated with the second and higher harmonics are emitted perpendicular to the direction toward the observer ($\theta' \simeq \pi/2$), and particles with $\theta' \simeq 0$ are virtually absent (Fig. 3). The frequencies of these photons increase only by the factor γ, so that the maxima of higher harmonics are located at

$$\omega = \gamma\omega_n = \gamma n\omega_H, \quad n > 1. \tag{30}$$

In particular, the maxima of the first and second harmonics coincide. Since the harmonics overlap, their maxima are shifted, and relation (30) is not precisely satisfied. The resulting spectrum differs substantially from a spectrum with equidistant features.

We have considered above only radiation by electrons whose velocities along the field are exactly the same. At the same time, a real ensemble of electrons will certainly be characterized by some distribution of the momentum component along the field. Therefore, the total velocity distribution of the emitting electrons can be written

$$dn = f(p_\parallel) exp\left(-\frac{m_e v^2}{2T}\right) dp_\parallel d\left(\frac{m_e v^2}{2T}\right). \tag{31}$$

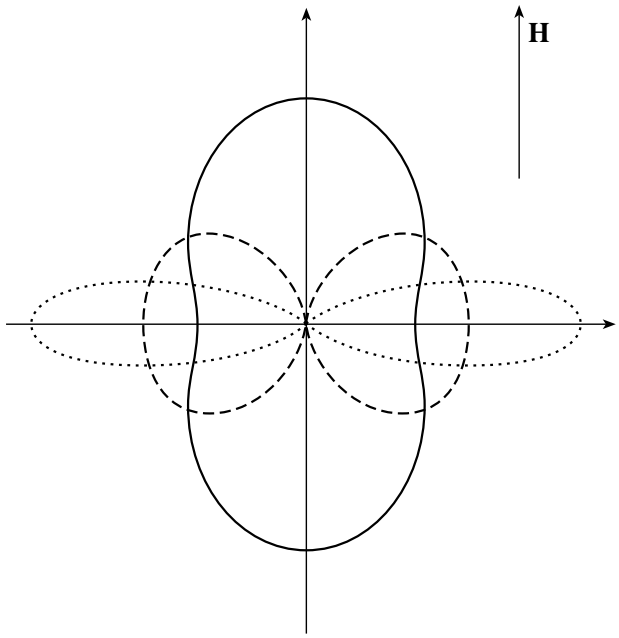

Fig. 3. Diagram of the directivity of the first (solid curve), second (dashed curve), and tenth (dotted curve) cyclotron harmonics. The scales are arbitrary. The direction of the magnetic field is marked by the arrow.

[6] Here and below, we assume that the laboratory frame moves ultrarelativistically with respect to the comoving frame.

This distribution will considerably affect the observed spectrum of an X-ray pulsar. In particular, the first harmonic may be transformed from a sharp peak (as in Fig. 2) to a broad line. Therefore, it is interesting to answer the following question: can the distribution $f(p_\parallel)$ in (31) be very broad for a real source, and can it result in the transformation of separate harmonics into a continuous spectrum?

As is noted above, ultrarelativistic particles can be produced by collisionless shocks formed in the accretion flow. They do not leave the shock, and oscillate at the shock front. Let us suppose that such oscillations are roughly harmonic. Then, the particle momentum will vary as $p = p_0 \cos \alpha t$, where α is the oscillation frequency. Since the oscillations are ultrarelativistic, the momenta of the particles are proportional to their Lorentz factors γ ($p \sim \gamma$), so that the γ values vary in the same way:

$$\gamma = \gamma_0 \cos \alpha t. \quad (32)$$

The distribution of the corresponding ensemble of particles along the field can be represented

$$dn = \frac{N}{\pi\gamma_0\sqrt{1-\left(1-\frac{\gamma}{\gamma_0}\right)^2}}d\gamma; \quad \gamma \in [1;\gamma_0]; \quad (33)$$

i.e., the distribution is characterized by a sharp peak with its maximum at $\gamma \simeq \gamma_0$. Of course, real distributions can differ considerably from (33); nevertheless, they can be sufficiently narrow that the individual cyclotron harmonics are not transformed into a continuous spectrum.

4. DISCUSSION

Among all the X-ray pulsars in which high cyclotron harmonics were found, of greatest interest is the source X0115+63. Its spectrum exhibits harmonics of at least three orders. Let us consider in more detail the high-energy spectra of these sources, since they contain the features of the most interest to us. The spectrum in this region possesses a quasi-power-law character with an exponential cut-off (if cyclotron features are not taken into consideration), as is very typical for X-ray pulsars. Spectra of this form are supposed to be produced by Comptonization of relatively cool radiation by hot electrons, considered in detail [15]. The resulting spectrum for energies above the cut-off at E_c is approximately a Wien spectrum with the characteristic temperature T_c equal to the temperature of the hot electrons T_e [15]. It is formed by photons that have undergone a large number of collisions with the electrons, and have therefore been heated to the temperature T_e. On the other hand, the photons forming the power-law part of the spectrum have experienced only a few collisions with hot electrons, and have not reached thermal equilibrium. As a result, the power-law spectrum at each specific frequency contains fewer photons than a Planck spectrum with temperature T_e.

To form an absorption line in such a spectrum, the electrons must have a temperature below T. Even if we assume that the electron-velocity distribution differs considerably from Maxwellian, the situation is unlikely to change fundamentally. In any case, the characteristic energy of motion of the electrons across the magnetic field should not exceed T_e. According to [6], the temperature T_e for X0115+63 is 17.4. If the electron temperature is unimportant (if it is much lower than $m_e c^2 \approx 500$), then the ratio of their optical depths for cyclotron radiation (absorption) at two neighboring harmonics is[7]

$$\frac{\tau_{n+1}}{\tau_n} = \frac{n(n+2)}{n+\frac{3}{2}}\left(1+\frac{1}{n}\right)^{2n+2}\left(\frac{kT}{2m_e c^2}\right). \quad (34)$$

Consequently, if the electron temperature satisfies the condition

$$T < T_e \sim 20\ \mathrm{keV}, \quad (35)$$

then their optical depth at the nth a harmonic should decrease rapidly with increasing harmonic number. In particular, the optical-depth ratio for the second and first harmonics should not exceed $\frac{2}{5}$. Therefore, the second harmonic must be formed closer (at any rate, not farther) to the neutron-star surface than the first harmonic (because the matter is much more transparent to the radiation corresponding to the second harmonic). In this case, the second harmonic is formed in a region with a strong magnetic field and must be shifted to the hard range from the first harmonic. In other words, the energy ratio of the maxima of the first and second harmonics must be no less than 2 ($\frac{w_2}{w_1} \geq 2$). The authors of [9] actually reported that the cyclotron lines in the spectrum of X0115+63 are not equidistant (they considered the lines as absorption ones). However, the second harmonic is shifted to the soft range (the $\frac{w_2}{w_1}$ ratio is 1.7). Therefore, the cyclotron feature observed in the spectrum of this object is unlikely to be an absorption one. The experimental spectrum of X0115+63 (see, e.g., [10]) shows that even if the observed lines are assumed to be emission ones, they are not equidistant.

In addition, the lines are very broad and overlapping. All these characteristic features of the cyclotron

[7] The Kirchhoff law for radiation is used to generalized formula (34) below to absorption. Strictly speaking, this law holds only under local thermodynamic equilibrium. However, on the one hand, our analysis is qualitative and, on the other hand, the Kirchhoff law holds good for the problem under consideration.

spectrum are usually explained by the complex geometry of the emitting region (a hot spot on the surface of the neutron star). The broadening of the line and its transformation to a broad band are thought to be due to variations in the intensity of the magnetic field within the emitting region,[8] while the higher harmonics are formed in parts of the hot spot where the magnetic field is weaker than in the regions of formation of the main harmonic, so that their frequencies are shifted with respect to the fundamental frequency, $\omega_n < n\omega_1$.

Unfortunately, this explanation runs into serious problems. It is not difficult to estimate the relative broadening of lines due to the above mechanism. First, we have

$$\frac{\Delta\omega}{\omega} = \frac{\Delta H}{H}. \qquad (36)$$

Let us assume that a neutron star is characterized by the standard radius ($R = 10$ km) and mass ($M = 1.4M_\odot$), a dipolar magnetic field

$$H_r = \frac{\mathbf{m}}{r^3}\cos\theta, \quad H_\theta = -\frac{\mathbf{m}}{2r^3}\sin\theta, \qquad (37)$$

and a temperature at the foot of the accretion column $T_e \sim 10$ keV. Then, the height of the hot region is approximately determined by the usual barometric formula

$$\Delta r = \frac{kT_e}{m_H g},$$

where k is the Boltzmann constant, m_H the mass of the hydrogen atom (the dominant component of the accretion gas), and g the free-fall acceleration at the neutron-star surface. Substituting the numerical values, we obtain $\Delta r \sim 0.7\ m$. The relative variation in the magnetic field of the form (37) within this distance from the neutron-star surface is

$$\frac{\Delta H}{H} = 3\frac{\Delta r}{R} \sim 2\times 10^{-4}.$$

Consequently, $\dfrac{\Delta\omega}{\omega} \sim 2\times 10^{-4}$. Further, let us estimate the variations in the magnetic field over the spot. If the angular scale of the spot is χ, then, as obviously follows from (37), the corresponding relative variation in the magnetic field will be

$$\frac{\Delta H}{H} = \frac{3}{8}\chi^2.$$

According to [16], the angular size of the spot is

$$\chi \simeq \sin\chi = 7.455\times 10^{-4} H^{-\frac{2}{7}} L^{\frac{1}{7}},$$

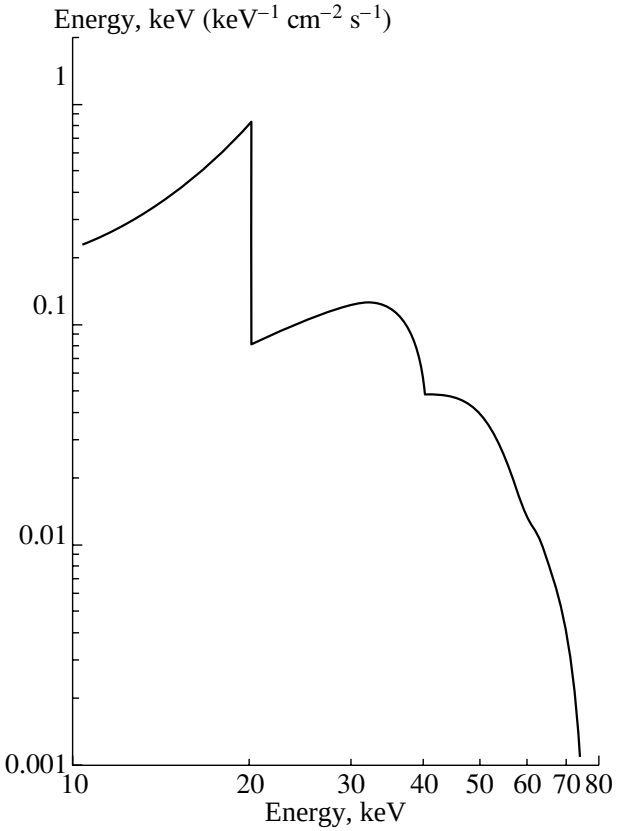

Fig. 4. Energy spectrum of cyclotron radiation by an X-ray pulsar containing four harmonics. The transverse temperature of the emitting electrons is $T = 20$ keV. The scales are logarithmic along both axes.

where L is the luminosity of the X-ray pulsar. The luminosity of the pulsar X0115+63 is $L \simeq 10^{37}$ erg/s [2], while the magnetic field (derived from the energy of the cyclotron line using the nonrelativistic formula) would be at least 7×10^{11} G. Hence, $\chi \simeq 0.06$, i.e., $\dfrac{\Delta\omega}{\omega} = \dfrac{\Delta H}{H} \approx 1.4\times 10^{-3}$.

Therefore, the broadening or shift of the cyclotron-line frequency due to the source geometry is $\dfrac{\Delta\omega}{\omega} \leq 2\times 10^{-3}$. Of course, this is completely inadequate to explain the observed width of the lines $\left(\dfrac{\Delta\omega}{\omega} \sim 0.2\right)$. Moreover, to reproduce the non-equal spacing of the lines, it is necessary to suppose a complex (and, therefore, quite artificial) temperature distribution over the spot.

Thus, we cannot explain the observed width of the cyclotron lines, their deviation from an equidistant distribution, or their overlapping using only geometrical considerations. The most probable and natural

[8] In the nonrelativistic case, cyclotron radiation (absorption) occurs at a single frequency, which depends only on the magnetic-field intensity, in accordance with (2).

explanation of all these properties is that the electrons giving rise to the cyclotron lines are ultrarelativistic, and so have a very anisotropic distribution of the form (31). Radiation produced by such particles possesses precisely these features. Therefore, the presence of several broad cyclotron lines separated by unequal distances in an X-ray pulsar's spectrum represents a strong argument in favor of our mechanism for the formation of the cyclotron features. The spectrum of cyclotron radiation calculated using (25) is plotted in Fig. 4 on a log–log scale for a transverse temperature of 20 keV. There is an obvious similarity between the observed spectrum of X0115+63 in [10] and the calculated spectrum in Fig. 4. Although the first harmonic in Fig. 4 represents a sharp peak, whereas it is quite broad in the observed spectrum, this could be explained as an effect of the velocity distribution (31) of the emitting electrons along the magnetic field. This same distribution should also result in additional broadening of the cyclotron lines.

The calculated cyclotron spectrum displays another very interesting feature: as we can see in Fig. 2, the dips between neighboring harmonics are located at nearly equal distances from each other. Therefore, if these decreases in intensity in the observed spectrum were interpreted as absorption lines, they would appear to be equidistant. This is probably why the cyclotron features in the spectra of X0115+63, Vela X-1, 4U1907+09, and A0535+26 have been interpreted as a set of equidistant absorption lines. However, as is demonstrated above, they are more likely to be emission rather than absorption lines.

We have already noted that ultrarelativistic, anisotropic electrons can be produced by a collisionless shock in the accretion flow. As is shown in [17, 18], such shocks can indeed form, and the corresponding electrons will have a strongly anisotropic velocity distribution. This is due to the fact that the electrons are accelerated primarily along the magnetic field. In addition, the transverse component of their momenta rapidly decreases due to radiative cooling, while the component along the field changes relatively little. The estimation of the electron temperature performed in [17, 18] yielded values $T \sim 10$ MeV, corresponding to $\gamma \sim 20$; i.e., the electrons are ultrarelativistic.

Note also that there is considerable indirect evidence that the particles forming the cyclotron lines in the spectra of X-ray pulsars are ultrarelativistic, primarily associated with discrepancies in estimates of the magnetic fields of these objects [11].

The pulsar Her X-1 was considered in detail in [16]. In general, its rotation speeds up, but it experiences deceleration in some intervals [19]. Consequently, its period is close to the equilibrium period; i.e., the angular velocity of rotation of the neutron star together with its magnetosphere is close to Keplerian at the Alfven radius. We can use this fact to estimate the magnetic field of this pulsar, which turns out to be relatively small, about $H = 3 \times 10^{11}$ G. In addition, radio pulsars in binary systems possess anomalously weak magnetic fields, $H \simeq 10^8$–10^{11} G [20]. Their formation as the result of the evolution of a pair containing an X-ray pulsar could be naturally explained if these pulsars had comparatively small magnetic fields, $H < 3 \times 10^{11}$ G. Mihara *et al.* [21] explained the observed 35-day cycle of Her X-1 as the result of periodic eclipses of the emitting region on the neutron-star surface by the accretion disk. For this mechanism to operate, the disk must be located at a sufficiently short distance from the surface; i.e., the field must be $H < 10^{11}$ G.

At the same time, a cyclotron line at energy 35–56 keV is observed in the spectrum of Her X-1. If the nonrelativistic formula (2) is applied, this corresponds to a magnetic field of $(3$–$5) \times 10^{12}$ G. This contradiction appears to be due to the fact that the electrons emitting the cyclotron line are actually ultrarelativistic. In this case, as is shown above, the energy of the first (fundamental) harmonic increases by the factor 2γ. Therefore, the nonrelativistic formula gives a substantially (by a factor of 2γ) overestimated value for the magnetic field. Thus, adopting the hypothesis that the velocities of the electrons emitting the cyclotron line are ultrarelativistic along the field makes it possible to avoid contradictions between estimates of the pulsar's magnetic field given by various methods; this represents weighty indirect evidence in favor of our mechanism for the formation of the cyclotron radiation.

Another argument supporting this mechanism is the observation of correlated variations in the energies of the cyclotron lines of some pulsars and in their luminosities [22]. This correlation can be explain in a natural way by our model. The luminosity of the hot spot is comparable to the Eddington luminosity; i.e., the radiation pressure appreciably affects the accretion rate. If the luminosity increases, the accretion flow will be substantially slowed. As a result, the intensity of the shock decreases, and the mean energy of the ultrarelativistic particles produced by the shock decreases. Consequently, the energy of the cyclotron lines also decreases [11]. However, we emphasize that it is recent observations of higher harmonics and their deviations from an equidistant spectral distribution that represent the first direct evidence that the electrons emitting the cyclotron lines are ultrarelativistic and have very anisotropic distributions of the form (31).

In this mechanism for the formation of the cyclotron lines, we must know the electron Lorentz factor γ to determine the pulsar's magnetic field. In

turn, this factor depends on the physical conditions in the accretion flow, in particular, in the collisionless shocks. It is possible that different types of X-ray pulsars have different shock structures, and, therefore, different characteristic γ values.

Accretion in low-mass binary systems occurs via the flow of material through the inner Lagrange point when the donor star fills its Roche lobe, whereas accretion in massive systems occurs via the capture of material from the powerful stellar wind of the O–B companion, which does not fill its Roche lobe. It may not be a chance coincidence that all four sources displaying several cyclotron harmonics are associated with massive systems, so that the second type of accretion is realized. In this case, the conditions at the front of the collisionless shock may be more favorable for the ultrarelativistic electrons to acquire greater momentum transverse to the magnetic field, resulting in the appearance of higher cyclotron harmonics in the pulsar spectrum.

5. CONCLUSION

The presence of several cyclotron harmonics in the spectra of X-ray pulsars provides important information about their physical properties.

First, in several cases, it makes the interpretation of the cyclotron feature as an absorption line doubtfull.

Second, if the cyclotron emission lines are not equidistant, this argues strongly that the electrons emitting these lines are ultrarelativistic and have a very anisotropic distribution. Emission by such electrons displays characteristic properties that are not typical of emission by nonrelativistic particles and which are actually observed in the observed spectra of X-ray pulsars. There is also other (indirect) evidence supporting this mechanism for the cyclotron-line formation.

Thus, the cyclotron radiation of X-ray pulsars is most likely produced by anisotropic ultrarelativistic electrons. As a result, the nonrelativistic formula (2) considerably overestimates the pulsar magnetic fields. The magnitude of this overestimation depends on the distribution function (31) of the emitting electrons in the corresponding shocks. We cannot rule out the possibility that this function may be different for different types of X-ray pulsars. Subsequent studies of the physical processes occurring in the accretion flows of these objects may help shed light on this problem.

Thus, investigations of the spectra of X-ray pulsars containing several cyclotron harmonics may enable us to refine estimates of their magnetic fields and thereby to resolve questions concerning the structure and evolution of these objects.

ACKNOWLEDGMENTS

I am grateful to G.S. Bisnovatyi-Kogan for help in preparation of this article and, in particular, for several fruitful ideas. I would also like to thank A. Santangelo for a helpful discussion. This work was supported by the Russian Foundation for Basic Research, project nos. 01-02-1690 and 02-02-06596, the "Astronomy" Program, and the Federal Program "Fundamental Research in the Field of Physical Sciences."

REFERENCES

1. J. Trümper, W. Pietsch, C. Reppin, *et al.*, Astrophys. J. **219**, L105 (1978).
2. N. E. White, J. H. Swank, and S. S. Holt, Astrophys. J. **270**, 711 (1983).
3. I. Kreykenbohm, P. Kretschmar, J. Wilms, *et al.*, Astron. Astrophys. **341**, 141 (1998).
4. M. Orlandini, D. dal Fiume, F. Frontera, *et al.*, Astron. Astrophys. **332**, 121 (1998).
5. G. Cusumano, T. Di Salvo, M. Orlandini, *et al.*, Astron. Astrophys. **338**, 79 (1998).
6. A. Santangelo *et al.*, in *Proceedings of the Taomina Integral Workshop, 1999.*
7. E. Kendziorra, P. Kretschmar, H. C. Pan, *et al.*, Astron. Astrophys. **291**, 31 (1994).
8. K. Iwasawa, K. Koyama, and J. Halpern, Publ. Astron. Soc. Jpn. **44**, 9 (1992).
9. W. A. Heindl, W. Coburn, D. E. Gruber, *et al.*, Astrophys. J. **521**, L49 (1999).
10. A. Santangelo, A. Segreto, S. Giarrusso, *et al.*, Astrophys. J. **523**, L85 (1999).
11. A. N. Baushev and G. S. Bisnovatyĭ-Kogan, Astron. Zh. **76**, 283 (1999) [Astron. Rep. **43**, 241 (1999)].
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics* (Nauka, Moscow, 1980).
13. V. L. Ginzburg, *Theoretical Physics and Astrophysics* (Nauka, Moscow, 1975; Pergamon, Oxford, 1979).
14. V. V. Borkus, A. S. Kaniovsky, R. A. Sunyaev, *et al.*, Pis'ma Astron. Zh. **24**, 415 (1998) [Astron. Lett. **24**, 350 (1998)].
15. R. A. Sunyaev and L. G. Titarchuk, Astron. Astrophys. **86**, 121 (1980).
16. G. S. Bisnovatyĭ-Kogan, Byull. Akad. Nauk Gruz. SSR, Abastumanskaya Astrofiz. Obs. **58**, 175 (1985).
17. Ya. B. Zel'dovich and N. I. Shakura, Astron. Zh. **46**, 225 (1969) [Sov. Astron. **13**, 175 (1969)].
18. G. S. Bisnovatyĭ-Kogan and A. M. Fridman, Astron. Zh. **46**, 721 (1969) [Sov. Astron. **13**, 566 (1969)].
19. V. M. Lipunov, Astron. Zh. **64**, 321 (1987) [Sov. Astron. **31**, 167 (1987)].
20. V. Boriakoff, R. Buecheri, and F. Fauci, Nature **304**, 417 (1983).
21. E. K. Sheffer and I. F. Kopaeva, Astron. Zh. **69**, 82 (1992) [Sov. Astron. **36**, 41 (1992)].
22. T. Mihara, K. Makishima, and F. Nagase, in *Proceedings of the International Workshop "All-sky X-ray Observations in the Next Decade," 1997,* p. 135.

Translated by Yu. Dumin

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 840–850. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 932–943.
Original Russian Text Copyright © 2002 by Smirnov.

On the Atomic Constants of TaI

Yu. M. Smirnov

Moscow Power Engineering Institute, ul. Krasnokazarmennaya 17, Moscow, 111250 Russia

Received February 1, 2002

Abstract—Cross sections for the excitation of transitions in the tantalum atom not related to the ground state have been measured using extended crossed beams and optical spectroscopy. The total excitation cross sections for 23 energy levels of TaI have been determined. The accuracy of known transition probabilities that can be used to find branching ratios is analyzed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Atomic constants form the basis for solving a large number of problems in modern astrophysics. One subject of primary importance is, of course, the probabilities of radiative transitions. However, as the scope of investigations expands and the number of objects studied increases, reference data on the cross sections of electron–atom collisions also often become important.

Unfortunately, neither the completeness nor the accuracy of available data on atomic constants are satisfactory. This is especially true with regard to cross sections Q for the excitation of atoms via electron collisions [1]. While both experimental and theoretical results have been obtained by many groups for inert gases and some low-temperature elements, data of only one group are often available for many metals with moderate evaporation temperatures (1000–2000 K). For elements with evaporation temperatures of 2000–3000 K, there have been only a few experimental studies carried out using the extended crossed-beam method and several theoretical studies by R.K. Peterkop, in which the cross sections were calculated using the Born method. Finally, there are no published results for higher-temperature elements ($T > 3000$ K), in particular, for the tantalum atom, which we consider here.

The radiative transition probabilities A_{ki} have been more thoroughly studied, since the necessary experiments can be conducted using gas-discharge radiation sources, while available methods for studying electron–atom collisions are based on the use of crossed beams. The intensity of optical radiation in this latter case is usually several orders of magnitude lower than in gas-discharge sources. The situation becomes especially difficult when determining atomic constants for high-temperature elements.

Extensive data on the transition probabilities of the tantalum atom were obtained in [2] using photographic photometry. A modern method, based on a combination of photoelectric measurements of the branching ratios (BR) and the radiative lifetimes τ, was used in the later work [3]. Reasonably accurate values of τ were taken from earlier work by the same group [4]. In this way, the probabilities of transitions from 35 fairly low levels of TaI in the energy range 18 500–35 500 cm^{-1} were determined in [3].

Our present work deals with experimental determinations of 73 excitation cross sections for tantalum lines that do not originate from the ground level.

2. THE EXPERIMENTAL

We used an extended crossed-beam method in which the optical emission by excited atoms from the region of beam crossing is detected.

To form the atomic beam, tantalum was evaporated from an autocrucible produced by melting tantalum rods by heating them with an electron beam. The general design of the apparatus and a detailed discussion of the extended crossed-beam method can be found in [5, 6]. It is currently the only setup enabling studies of atomic excitation for any evaporation temperature.

The action of the autocrucible is made possible by the very high power loss via the surface emission of the melted metal. The losses via heat transfer from the metal to the graphite substrate were very small due to the small contact area between the tantalum rods and substrate. The absence of direct contact between the melt and substrate enabled us to considerably increase the lifetime of the substrate and ensured the absence of carbides in the melt and atomic beam. The density of atoms in the region where the electron and atom beams intersected was 2.1×10^9 cm^{-3} at the melt surface temperature of 3400 K.

When tantalum is evaporated, not only the ground level $5d^3 6s^2 a^4 F_{3/2}$, but also other low-lying levels of the same term with $J = 5/2$, $7/2$, and $9/2$ are

1063-7729/02/4610-0840$22.00 © 2002 MAIK "Nauka/Interperiodica"

populated. If the distribution of atoms in these levels is assumed to be close to that for thermodynamic equilibrium, we obtain the following estimates for the populations as fractions of the total density of atoms in the beam (where the numbers in parentheses represent the level energies in cm^{-1}): 40.6% for $a^4F_{3/2}(0)$, 26.1% for $a^4F_{5/2}(2010)$, 15.3% for $a^4F_{7/2}(3963)$, and 9.5% for $a^4F_{9/2}(5621)$. The total population of the ground-term levels is 91.5% of the total number of atoms in the beam; another 6.0% is associated with the three quartet levels $5d^36s^2a^4P_{1/2,3/2,5/2}$, located above 6000 cm^{-1}. Finally, the last 2.5% corresponds to the total population of the five levels of the sextet term $5d^46sa^6D$ ($J = 1/2$–$9/2$) with energies above 9700 cm^{-1}. Since almost all real objects have similar temperature distributions, using these results as reference data when investigating specific problems should not lead to appreciable errors.

The spectral resolution of our equipment is one of the highest achieved in beam experiments, about 0.1 nm. The characteristic width of the electron energy distribution in the range 20–250 eV is below 1.0 eV (for 90% of the beam electrons). The current density of the electron beam is below 1.0 mA/cm^2.

When we work with high-temperature elements, the presence of bright background radiation from the surface of the melted metal is an important factor, which considerably restricts the spectral range that can be studied. The power of this emission in our experiment was several kW. None of the diaphragms can completely eliminate scattered radiation, especially when the electron and atom beams are considerably extended ($\sim$ 190 mm) along the line of sight of the optical system. Due to this factor, measurements with tantalum can be conducted only at $\lambda < 370$ nm. However, the following two circumstances must also be taken into consideration.

(1) Placing astronomical instruments beyond the Earth's atmosphere and constructing orbiting telescopes removes certain previous restrictions on the spectral ranges used in observational astronomy.

(2) Using reliable information on the branching ratios makes it possible to find the excitation cross sections for spectral lines in the visible and infrared with satisfactory accuracy if the excitation cross section for at least one transition in the ultraviolet was measured for the corresponding upper level.

The errors in the relative values of the cross sections are 10–25%. The absolute values of the cross sections are determined with an accuracy of $\pm 25 \ldots \pm 40\%$. More complete information on the techniques and procedures associated with experiments with extended crossed beams and a discussion of various sources of errors are given in [5, 6].

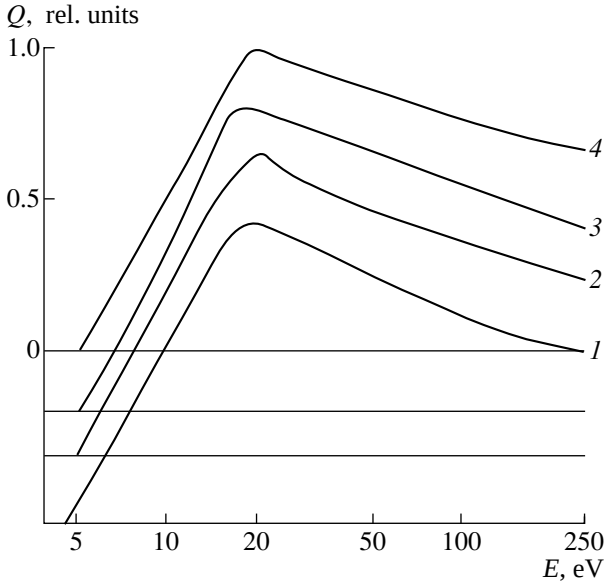

Fig. 1. Optical excitation functions of tantalum. The curve numbers correspond to column no. 9 in Table 1.

3. EXCITATION CROSS SECTIONS OF TaI: RESULTS AND DISCUSSION

Optical emission spectrum generated by inelastic collisions between the beam of monoenergetic electrons and the tantalum atoms was detected when the energy of the exciting electrons was 50 eV. We measured the dependence of the excitation cross sections on the energy of the colliding electrons (i.e., the optical excitation functions, OEF) for the most intense lines at energies of 0–250 eV.

Over 100 excitation cross sections for TaI spectral lines were obtained in the interval 230–370 nm, none of which can be reliably identified with TaII. Results for resonant transitions, as well as competing nonresonant lines, are discussed in [7]. We describe here our data for transitions from upper levels that are not connected to the ground level via intense allowed transitions in this spectral range.

The results of these measurements are presented in Table 1, which contains the wavelength λ, type of transition and inner quantum number J, energies of the low E_{low} and upper E_{up} levels, cross sections for an electron energy of 50 eV Q_{50} and at the maximum of the OEF Q_{max}, and the position of this maximum $E(Q_{max})$. The numbers in the OEF column correspond to the notation in Fig. 1.

The exact wavelengths, types of transitions, J, and level energies were taken primarily from [8, 9]. Only the energies, J, and parities are known for most of the upper levels, while the configurations and terms are not established. We introduced some additions to this classification using the data of [10, 11] on the

Table 1. Excitation cross sections for spectral lines of the tantalum atom

λ, nm	Transition	J	E_{low}, cm^{-1}	E_{up}, cm^{-1}	Q_{50}, $10^{-18}cm^2$	Q_{max}, $10^{-18}cm^2$	$E(Q_{max})$, eV	OEF	$I\lambda/gA$
1	2	3	4	5	6	7	8	9	10
238.000	$5d^36s^2a^4F-$	9/2–7/2	5621	47625	19.7				–
239.630	$5d^36s^2a^4F-$	9/2–11/2	5621	47339	15.5				12.0
243.991	$5d^36s^2a^4F-$	5/2–5/2	2010	42982	20.8				0.465
244.394	$5d^36s^2a^4P-$	3/2–3/2	6068	46974	43.7				–
244.413	$5d^46s\ a^6D-$	3/2 - 3/2	9975	50877					–
247.462	$5d^36s^2a^4F-$	5/2 - 3/2	2010	42408	68.1				6.09
251.978	$5d^36s^2a^4P-$	1/2 - 3/2	6049	45723	22.8				4.85
252.635	$5d^36s^2a^4F-5d^46p\ y^6F^\circ$	5/2–5/2	2010	41580	46.2				2.55
253.623	$5d^46sa^6D-$	5/2–5/2	11243	50660	30.4				–
255.505	$5d^36s^2a^4F-$	7/2–9/2	3963	43090	23.1				4.12
255.771	$5d^36s^2a^2P-$	1/2–3/2	11792	50877	42.2				–
256.210	$5d^36s^2a^4F-$	7/2–5/2	3963	42982	60.8				4.50
257.778	$5d^36s^2a^4F-$	9/2–9/2	5621	44402	52.4				3.09
258.561	$5d^36s^2a^2G-$	7/2–5/2	9705	48369	19.5				1.19
259.308	$5d^36s^2a^4F-$	9/2–9/2	5621	44173	46.1				4.42
259.526	$5d^36s^2a^4F-$	9/2–11/2	5621	44141	64.0				1.18
260.863	$5d^36s^2a^4F-$	5/2–7/2	2010	40333	212.	246.	17	3	34.0
261.566	$5d^36s^2a^4F-$	5/2–3/2	2010	40230	43.4				6.23
263.667	$5d^36s^2a^4F-$	7/2–5/2	3963	41879	124.				3.62
263.690	$5d^36s^2\ a^4F-$	9/2–7/2	5621	43533					1.76
265.327	$5d^36s^2a^4F-$	5/2–5/2	2010	39688	286.	358.	20	2	–
266.134	$5d^36s^2a^4F-$	9/2–11/2	5621	43185	159.				2.13
269.681	$5d^36s^2a^4P-5d^46p\ y^4P^\circ$	5/2–5/2	9253	46323	26.5				3.51
270.669	$5d^36s^2a^2G-$	9/2–7/2	10690	47625	42.3				0.521
271.718	$5d^36s^2a^4F-$	7/2–7/2	3963	40755	25.9				13.6
272.076	$5d^36s^2a^4F-$	5/2–5/2	2010	38753	32.1				0.124
272.778	$5d^36s^2a^2G-$	9/2–11/2	10690	47339	28.7				2.48
274.878	$5d^36s^2a^4F-$	7/2–7/2	3963	40333	142.	165.	17	3	12.5
275.831	$5d^36s^2a^4F-$	5/2–7/2	2010	38253	80.0				13.1
280.207	$5d^36s^2a^4F-$	7/2–7/2	3963	39641	86.7				36.5
281.512	$5d^36s^2a^4P-5d^46p\ y^6F^\circ$	3/2–5/2	6068	41580	31.0				28.2
282.118	$5d^36s^2a^4P-5d^46p\ y^4P^\circ$	5/2–3/2	9253	44689	10.7				–
282.718	$5d^36s^2a^4F-$	9/2–7/2	5621	40981	34.1				65.7
283.364	$5d^46sa^6D-5d^46p\ x^6D^\circ$	3/2–5/2	9975	45255	18.0				16.3
284.425	$5d^36s^2a^4P-$	1/2–3/2	6049	41197	26.8				18.4
284.535	$5d^36s^2a^4F-$	9/2–7/2	5621	40755	37.5				17.0
285.098	$5d^36s^2a^4F-5d^36s6p\ y^4G^\circ$	9/2–9/2	5621	40686	271.	319.	19	4	7.13
285.728	$5d^36s^2a^2G-$	7/2–7/2	9705	44693	20.6				–
286.202	$5d^46sa^6D-5d^46p\ y^4P^\circ$	1/2–3/2	9758	44689	29.8				1.22

Table 1. (Contd.)

1	2	3	4	5	6	7	8	9	10
286.865	$5d^36s^2a^4F-5d^36s6p\ z^4G^\circ$	7/2–9/2	3963	38813	40.2	48.7	19	1	28.7
288.002	$5d^36s^2a^4F-$	9/2–7/2	5621	40333	60.3	70.0	17	3	14.9
290.036	$5d^36s^2a^2G-$	7/2–9/2	9705	44173	40.0				14.9
291.549	$5d^36s^2a^4F-$	7/2–7/2	3963	38253	37.2				33.6
291.712	$5d^36s^2a^4P-5d^46p\ y^6F^\circ$	3/2–3/2	6068	40339	14.1				10.8
292.646	$5d^36s^2a^4P-$	3/2–3/2	6068	40230	22.5				37.5
294.006	$5d^36s^2a^2G-$	9/2–9/2	10690	44693	79.4				2.26
295.192	$5d^36s^2a^4F-5d^36s6p\ z^6S^\circ$	5/2–5/2	2010	35876	26.2				98.5
296.332	$5d^36s^2a^4F-$	5/2–7/2	2010	35746	65.3				26.2
296.554	$5d^36s^2a^4F-$	5/2–3/2	2010	35720	82.2				26.2
297.556	$5d^36s^2a^4F-$	7/2–7/2	3963	37561	24.5				37.2
297.754	$5d^46sa^6D-$	3/2–5/2	9975	43550	34.4				–
298.119	$5d^36s^2a^2G-$	7/2–5/2	9705	43239	29.8				4.07
298.858	$5d^36s^2a^2G-$	9/2–11/2	10690	44141	20.8				2.72
301.188	$5d^36s^2a^4F-5d^36s6p\ z^4G^\circ$	9/2–9/2	5621	38813	133.	161.	19	1	23.2
301.909	$5d^36s^2a^4F-5d^36s6p\ z^4H^\circ$	7/2–9/2	3963	37076	14.9				–
302.751	$5d^46sa^6D-5d^46p\ x^6D^\circ$	7/2–5/2	12234	45255	24.8				3.44
303.891	$5d^36s^2a^2P-5d^46p\ y^4P^\circ$	1/2–3/2	11792	44689	9.0				–
304.596	$5d^46sa^6D-$	9/2–7/2	13351	46172	22.0				8.78
322.132	$5d^46sa^6D-$	3/2–5/2	9975	41010	15.2				25.4
328.984	$5d^36s^2a^4P-$	5/2–7/2	9253	39641	17.4				–
331.116	$5d^36s^2a^4F-$	9/2–11/2	5621	35813	34.9				33.1
331.884	$5d^36s^2a^4F-5d^36s6p\ z^4H^\circ$	5/2–7/2	2010	32132	31.5				61.1
335.151	$5d^46sa^6D-5d^46p\ y^6F^\circ$	1/2–3/2	9758	39587	11.2				5.49
336.164	$5d^46sa^6D-$	9/2–9/2	13351	43090	26.3				10.8
337.605	$5d^46sa^6D-5d^46p\ y^6F^\circ$	3/2–3/2	9975	39587	17.4				10.7
343.450	$5d^36s^2a^2G-5d^36s6p\ z^4G^\circ$	7/2–9/2	9705	38813	23.9	29.0	19	1	23.5
346.012	$5d^36s^2a^2F-$	5/2–5/2	17224	46117	17.5				–
347.390	$5d^46sa^6D-$	3/2–5/2	9975	38753	9.6				11.2
349.785	$5d^36s^2a^4F-5d^36s6p\ y^4G^\circ$	5/2–7/2	2010	30590	22.0				94.1
350.498	$5d^36s^2a^4F-5d^36s6p\ y^6D^\circ$	7/2–5/2	3963	32486	10.4				164.
353.158	$5d^36s^2a^4P-$	5/2–7/2	9253	37561	12.4				16.7
359.564	$5d^36s^2a^2D-$	3/2–5/2	10950	38753	17.7				22.9
362.662	$5d^36s^2a^4F-5d^36s6p\ y^4G^\circ$	7/2–9/2	3963	31530	12.3				51.4
364.206	$5d^36s^2a^4F-5d^36s6p\ y^4G^\circ$	9/2–11/2	5621	33070	22.5				70.1
366.689	$5d^36s^2a^2P-$	3/2–3/2	15903	43167	13.8				–
368.973	$5d^36s^2a^2G-$	7/2–7/2	9705	36799	10.2				34.1
369.305	$5d^36s^2a^2D-5d^36s6p\ z^4G^\circ$	5/2–7/2	12865	39936	16.3				19.0

Table 2. Total excitation cross sections of the energy levels of the tantalum atom

λ, nm	E_{up}, cm^{-1}	Q_{50}, $10^{-18}\,cm^2$	BR[3]	ΣQ_{50}, $10^{-18}\,cm^2$
349.785	30590	22.0	0.676	32.5
362.662	31530	12.3	0.870	14.1
350.498	32486	10.4	0.164	63.5
364.206	33070	22.5	0.957	23.5
297.556	37561	24.5		36.9
353.158		12.4		
257.831	38253	80.0		117.2
291.549		37.2		
272.076	38753	32.1		59.4
347.790		9.6		
359.564		17.7		
286.865	38813	40.2		197.1
301.188		133.		
343.450		23.9		
335.151	39587	11.2		28.6
337.605		17.4		
280.207	39641	86.7		104.1
328.984		17.4		
261.566	40230	43.4		65.9
292.646		22.5		
260.863	40333	212.		414.3
274.878		142.		
2898.002		60.3		
271.718	40755	25.9		63.4
284.535		37.5		
252.635	41580	46.2		77.2
281.512		31.0		
243.991	42982	20.8		81.6
256.210		60.8		
255.505	43090	23.1		49.4
336.164		26.3		
259.526	44141	64.0		84.8
298.858		20.8		
259.308	44173	46.1		86.1
290.036		40.0		
282.118	44689	10.7		49.5
286.202		29.8		
303.891		9.0		
285.728	44693	20.6		100.0
294.006		79.4		
283.364	45255	18.0		42.8
302.751		24.8		
239.630	47339	15.5		44.2
272.778		28.7		
238.000	47625	19.7		62.0
270.669		42.3		

hyperfine structure of TaI lines. Note that, although the title of Table 1 from [10] states that the transitions were taken from [8, 9], about one-third of the presented lines and corresponding transitions are absent from these papers. The actual source of information on these lines was probably [12] or later editions of the same tables.

Each of [8, 9, 11] refutes the existence of some levels found in previous studies. In addition, the existence of some levels has been confirmed but their identifications changed. In some cases, the values of J have also changed. The identified levels of TaI are presented in [8, 9] in pure LS-coupling notation. There is undoubtedly a substantial mixing of configurations in tantalum, whose electronic shell is very complex. The presence of significant perturbations was noted in [9], but there has been no more adequate treatment of most of the TaI energy levels until now. Only [13] presents a parametric calculation of 40 even levels of TaI located below 27 200 cm^{-1} and associated with the $5d^36s^2 + 5d^46s + 5d^5$ configurations. This analysis was continued in [14], where levels known previously were augmented by ten new levels found experimentally. In addition, the positions of a number of high-lying even levels that were not detected experimentally were calculated theoretically. A similar study of odd levels of TaI has not been carried out.

The excitation cross sections of spectral lines Q_{ki} can be derived directly from experiments in which information on electron–atom collisions is obtained by detecting optical signals emitted by the excited atoms. The excitation cross sections of energy levels q_k are used as the basic quantities in both theoretical studies of collisional processes and in most practical experiments. These quantities are linked by the relation

$$Q_{ki} = \left(q_k + \sum_l Q_{lk}\right) \frac{A_{ki}}{\sum\limits_m A_{km}}, \qquad (1)$$

where A_{ki} and A_{km} are the transition probabilities $k \to i$ and $k \to m$. The sum in parentheses takes into account the contribution of cascade transitions from all upper levels l to the population of level k. The factor following the parentheses represents the branching ratio BR_{ki}, which specifies the contribution of the transition $k \to i$ to the deactivation of level k via spontaneous radiative transitions.

Since $A_{ki}/A_{km} = Q_{ki}/Q_{km}$, (1) can be reduced to the form

$$q_k = \sum_m Q_{km} - \sum_l Q_{lk}. \qquad (2)$$

The first sum represents the total excitation cross section for level k, which takes into account both the

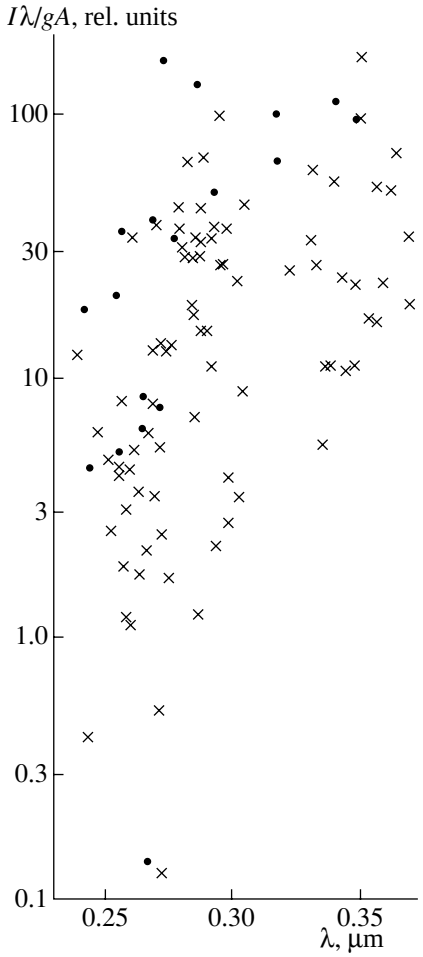

Fig. 2. $I\lambda/gA$ as a function of λ for the data of [2, 12]. Resonant transitions are presented by dots and nonresonant transitions by crosses.

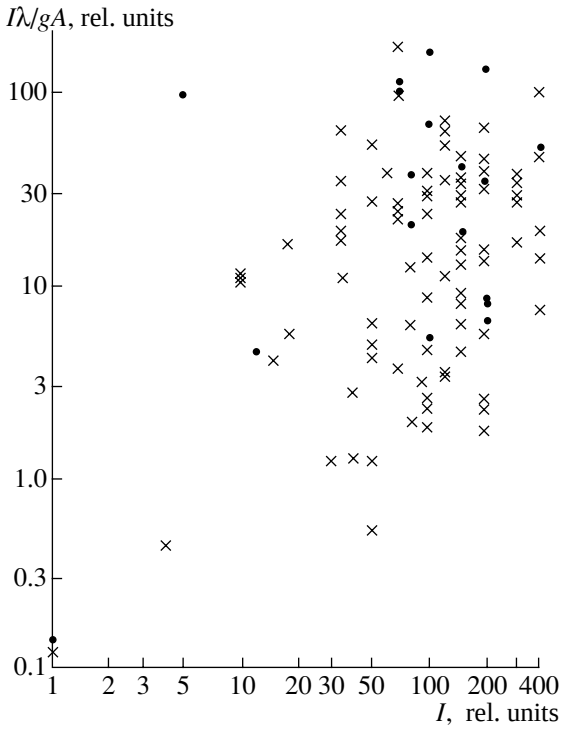

Fig. 3. Same as Fig. 2 for $I\lambda/gA$ as a function of I.

direct population of this level by electron collisions from the initial state and its population by cascade processes. It follows from (2) that the excitation cross section of the energy level q_k can be determined by measuring the excitation cross sections for spectral lines, and no additional information is necessary. Unfortunately, the accuracy of the corresponding values of q_k depends on the completeness of experimental information on the terms appearing in both sums (as well as on the accuracy of measurements of Q_{ki}).

In general, obtaining information on the cross sections that is sufficiently complete for the application of (2) requires measurements conducted over a wide range of wavelengths and intensities. The largest difficulties are usually encountered in the infrared, where detector sensitivities are lower and the influence of the background is larger than in the visible and ultraviolet. The most favorable situation is in the vacuum ultraviolet: although the instruments must be installed in a vacuum, extremely high sensitivity can be attained by counting individual photons.

However, there is the alternative of determining q_k by measuring the excitation cross section for only one transition from level k. As follows from (1), we can use the absolute, or even relative, transition probabilities and the branching ratios obtained in independent measurements. A_{ki} or BR_{ki} are usually measured using gas discharges as sources of radiation. As was noted above, the intensities of gas-discharge sources are typically several orders of magnitude higher than those of crossed-beam sources. Therefore, gas-discharge experiments can potentially provide substantially more extensive and accurate information on branching ratios, covering the maximum possible number of spectral lines.

A decisive factor limiting the measurable spectral range in our experiments is the extremely intense background radiation from the surface of the melted metal. Another factor hampering measurements at 306–320 nm is the radiation of intense hydroxyl bands, due to the dissociative excitation of OH by e–H_2O collisions [15]. Water vapor is inevitably present in metallic vacuum systems, even after prolonged heating. The situation was especially difficult in our case due to the small density of tantalum atoms in the beam. Finally, measurements in the vacuum ultraviolet are not interesting for tantalum, since the threshold of our apparatus, 190 nm, corresponds to

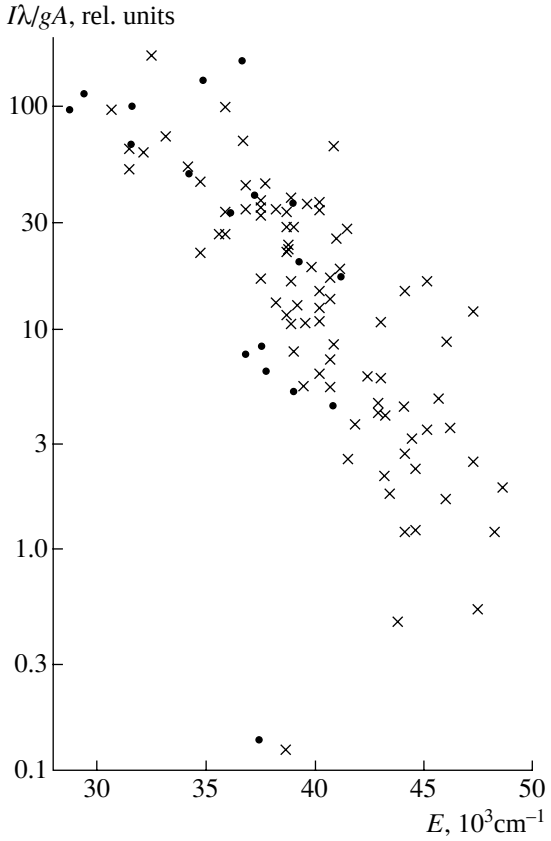

Fig. 4. Same as Fig. 2 for the $I\lambda/gA$ as a function of E_{up}.

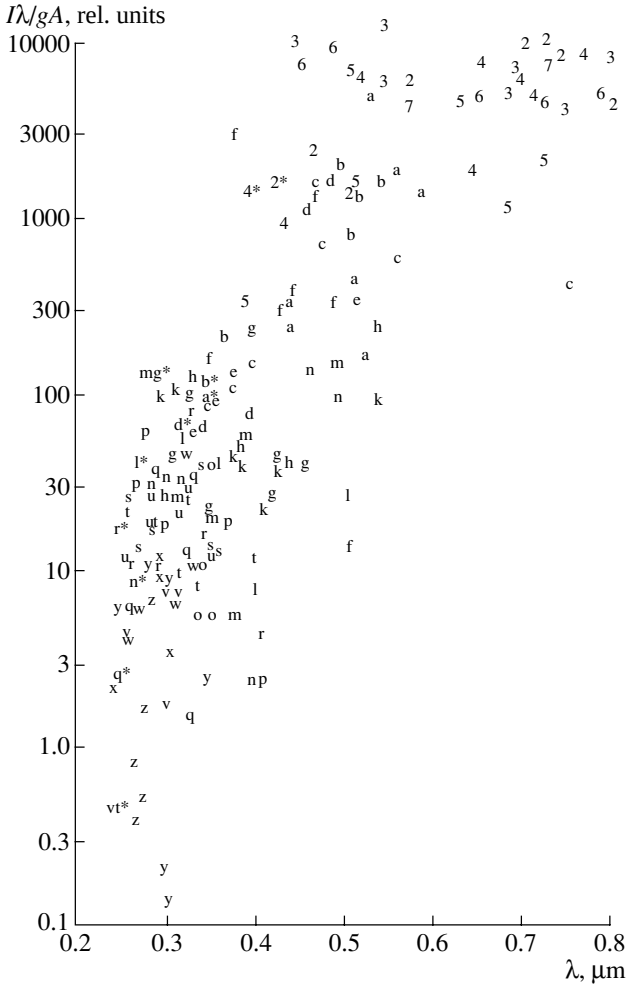

Fig. 5. $I\lambda/gA$ as a function of λ over a wide wavelength range. Resonant lines are marked by asterisks. The various plots show data for the following lines: (a) 28 689 cm^{-1}, (b) 29 343, (c) 30 591, (d) 31 428, (e) 32 132, (f) 32 486, (g) 34 792, (h) 35 746, (k) 35 876, (l) 37 145, (m) 37 523, (n) 37 630, (o) 39 587, (p) 39 936, (q) 40 230, (r) 40 339, (s) 40 755, (t) 41 010, (u) 41 197, (v) 42 982, (w) 43 090, (x) 45 255, (y) 46 172, (z) 47 625, (7) 23 355, (2) 23 363, (3) 24 243, (4) 25 181, (5) 25 512, and (6) 25 926. (All photon energies are given in cm^{-1}.)

a resonant photon energy of 52 500 cm^{-1}, and the highest known tantalum energy level does not reach this value [9].

Despite these restrictions, branching can be taken into account for a number of the TaI levels studied here using the measured cross sections for competing transitions, which are usually most intense for upper TaI levels precisely in the spectral range under investigation. Moreover, we can use the information on branching ratios obtained in [3] for some levels. Unfortunately, the contribution of cascade level population cannot be taken into account at this stage, since all cascade transitions to the considered levels of TaI are beyond the studied spectral range, in the visible and infrared. Furthermore, even spectroscopic information on such transitions is very sparse: only seven transitions whose lower levels are above 30 000 cm^{-1} were presented in [9], and there are no such transitions in [8].

The total excitation cross sections of TaI energy levels for an exciting-electron energy of 50 eV are presented in Table 2. These cross sections are quite large, and $\Sigma Q_{50} \geq 1.0 \times 10^{-16}$ cm^2 for the five levels. Unfortunately, the overlap between the levels studied in [3] and in the present work is small. As a result, we were able to use only four branching ratios from [3].

Considerably more extensive data on TaI transition probabilities are presented in [2]. These could be used, in particular, to obtain the branching ratios for upper levels for which we were able to detect only one transition. Unfortunately, the transition probabilities for atoms and singly-charged ions of many elements obtained in [2] differ substantially from the data of other authors. Nevertheless, the data in [2] are useful when these are the only data available. This situation occurs most frequently for fairly high-lying levels of

atoms with complex electron shells. This is precisely the case here, since information about the transition probabilities or branching ratios for TaI levels with energies above 35 500 cm^{-1} is completely absent.

4. TRANSITION PROBABILITIES FOR THE TANTALUM ATOM

It is potentially useful to analyze the discrepancies between the data of [2] and the results of other authors. The only possibility to do this for high-lying levels of tantalum is to compare with the data of [12], which contains the relative intensities of almost all known spectral lines of TaI. The intensities presented in [8, 9] were taken from [12].

The radiation source used in [2] was a direct-current arc discharge between copper electrodes. These electrodes were produced by pressing copper dust doped with the required element, in the ratio of one atom of doping element per 1000 copper atoms. The spectrum was detected using a photographic apparatus with exposures of 1 s to 5 min. In addition, a rotating four-step attenuator was used. The relative intensities of the lines of the studied metal were first compared with reference copper lines chosen in various parts of the spectrum. Next, they were reduced to a true relative-intensity scale. This scale was established using precision measurements of the true relative intensities of approximately 200 copper lines carried out with the standard method of multicolor photographic photometry. A tungsten lamp with a ribbon filament and a hydrogen lamp were used as reference sources in various spectral intervals. At short wavelengths $\lambda < 250$ nm, we used a reasonable extrapolation of the transfer coefficients of [2, p. IX]. The errors in the obtained intensities were estimated to be 25%.

The transition probabilities were calculated with the formula

$$gA_{\rm abs} = 0.667 \times 10^{16} C \frac{u}{np} I\lambda e^{E/kT}, \qquad (3)$$

where g is the statistical weight of the upper level, C is a normalization factor, u is the sum over the states, n is the number density of atoms, p is the time of localization of the studied atoms in the discharge zone relative to the time of localization of the copper atoms, I is the line intensity, E is the energy of the upper level, k is the Boltzmann constant, and T is the absolute temperature. We assumed a Boltzmann distribution.

Since all transitions are grouped by their basic characteristic—the common upper level—when determining the branching ratios, the quantities C, u, n, p, E, k, and T are constant for each such group. Therefore,

$$I\lambda/gA = \text{const} \qquad (4)$$

for all transitions with a common upper level.

Therefore, the dependences of $I\lambda/gA$ on the wavelength and line intensity for each group of transitions with a common upper level will form a horizontal line, while the dependence on the level energy will be given by a single point. There should be a set of horizontal lines in the $I\lambda/gA = f(I, \lambda)$ dependences for several levels with different energies. The only process in the radiation source that can lead to deviation from a linear dependence is reabsorption, which is most efficient for resonant radiation. However, Corliss and Bozeman [2] suggest that this factor can be neglected, since the density of the studied atoms is only 0.001 relative to the copper atoms, i.e., it is quite small.

The relative intensities presented in the tables of [12] were obtained in [16, 17] using a method and apparatus very similar to those of [2], i.e., multicolor photographic photometry. According to [18], the typical relative-intensity errors are 20–30%. The most important factor affecting the differences in the line intensities is the conditions for their excitation in the source of radiation. Fortunately, these excitation conditions are not important for the branching ratios, since the lines are grouped according to their common upper level. Therefore, the ratio of $I\lambda$ (data from [12]) to gA (data from [2]) should be the same for each set of lines with a common upper level.

The values of this ratio for upper levels of TaI for the transitions studied here and in [7] are presented in the figures. We adopted the values of I from [8, 9], and λ is expressed in μm. Figures 2–4 show the dependences of $I\lambda/gA$ on λ, I, and $E_{\rm up}$. As is indicated above, the first two of these dependences should form sets of horizontal lines, each corresponding to a particular level $E_{\rm up}$. These lines should collapse to points in the last dependence.

As we can see in Fig. 2, the $I\lambda/gA$ ratios do not form any regular structure, and are located in a broad sloping band. There is a clear tendency for the points corresponding to resonant lines to occupy higher positions. The most likely origin for this behavior is the presence of appreciable reabsorption of the resonant lines under the conditions considered in [2].

The region filled with dots and crosses in Fig. 3 is more compact, but still shows no structure. Four values located far from the center of this region stand out. Two of these, in the bottom left-hand corner, correspond to the extremely weak (according to [12]) 266.862 nm and 272.076 nm lines, which have a relative intensity of unity. The most likely reason for their anomalous locations is a trivial misprint in [12]: $I = 1$ was printed instead of $I = 100$. The cross that is close to them (the 243.991 nm line) can be explained in the same way: $I = 4$ was printed instead of $I = 40$ in [12].

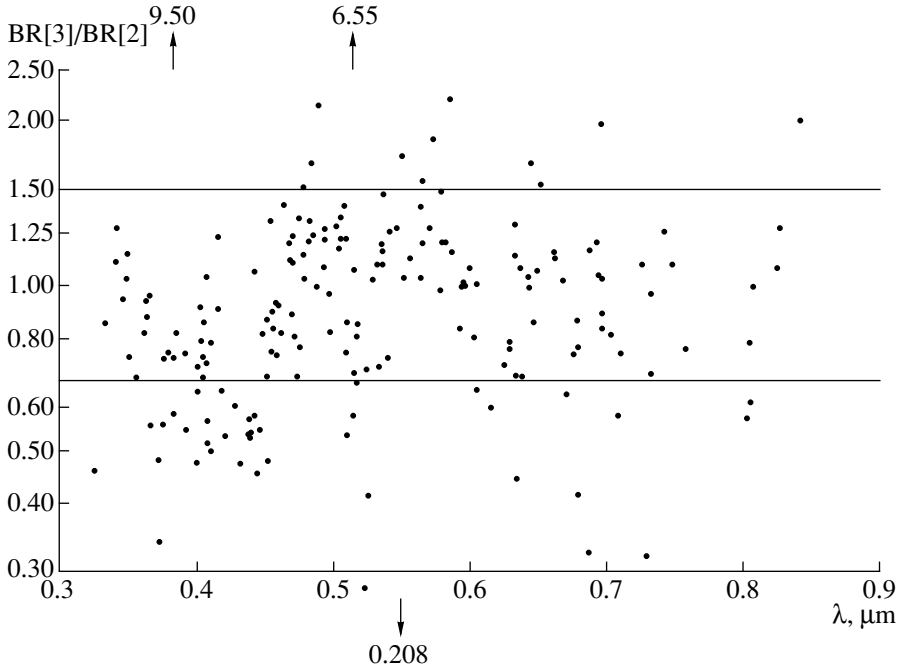

Fig. 6. BR[3]/BR[2] as a function of λ.

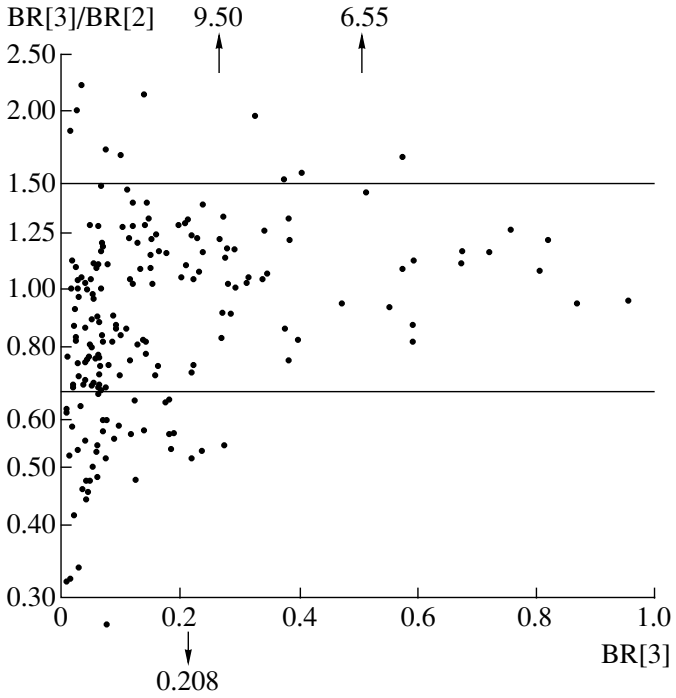

Fig. 7. BR[3]/BR[2] as a function of BR[3].

On the other hand, we could find no satisfactory explanation for the behavior of the 348.462 nm resonant line (the dot in the top left-hand corner of Fig. 3), especially given its ordinary behavior in Figs. 2 and 4.

The distribution of crosses and dots in Fig. 4 is more ordered than in Figs. 2 and 3. The only obvious exception is the three points with very low intensities discussed above. Unfortunately, the main feature that should characterize the dependence of $I\lambda/gA$ on $E_{\rm up}$ is not visible, namely, the collapse of the lines for each

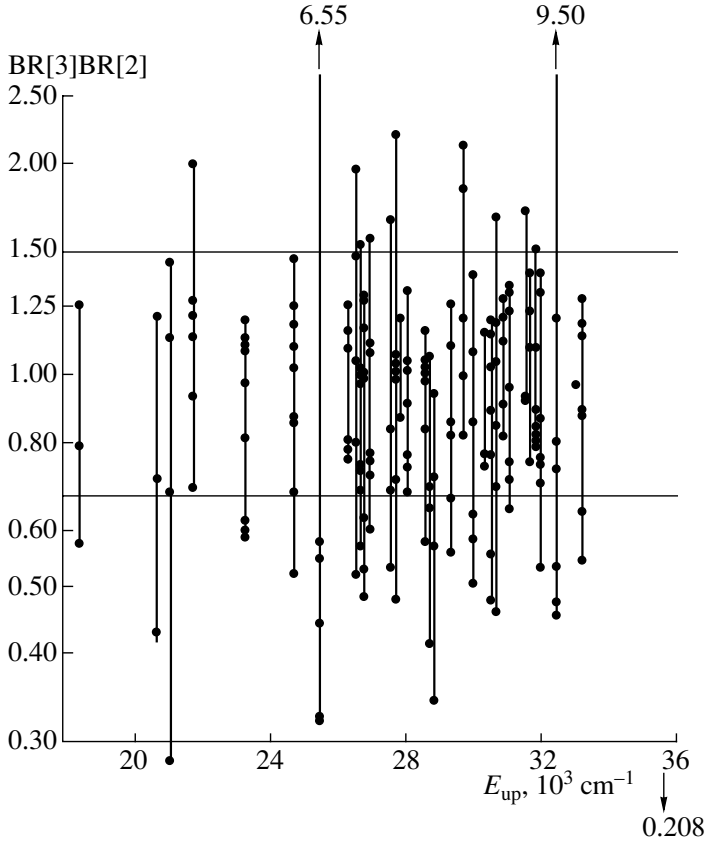

Fig. 8. BR[3]/BR[2] as a function of E_{up}.

level to points. The vertical scatter of dots and crosses for any value of E_{up} characterizes the errors in the branching ratios. We can see in Fig. 4 that these errors are large for most levels, reaching an order of magnitude in some cases. Recall that Figs. 2–4 present only measurements for spectral lines studied in the present work and in [7]; more extensive material is considered in [2, 12].

Figure 5 presents the dependence of $I\lambda/gA$ on λ for a larger number of lines. Most importantly, this figure includes low-intensity transitions from the same levels as in Figs. 2–4, as well as spectral lines beyond the range for our work ($\lambda < 370$ nm), in the visible, up to 900 nm. Transitions from common upper levels are denoted by letters, and resonant transitions are marked by asterisks. We considered only levels from which at least four competing transitions originated. In addition, we added transitions from six low-lying levels, enabling us to extend the dependence to the red. Since the number *1* is very similar to the letter *l*, we have replaced it with *7*. The general pattern confirms the less well defined behavior in Fig. 2. At the same time, the individual notations for transitions from common upper levels indicate that the real scatter of $I\lambda/gA$ for transitions from all levels is considerable (up to two orders of magnitude) when the less intense transitions are included. Therefore, the use of the data of [2] to find branching ratios and total excitation cross sections of energy levels is not advisable.

We can investigate the behavior of the branching ratios for low-lying TaI levels using the data of [3], which presents individual errors for each branching ratio; the relative errors for most lines that are not very weak do not exceed 10%. On the other hand, the standard deviation of the relative oscillator strengths of transitions from low-lying levels in [2] is $\Delta(\log gf_{ik}) = 0.14$ [2, Table 10]. In this case, the total error derived from the data of [2, 3] will be $\Delta(\log gf_{ik}) \approx 0.18$, which corresponds to a factor of 1.5.

Figures 6–8 show the dependences of BR[3]/BR[2] on λ, BR[3], and E_{up}. Overall, these are analogous to the relations in Figs. 2–4, although Fig. 7 presents the dependence on the branching ratios from [3] rather than the line intensity. All data from [3] were used to construct these plots. Since the branching ratios of the same transition are compared for each BR[3]/BR[2], the individual properties of the levels and transitions from them should not be important.

This means that the plots in Figs. 6 and 7 should be horizontal lines with BR[3]/BR[2] $\equiv 1$, and the plot in Fig. 8 should form a set of points with this same value. The vertical scatter of points in Figs. 6–8 is appreciably less than in Figs. 2–4, probably associated with the fact that a considerable contribution to the errors in Figs. 2–4 is made by the data of [12]. In addition, the normalization $\sum_i \mathrm{BR}_{ki} \equiv 1$ for each upper level k imposes constraints on the range of errors in the branching ratios. Nevertheless, an appreciable fraction of the points in Figs. 6–8 also lie beyond the total error corridor for the results of [2, 3] indicated above. Figure 7 shows that this is due primarily to low-intensity lines with BR < 0.2.

5. CONCLUSIONS

We have experimentally studied the excitation cross sections of TaI transitions not connected to the ground level for the first time and have determined the total excitation cross sections (including the contribution of cascade level population) for 23 levels of tantalum. Since branching ratios are sometimes needed to calculate the total cross sections or determine excitation cross sections that cannot be measured experimentally, we analyzed the accuracies of known probabilities of radiative transitions in TaI that can be used to find the required branching ratios. There is no doubt that photoelectric recording of spectra is more accurate and reliable than photographic recording, but photoelectric data remain unavailable for many transitions in atoms with complex electron shells. Our analysis indicates that discrepancies between the data obtained by different authors obtained using similar methods can be appreciably greater than the quoted total errors, even when considering relative intensities. Therefore, the errors in the abundances of metals in stellar atmospheres may in some cases depend more strongly on errors in the atomic constants used than on the errors in the observational data.

REFERENCES

1. D. W. O. Heddle and J. W. Gallagher, Rev. Mod. Phys. **61**, 221 (1989).
2. C. H. Corliss and W. R. Bozeman, *Experimental Transition Probabilities for Spectral Lines of Seventy Elements* (National Bureau of Standards, Washington, DC, 1962; Mir, Moscow, 1968), Monograph 53.
3. E. A. Den Hartog, D. W. Duquette, and J. E. Lawler, J. Opt. Soc. Am. B **4**, 48 (1987).
4. S. Salih, D. W. Duquette, and J. E. Lawler, Phys. Rev. A **27**, 1193 (1983).
5. Yu. M. Smirnov, *Physics of Electronic and Atomic Collisions* [in Russian] (Fiz.-Tekh. Inst. Akad. Nauk SSSR, Leningrad, 1985), p. 183.
6. Yu. M. Smirnov, J. Phys. II **4**, 23 (1994).
7. Yu. M. Smirnov, Dokl. Akad. Nauk **384**, (2002).
8. P. F. A. Klinkenberg, G. J. van den Berg, and J. C. van den Bosch, Physica (Amsterdam) **16**, 861 (1950).
9. G. J. van den Berg, P. F. A. Klinkenberg, and J. C. van den Bosch, Physica (Amsterdam) **18**, 221 (1952).
10. H. Hammerl, G. H. Guthohrlein, M. Elantkovska, *et al.*, Z. Phys. D **33**, 97 (1995).
11. H. Mocnik, B. Arcimowicz, W. Salmhofer, *et al.*, Z. Phys. D **36**, 129 (1996).
12. G. R. Harrison, *Massachusetts Institute of Technology Wavelength Tables* (The Technology Press, MIT, Cambridge, 1939).
13. J. F. Wyart, Phys. Scr. **18**, 87 (1978).
14. J. Dembczynski, B. Arcimowicz, G. H. Guthöhrlein, and L. Windholz, Z. Phys. D **39**, 143 (1997).
15. A. N. Kuchenev and Yu. M. Smirnov, in *Plasmochemistry-89* [in Russian] (Inst. Neftekhim. Sinteza Akad. Nauk SSSR, Moscow, 1989), Part II, p. 227.
16. C. C. Kiess and H. K. Kiess, J. Res. Natl. Bur. Stand. **11**, 277 (1933).
17. C. C. Kiess and E. Z. Stowell, J. Res. Natl. Bur. Stand. **12**, 459 (1934).
18. A. N. Zaĭdel', G. V. Ostrovskaya, and Yu. I. Ostrovskiĭ, *The Technique and Practice of Spectroscopy* [in Russian] (Nauka, Moscow, 1972), p. 311.

Translated by Yu. Dumin

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 851–857. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 944–951.
Original Russian Text Copyright © 2002 by Gasanov, Luk'yanov.

The Libration Points for the Motion of a Star Inside an Elliptical Galaxy

S. A. Gasanov and L. G. Luk'yanov

Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899 Russia

Received August 1, 2001; in final form, May 25, 2002

Abstract—We consider the three-dimensional motion of a star inside an inhomogeneous, rotating, elliptical galaxy with a homothetic density distribution. The libration points inside a triaxial, rotating, inhomogeneous, gravitating ellipsoid are determined and their stability studied. We determine zero-velocity surfaces and regions of possible motion both inside and outside the galaxy. The motions are shown to be stable in the sense given by Hill. The obtained zero-velocity surfaces enable the construction of a Roche model for a galaxy in the form of either a triaxial ellipsoid or an oblate ellipsoid of rotation.
© 2002 MAIK "Nauka/Interperiodica".

1. INTRODUCTION

In [1], we considered the two-dimensional motion of a particle inside an inhomogeneous, rotating, ellipsoidal body with a homothetic density distribution. We derived an expansion of the force function to fourth order in the second eccentricities of the ellipsoid, taken to be small parameters. An expression for the perturbing function was obtained and the equations of the perturbed motion presented in terms of canonical elements were solved.

In [1], we considered an inhomogeneous body T constrained by the ellipsoidal surface

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \qquad a\geq b\geq c \tag{1}$$

and possessing a homothetic (ellipsoidal) density distribution. We took the density ρ of T to be a function of a parameter $p=p(x,y,z)$ and to vary continuously from the center to the outer surface. Based on this, we specified the density of T in the form of a converging power series in the parameter p:

$$\rho=\rho(p)=\rho_0+\sum_{n=1}^{\infty}\rho_n\varepsilon^n p^n \tag{2}$$
$$(0<\varepsilon<1, \quad 0\leq p\leq 1),$$

where ε is a small parameter describing the density distribution. In addition, the following conditions were satisfied:

$$\rho(0)=\rho_0>0, \quad \rho(1)=\rho_f, \quad \rho_n<0 \tag{3}$$
$$(n=1,2,\dots).$$

Here, ρ_0 and ρ_f are the density at the center of the ellipsoidal body T and at its surface, respectively.

Let us consider a star to be a passively gravitating particle P moving inside an elliptical galaxy, which we will approximate by a triaxial ellipsoidal body T rotating with a constant and relatively small angular velocity Ω about the z axis. The coordinates of the star x, y, z are given in a Cartesian coordinate system with the origin at the center of the galaxy and the x, y, and z axes directed along the principal axes of the ellipsoid (1). As follows from the definitions of the second eccentricities λ and μ of the ellipsoid (1),

$$a^2=c^2(1+\lambda^2), \quad b^2=c^2(1+\mu^2) \quad (\mu^2\leq\lambda^2<1). \tag{4}$$

In the two-dimensional case [1], we derived the total potential V in the form of a power series in the coordinates. In the three-dimensional problem, this potential has the form

$$V=V_0+R, \quad V_0=V_{00}-\frac{1}{2}\Big(V_{01}x^2+V_{02}y^2+V_{03}z^2\Big),$$

$$\begin{aligned}R=&\frac{\varepsilon}{4}\Big(\bar R_1x^4+\bar R_2y^4+\bar R_3z^4+2\bar R_4x^2y^2\\&+2\bar R_5x^2z^2+2\bar R_6y^2z^2\Big)+\frac{\varepsilon^2}{6}\Big(S_1x^6+S_2y^6\\&+S_3z^6+3S_4x^2y^4+3S_5x^4y^2+3S_6x^2z^4\\&+3S_7x^4z^2+6S_8x^2y^2z^2+3S_9y^4z^2\\&+3S_{10}y^2z^4\Big)+\dots,\end{aligned} \tag{5}$$

where V_0 is the gravitational potential in the case of unperturbed motion, $V_{00}=\text{const}$ is the potential at the center of the ellipsoid, and R is the perturbing function. All the positive coefficients V_{0i} $(i=0,1,2,3)$, $\bar R_k$ $(k=1,2,\dots,6)$, S_j $(j=1,2,\dots,10)$ are polynomials in the second eccentricities λ and μ, and are presented in full form in the Appendix.

We will assume that λ and μ are small. Consequently, the shape of the galaxy T will be close to

spherical, and we can retain only terms to λ^2 and μ^2 inclusive in the expressions for the coefficients V_{0i} and $\bar{R}_k$. In this case, we have the following relations between the coefficients $\bar{R}_k$:

$$\bar{R}_2 = \alpha^2\bar{R}_1, \quad \bar{R}_3 = \beta^2\bar{R}_1, \quad \bar{R}_4 = \alpha\bar{R}_1, \quad (6)$$
$$\bar{R}_5 = \beta\bar{R}_1, \quad \bar{R}_6 = \alpha\beta\bar{R}_1, \quad \bar{R}_4^2 = \bar{R}_1\bar{R}_2,$$
$$\bar{R}_5^2 = \bar{R}_1\bar{R}_3, \quad \bar{R}_6^2 = \bar{R}_2\bar{R}_3,$$

where

$$\alpha = 1 + \frac{5}{7}(\lambda^2 - \mu^2), \qquad \beta = 1 + \frac{5}{7}\lambda^2. \quad (7)$$

The equations of motion of a star inside an inhomogeneous, gravitating galaxy in a coordinate system xyz that rotates about the z axis with constant angular velocity Ω can be presented in the form

$$\frac{d^2x}{dt^2} - 2\Omega\frac{dy}{dt} = \frac{\partial U}{\partial x}, \quad (8)$$
$$\frac{d^2y}{dt^2} + 2\Omega\frac{dx}{dt} = \frac{\partial U}{\partial y},$$
$$\frac{d^2z}{dt^2} = \frac{\partial U}{\partial z}.$$

Here, we have assumed

$$U = V + \frac{\Omega^2}{2}(x^2 + y^2), \quad (9)$$

where the potential V is specified by (5).

In the general case, the Poincaré inequality, which is a prerequisite for the equilibrium of a rotating ellipsoid, has the form [2, 3]

$$\Omega^2 < 2\pi G\rho K, \quad (10)$$

where $K < 1$ depends only on the shape of the ellipsoid and G is the gravitational constant. Using the potential (5), the inequality (10) can be rewritten in the form [1]

$$\Omega^2 < g^2 - \varepsilon\bar{g}^2, \quad (11)$$

where

$$g^2 = \frac{Q\rho_0}{6}\Big[1 - \frac{3}{10}(\lambda^2+\mu^2) + \frac{3}{56}(3\lambda^4+2\lambda^2\mu^2+3\mu^4)\Big],$$
$$\bar{g}^2 = -\frac{Q\rho_1}{20}\Big[1 - \frac{15}{14}(\lambda^2+\mu^2) + \frac{5}{24}(5\lambda^4+6\lambda^2\mu^2+5\mu^4)\Big].$$

Here, $Q = 4\pi Gab/c^2$ or, as follows from (4), $Q = 4\pi \times G\sqrt{(1+\lambda^2)(1+\mu^2)}$. If we set $\varepsilon = 0$ and restrict our consideration to zeroth-order terms in the expression for g^2, we will obtain $K = ab/(3c^2)$.

Also, as follows from (4),

$$\Omega^2 < g^2 < V_{01} \le V_{02}, \quad (12)$$
$$V_{01} - \Omega^2 = g_1^2 \le g_2^2 = V_{02} - \Omega^2, \quad g_3^2 = V_{03}.$$

Thus, the angular velocity of rotation Ω of the coordinate system satisfies the inequalities (11) or (12).

In this formulation, the solutions of (8) depend on four free parameters: λ, μ, Ω, and ρ_1/ρ_0. Further, we will assume that these parameters are small, and all subsequent operations will take this into account.

2. STEADY-STATE SOLUTIONS FOR THE EQUATIONS OF MOTION

In order to find the steady-state solutions

$$x = x_0 = \text{const}, \qquad y = y_0 = \text{const}, \qquad z = z_0 = \text{const} \quad (13)$$

we derive the following system of algebraic equations from (8):

$$\frac{\partial U}{\partial x} = x_0\Big(-g_1^2 + R_1x_0^2 + R_4y_0^2 + R_5z_0^2\Big) = 0, \quad (14)$$
$$\frac{\partial U}{\partial y} = y_0\Big(-g_2^2 + R_4x_0^2 + R_2y_0^2 + R_6z_0^2\Big) = 0,$$
$$\frac{\partial U}{\partial z} = z_0\Big(-g_3^2 + R_5x_0^2 + R_6y_0^2 + R_3z_0^2\Big) = 0,$$

where

$$R_k = \varepsilon\bar{R}_k \qquad (k = 1, 2, \ldots, 6). \quad (15)$$

Let us consider all possible real solutions for the system (14) specifying certain points in the space xyz, usually called libration points. We will denote these as $\mathrm{L}_k(x_0, y_0, z_0)$ $(k = 1, 2, \ldots)$.

(1) Libration point L_1. This is the zero solution of system (14). Thus, $\mathrm{L}_1 = \mathrm{L}_1(0, 0, 0)$.

(2) Libration points L_2 and L_3. These points are specified by system (14) for $y_0 = 0$ and $z_0 = 0$; i.e.,

$$\mathrm{L}_2 = \mathrm{L}_2\Big(+\frac{g_1}{\sqrt{R_1}}, 0, 0\Big), \quad \mathrm{L}_3 = \mathrm{L}_3\Big(-\frac{g_1}{\sqrt{R_1}}, 0, 0\Big).$$

(3) Libration points L_4 and L_5 are obtained from (14) when $x_0 = 0$ and $z_0 = 0$:

$$\mathrm{L}_4 = \mathrm{L}_4\Big(0, +\frac{g_2}{\sqrt{R_2}}, 0\Big), \quad \mathrm{L}_5 = \mathrm{L}_5\Big(0, -\frac{g_2}{\sqrt{R_2}}, 0\Big).$$

(4) Libration points L_6 and L_7 are obtained from (14) when $x_0 = 0$ and $y_0 = 0$:

$$\mathrm{L}_6 = \mathrm{L}_6\Big(0, 0, +\frac{g_3}{\sqrt{R_3}}\Big), \quad \mathrm{L}_7 = \mathrm{L}_7\Big(0, 0, -\frac{g_3}{\sqrt{R_3}}\Big).$$

Using the equalities (6), it can easily be shown that the differential equations (8) admit only seven steady-state solutions L_k, $(k = 1, 2, \ldots, 7)$. No other steady-state solutions exist.

We can easily verify that all the libration points except for L_1 are located outside the ellipsoid (1).

Indeed, as follows from (3), (11), and (12), for these points,

$$x_{23}^2 \geq \frac{5}{3}\Big(1 - \frac{4}{35}\lambda^2 + \frac{2}{35}\mu^2 - \frac{\Omega^2}{\chi_0}\Big)a^2 \geq \frac{13}{12}a^2, \quad (16)$$

$$y_{45}^2 \geq \frac{5}{3}\Big(1 + \frac{2}{35}\lambda^2 - \frac{4}{35}\mu^2 - \frac{\Omega^2}{\chi_0}\Big)b^2 \geq \frac{13}{12}b^2, \quad (17)$$

$$z_{67}^2 \geq \frac{5}{3}\Big(1 + \frac{2}{35}\lambda^2 + \frac{2}{35}\mu^2\Big)c^2 \geq \frac{5}{3}c^2. \quad (18)$$

Thus, the only physically meaningful libration point L_1 is inside the gravitating ellipsoid T.

3. STABILITY OF THE LIBRATION POINTS L_1

Let us consider the linearized system of the equations of motion in the vicinity of the libration points L_1

$$\frac{d^2x}{dt^2} - 2\Omega\frac{dy}{dt} = -g_1^2 x \quad (19)$$

$$\frac{d^2y}{dt^2} + 2\Omega\frac{dx}{dt} = -g_2^2 y$$

$$\frac{d^2z}{dt^2} = -g_3^2 z.$$

The characteristic equation of system (19) will be

$$(\Lambda^4 + p\Lambda^2 + q)(\Lambda^2 + r) = 0, \quad (20)$$

where

$$p = 4\Omega^2 + g_1^2 + g_2^2, \quad q = g_1^2 g_2^2, \quad r = g_3^2. \quad (21)$$

The roots of (20) can be determined from the formula

$$\Lambda_{1,2,3,4} = \pm\sqrt{-\frac{p}{2} \pm \sqrt{\Delta}}, \qquad \Lambda_{5,6} = \pm\sqrt{-r}, \quad (22)$$

$$\Delta = \frac{p^2}{4} - q.$$

The necessary condition for stability in the sense given by Lyapunov is that the following inequalities be simultaneously satisfied:

$$\Delta \geq 0, \quad p \geq 0, \quad q \geq 0, \quad r \geq 0. \quad (23)$$

The inequalities (23) guarantee that all roots of (20) will be either purely imaginary or zero. Violation of any of the inequalities (23) leads to instability, since real or complex roots with a positive real part then appear among the roots of the characteristic equation.

It is evident that $p > 0$, $q > 0$, and $r > 0$ for the libration points L_1, and

$$\Delta = \frac{1}{4}(g_1^2 - g_2^2)^2 + 2\Omega^2(2\Omega^2 + g_1^2 + g_2^2) > 0. \quad (24)$$

Thus, the Lyapunov stability condition is satisfied, and the libration point L_1 will be stable in this sense in a linear formulation.

To investigate the stability in a non-linear formulation, we consider the Jakoby-type integral for (19):

$$\frac{1}{2}\Big(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\Big) = U + h, \quad (25)$$

where h is the constant of integration. This integral can be written in the form

$$\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + g_1^2x^2 + g_2^2y^2 + g_3^2z^2 - 2R = 2h + 2V_{00}. \quad (26)$$

We can see from this last equation that there exists a small but finite region near the origin of the coordinate system where the left-hand side of the integral (26) is a positively defined function. Taking this to be the Lyapunov function, we will satisfy all requirements of Theorem 1 of the direct method of Lyapunov. Based on this theorem, we can assume that the libration point L_1 is stable in the sense given by Lyapunov.

The same result can be obtained using the Lagrange theorem, since the force function U has an isolated maximum at the point L_1.

4. THE ZERO-VELOCITY SURFACES

No clear-cut description of a method for the construction of zero-velocity surfaces is available in the literature on celestial mechanics and stellar dynamics. Therefore, we will first consider a procedure for constructing zero-velocity surfaces, called Hill surfaces in celestial mechanics and level surfaces in the theory of gravitation.

The fact that the kinetic energy is non-negative makes it possible to determine the region of possible motion from the integral (25) by using the inequality

$$U \geq -h. \quad (27)$$

The zero-velocity surface is the boundary of the region of possible motion

$$U = C, \quad (28)$$

where $C = -h$ is an arbitrary constant.

To properly construct the family of zero-velocity surfaces (28), it is necessary to know the following:

– the singular points of the set of surfaces,

– the type of these singularities,

– the values for the constant C calculated at each singular point, as well as the signs of the inequalities between them, and

– the dependence of the force function at infinity and at the centers of gravity.

Singular points of the family (28) are those at which it is not possible to construct a single tangential plane. To determine such points, we will derive a system of algebraic equations that exactly coincides with the system (14). Therefore, the singular

points are the libration points—rigorous, particular solutions for the initial equations of motion. Note that the infinitely remote libration points [4] must also be known if we are to correctly construct the surfaces (28).

Two types of singular points can exist (excluding singular lines): conical and isolated. To determine the type of a singular point, and with it the behavior of zero-velocity surfaces in its vicinity, the function U can be expanded in a Taylor series in the vicinity of the libration point. Using (14), we obtain in this way

$$U - U_0 = \frac{\partial^2 U(\mathrm{L}_k)}{\partial x^2}(x - x_0)^2 + \frac{\partial^2 U(\mathrm{L}_k)}{\partial y^2}(y - y_0)^2 + \frac{\partial^2 U(\mathrm{L}_k)}{\partial z^2}(z - z_0)^2 + \cdots = C - C_0, \quad (29)$$

where x_0, y_0, and z_0 are the coordinates of the libration point L_k and $U_0 = C_0$ is the potential U at this point.

The expansion (29) is given for $\varepsilon = 0$; otherwise, we must take into account the mixed derivatives for U that will arise from R in (5).

When $C = C_0$, we find from (29) the surface containing the point L_k. Maintaining only up to second-order terms, we can approximately represent this as the second-order surface

$$\frac{\partial^2 U(\mathrm{L}_k)}{\partial x^2}(X - x_0)^2 + \frac{\partial^2 U(\mathrm{L}_k)}{\partial y^2}(Y - y_0)^2 + \frac{\partial^2 U(\mathrm{L}_k)}{\partial z^2}(Z - z_0)^2 = 0, \quad (30)$$

where X, Y, and Z are the coordinates of the current point of this surface. If the signs of the coefficients in the last equality are different, the singular point is conical. The $X - x_0$, $Y - y_0$, or $Z - z_0$ axis corresponding to the term in (30) with the sign that does not coincide with those of the other two terms is the cone axis. When the coefficients in (30) have the same sign, the singular point is isolated (an imaginary cone).

Knowing the type of singular point, the behavior of the zero-velocity surface for $C \neq C_0$ can be determined in the vicinity of this point. For a conical singular point inside the cone, this surface is close to a two-sheet hyperboloid, and the surface is close to a one-sheet hyperboloid if the singular point is outside the cone; in both cases, the cone is an asymptote for the hyperboloids.

In the vicinity of an isolated singular point, the zero-velocity surface for $C < C_0$ is close to an ellipsoid, while the surface does not exist for $C > C_0$. The signs of the inequality here may be opposite.

Knowing the coordinates for the singular points, the constants C (or h) at each point can easily be calculated and ranked in descending (ascending) order. The libration points will appear in this order during the construction of the family of surfaces (28) decreasing C from $+\infty$ to the minimum values.

The behavior of the force function at infinity is specified by the $\Omega^2(x^2 + y^2)/2$ term in (9); i.e., the surface is close to a right circular cylinder with its axis coincident with the z axis. The value for U at the center of the ellipsoid is V_{00}.

Such analyses indicate that the only libration point L_1 is situated inside the ellipsoid T. This is an isolated singular point, at which the constant C is equal to V_{00}.

All the zero-velocity surfaces form ovals enclosing the point L_1. Therefore, the regions of possible motion of a star inside an elliptical galaxy are always restricted. The motions inside the galaxy are stable in the sense given by Hill.

For simplicity, let us consider the construction of zero-velocity surfaces in the space outside the ellipsoid T when the ellipsoid is assumed to be homogeneous. A rigorous allowance for the homothetic mass distribution only affects the locations of the singular points and zero-velocity surfaces, while leaving all qualitative conclusions unchanged.

Batrakov [5] showed that four libration points exist in the outer region of a homogeneous, triaxial, rotating ellipsoid, located symmetrically about the center of the ellipsoid on extensions of the major and minor axes in the xy plane. The stability of these points was later studied by Abalakin [6], Zhuravlev [7] and others. The libration points L_2 and L_3, located on the extension of the major axis of the ellipsoid, are instable in the sense given by Lyapunov, while the points L_4 and L_5 are stable to first approximation. A non-linear analysis indicated that L_4 and L_5 are stable for most initial conditions in terms of the Lebesgue measure, with the exception of some resonant cases, when instability occurs.

We are not aware of any studies of zero-velocity surfaces and regions of possible motion for the problem at hand. The existence of libration points at infinity and a fictitious (i.e., physically meaningless) point at the center of the ellipsoid are not mentioned in [5–7].

The differential equations for the motion of a body outside a homogeneous, gravitating ellipsoid have the same form as (8), while the potential V is

$$V = GM\left(\frac{1}{r} + \varepsilon\frac{\alpha_0 x^2 + \beta_0 y^2 + \gamma_0 z^2}{r^5} + \ldots\right), \quad (31)$$

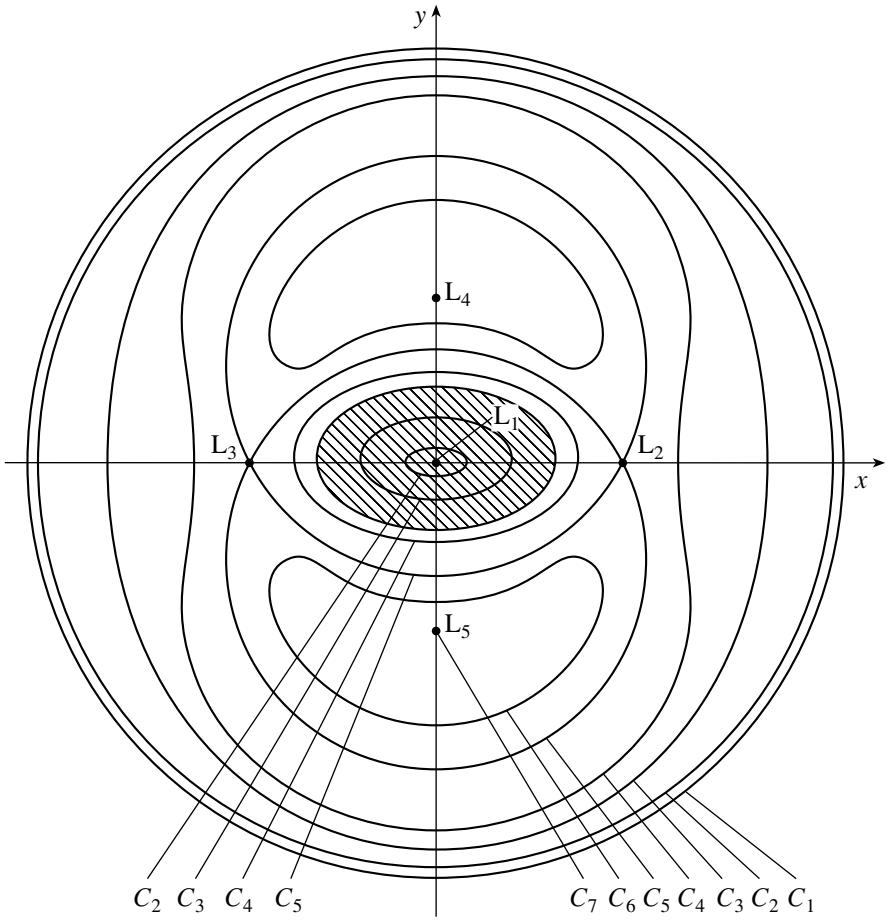

Fig. 1. Cross section of the zero-velocity surfaces in the xy plane. The cross section of the gravitating ellipsoid (the galaxy) is shaded. $C_i > C_j$ for $i < j$, $C_5 = C_{23}$, $C_7 = C_{45}$, $C_{11} = V_{00}$, $C_1 > C_{11} > C_2$.

where

$$r = \sqrt{x^2 + y^2 + z^2}, \qquad \varepsilon\alpha_0 = \frac{3}{10}\frac{\alpha'}{a_0^2}, \qquad (32)$$

$$\varepsilon\beta_0 = \frac{3}{10}\frac{\beta'}{a_0^2}, \qquad \varepsilon\gamma_0 = \frac{3}{10}\frac{\gamma'}{a_0^2},$$

$$a_0^3 = \frac{GM}{\Omega^2}, \qquad \alpha' = a^2 - R^2,$$

$$\beta' = b^2 - R^2, \qquad \gamma' = c^2 - R^2,$$

and M is the mass of the body T. It is assumed that ε is small and positive; i.e., the ellipsoid differs little from a sphere with radius R.

In this formulation, the equations of motion allow the existence of the following libration points:

$$\mathrm{L}_1 = \mathrm{L}_1(0,0,0), \quad \mathrm{L}_2 = \mathrm{L}_2(1+\varepsilon\alpha_0,0,0), \quad (33)$$
$$\mathrm{L}_3 = \mathrm{L}_3(-1-\varepsilon\alpha_0,0,0), \quad \mathrm{L}_4 = \mathrm{L}_4(0,1+\varepsilon\beta_0,0),$$
$$\mathrm{L}_5 = \mathrm{L}_5(0,-1-\varepsilon\beta_0,0), \quad \mathrm{L}_{\pm\infty} = \mathrm{L}_{\pm\infty}(0,0,\pm\infty).$$

Of course, the libration point L_1 here is physically meaningless. However, it coincides with the point considered above, which does have physical meaning.

To construct the zero-velocity surfaces, all the auxiliary calculations described above must be carried out. The singular points L_k $(k = 2,3,4,5)$ prove to be conical, while the infinitely remote points $\mathrm{L}_{\pm\infty}$ are isolated. Let the values for C at the corresponding libration points be

$$C(\mathrm{L}_2) = C(\mathrm{L}_3) = C_{23}, \quad C(\mathrm{L}_4) = C(\mathrm{L}_5) = C_{45},$$
$$C(\mathrm{L}_{\pm\infty}) = C_{\pm\infty} = 0.$$

It is easy to establish inequalities relating these values:

$$C_{23} > C_{45} > C_{\pm\infty}. \qquad (34)$$

Figures 1 and 2 illustrate the zero-velocity surfaces and regions of possible motion inside and outside the gravitating ellipsoid, by presenting the cross sections formed by the xy and yz planes, respectively.

We will consider variations of the surfaces and regions of possible motion as C is decreased, starting with $C = \infty$. We will use the potential (5) for surfaces inside the ellipsoid and the potential (31) outside the ellipsoid. The process begins with a surface that is

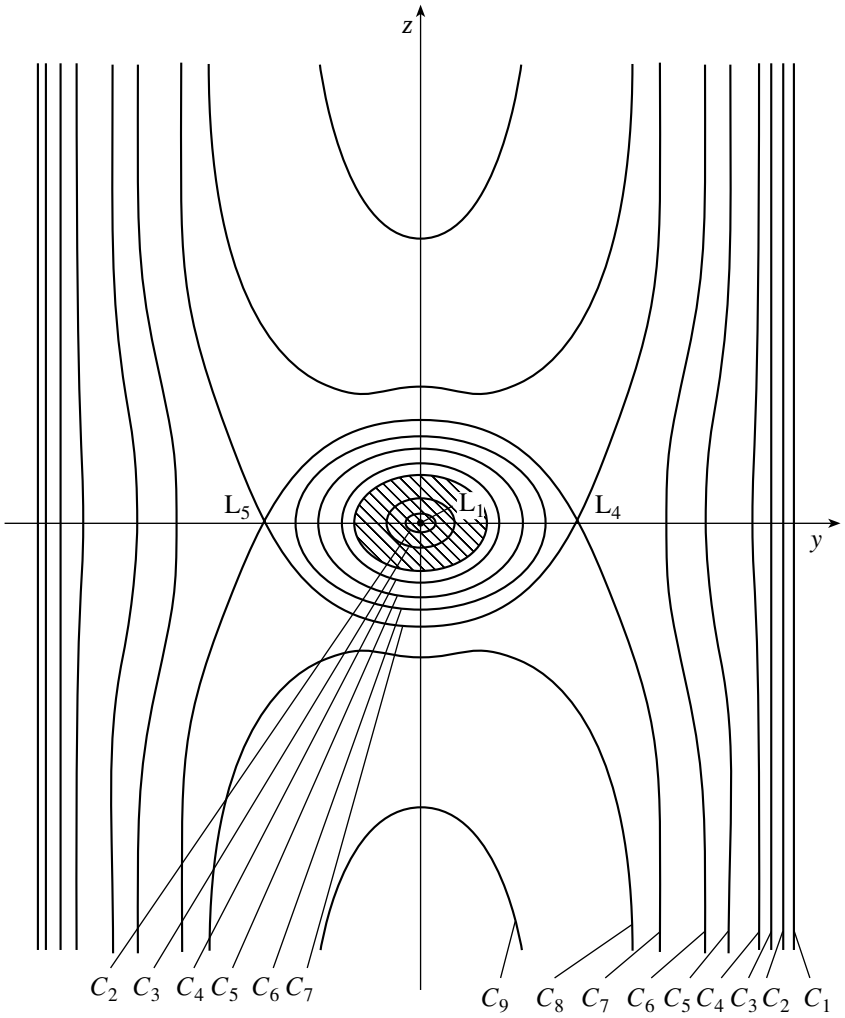

Fig. 2. Cross section of the zero-velocity surfaces in the yz plane. The notation is as in Fig. 1.

close to a cylinder with a very large radius and its axis coincident with the z axis (the curve C_1 in Figs. 1, 2).

When $C = V_{00}$, the point L_1 appears at the center of the ellipsoid, while some oval-shaped regions of possible motion around L_1 appear when $C < V_{00}$ (the curves C_2 and C_3 in Figs. 1 and 2). At some C, which depends on the size of the ellipsoid T, the oval curves around L_1 exceed the boundaries of the ellipsoid (the curve C_4). If the size of the gravitating ellipsoid is negligibly small, such ovals appear for very large C.

Further decreasing C to C_{23} results in the contact (or, more precisely, in the self-intersection) of the inner oval with the outer deformed cylinder. A "passage" is then formed through the cylinder, making it possible for a moving body to move to an infinite distance from the ellipsoid (the curves C_5, C_6 in Figs. 1, 2).

Further, both "stems" of the cylinder in the xy plane narrow to zero for C_{45} (the curve C_7 in Figs. 1, 2), after which the surface breaks into two parts and does not intersect the xy plane (the curve C_8 in Fig. 2). Each part then constricts into one of the infinitely remote libration points (the curve C_9 in Fig. 2).

If we consider an oblate ellipsoid of rotation ($a = b > c$) instead of a triaxial ellipsoid, a circle with its center at the coordinate origin will appear instead of the four singular points L_k ($k = 2, 3, 4, 5$). In this case, as C decreases, the zero-velocity surfaces will intersect the inner oval simultaneously along the entire singular circle for $C = C_{23} = C_{45}$ and will then similarly constrict to the infinitely remote points. These surfaces can be represented schematically by a rotation of the curves in Fig. 2 about the z axis.

Our analysis indicates that the motion of bodies in the outer space is stable for $C > C_{23}$ in the sense given by Hill. In other words, when this inequality is satisfied, the star will always be located in the vicinity of the elliptical galaxy, inside the $C = C_{23}$

zero-velocity surfaces (the finite region restricted by the curve C_5 in Figs. 1, 2).

When $C < C_{23}$, two "funnels" arise on the zero-velocity surfaces near L_2 and L_3, through which the star can escape to an infinite distance from the galaxy.

The $C = C_{23}$ surface can be used as a Roche model, while the two "funnels" around the libration points L_2 and L_3 can explain the existence of pairs of arms in spiral galaxies. If the shape of the galaxy is represented by an oblate ellipsoid of rotation, the entire circle becomes a "funnel," and a disk can arise around the galaxy instead of spiral arms.

Of course, we have assumed that the gravitating galaxy entirely fills its Roche lobe (the region inside the closed part of the curves C_5 in Figs. 1, 2), and that the angular velocity of its rotation reaches its limiting values in the Poincaré inequality.

5. CONCLUSION

We have considered the three-dimensional motion of a star inside an inhomogeneous, rotating, elliptical galaxy with a homothetic density distribution. We have derived a fourth-order expansion of the force function for this problem in the second eccentricities of the ellipsoid, taken to be small parameters. The coordinate expansion of the perturbing function included up to sixth order terms.

The steady-state solutions for the equations of motion (libration points) show that the libration point at the center of the galaxy is stable in the sense given by Lyapunov. We have determined the zero-velocity surfaces and regions of possible motion inside and outside the galaxy. The motion inside the galaxy is stable in the sense given by Hill; i.e., it occurs in some closed region. It is possible to construct a Roche model for a galaxy having the shape of a triaxial ellipsoid or oblate ellipsoid of rotation.

REFERENCES

1. S. A. Gasanov, Pis'ma Astron. Zh. **27**, 150 (2001) [Astron. Lett. **27**, 124 (2001)].
2. H. Poincaré, *Lecons sur les hypotheses cosmogoniques* (A. Hermann, Paris, 1911).
3. K. F. Ogorodnikov, *Dynamics of Stellar Systems* (Fizmatgiz, Moscow, 1958; Pergamon, Oxford, 1965).
4. L. G. Luk'yanov, Astron. Zh. **66**, 385 (1989) [Sov. Astron. **33**, 194 (1989)].
5. Yu. V. Batrakov, Byull. ITA **6**, 524 (1957).
6. V. K. Abalakin, Byull. ITA **6**, 543 (1957).
7. S. G. Zhuravlev, Astron. Zh. **51**, 1330 (1974) [Sov. Astron. **18**, 792 (1974)].

APPENDIX

$$V_{00} = \frac{c^2}{2}\Big(3\chi_0 - \frac{5}{2}\varepsilon\chi_1 c^2 - \frac{7}{3}\varepsilon^2\chi_2 c^4\Big) \times\Big[1 - \frac{1}{6}(\lambda^2+\mu^2)\Big],$$

$$V_{01} = \chi_0\Big[1 - \frac{3}{10}(3\lambda^2+\mu^2)\Big],$$

$$V_{02} = \chi_0\Big[1 - \frac{3}{10}(\lambda^2+3\mu^2)\Big],$$

$$V_{03} = \chi_0\Big[1 - \frac{3}{10}(\lambda^2+\mu^2)\Big],$$

$$\bar{R}_1 = \chi_1\Big[1 - \frac{5}{14}(5\lambda^2+\mu^2)\Big],$$

$$\bar{R}_2 = \chi_1\Big[1 - \frac{5}{14}(\lambda^2+5\mu^2)\Big],$$

$$\bar{R}_3 = \chi_1\Big[1 - \frac{5}{14}(\lambda^2+\mu^2)\Big],$$

$$\bar{R}_4 = \chi_1\Big[1 - \frac{15}{14}(\lambda^2+\mu^2)\Big],$$

$$\bar{R}_5 = \chi_1\Big[1 - \frac{5}{14}(3\lambda^2+\mu^2)\Big],$$

$$\bar{R}_6 = \chi_1\Big[1 - \frac{5}{14}(\lambda^2+3\mu^2)\Big],$$

$$S_1 = \chi_2\Big[1 - \frac{7}{18}(7\lambda^2+\mu^2)\Big],$$

$$S_2 = \chi_2\Big[1 - \frac{7}{18}(\lambda^2+7\mu^2)\Big],$$

$$S_3 = \chi_2\Big[1 - \frac{7}{18}(\lambda^2+\mu^2)\Big],$$

$$S_4 = \chi_2\Big[1 - \frac{7}{18}(3\lambda^2+5\mu^2)\Big],$$

$$S_5 = \chi_2\Big[1 - \frac{7}{18}(5\lambda^2+3\mu^2)\Big],$$

$$S_6 = \chi_2\Big[1 - \frac{7}{18}(3\lambda^2+\mu^2)\Big],$$

$$S_7 = \chi_2\Big[1 - \frac{7}{18}(5\lambda^2+\mu^2)\Big],$$

$$S_8 = \chi_2\Big[1 - \frac{7}{6}(\lambda^2+\mu^2)\Big],$$

$$S_9 = \chi_2\Big[1 - \frac{7}{18}(\lambda^2+5\mu^2)\Big],$$

$$S_{10} = \chi_2\Big[1 - \frac{7}{18}(\lambda^2+\mu^2)\Big], \quad \chi_0 = \frac{Q\rho_0}{3},$$

$$\chi_1 = -\frac{Q\rho_1}{5c^2}, \quad \chi_2 = -\frac{Q\rho_2}{7c^4}$$

Translated by K. Maslennikov

Astronomy Reports, Vol. 46, No. 10, 2002, pp. 858–865. Translated from Astronomicheskiĭ Zhurnal, Vol. 79, No. 10, 2002, pp. 944–951.
Original Russian Text Copyright © 2002 by Akulenko, Kumakshev, Markov, Rykhlova.

Forecasting the Polar Motions of the Deformable Earth

L. D. Akulenko[1], S. A. Kumakshev[1], Yu. G. Markov[2], and L. V. Rykhlova[3]

[1] *Institute for Problems of Mechanics, Russian Academy of Sciences, pr. Vernadskogo 101, Moscow, 117526 Russia*

[2] *Moscow Aviation Institute, Volokolamskoe sh. 4, Moscow, 125080 Russia*

[3] *Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia*

Received December 10, 2001; in final form, May 23, 2002

Abstract—A mathematical model for the complicated phenomenon of the polar oscillations of the deformable Earth that adequately describes the astrometric data of the International Earth Rotation Service is constructed using celestial mechanics and asymptotic techniques. This model enables us to describe the observed phenomena (free nutation, annual oscillations, and trends) simply and with statistical reliability. The model contains a small number of parameters determined via a least-squares solution using well-known basis functions. Interpolations of the polar trajectory for intervals of 6 and 12 yrs and forecasts for 1–3 yrs are obtained using the theoretical curve. The calculated coordinates demonstrate a higher accuracy than those known earlier. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

Numerous astrometric studies are based on dynamical theories of the Earth's rotation with respect to its center of mass [1–4]. Observations obtained beginning at the end of the 19th century show that the Earth's rotational axis changes its orientation in the course of time with respect to both fixed and inertial coordinate systems. The determination of the position of the instantaneous rotational axis and the motion of the Earth's center of mass is important for both theoretical and applied purposes [1].

Both components of the motion are extremely complicated for an exhaustive high-accuracy analysis because of the gravitational perturbations due to the Sun, Moon, Jupiter, and other planets and to various geophysical phenomena. Studies of the rotation about the center of mass are hindered by the difficulty in constructing high-accuracy dynamical models of the deformable Earth. A number of firmly detected and well determined components of the polar motion (the latitude variations) cannot be described using a model with a perfectly rigid Earth. Known theories of rotation [2–4] use various models for the internal structure of the planet taking into account numerous geophysical perturbations. However, there is no doubt that these theories do not adequately describe the Earth's rotation. There is no simple, reliable model containing relatively few parameters that can be determined observationally. It is well known [1] that, as a rule, even short-term forecasts (for about one hundred days) fail and require frequent (weekly) corrections based on observations (http://hpiers.obspm.fr/eoppc/eop/eopc04/eopc04-xy.gif).

We propose here a mathematical model that adequately represents the astrometric data of the International Earth Rotation Service (IERS) and suggests a rational explanation for the observed properties of the motion based on celestial mechanics. The model makes it possible to obtain high-accuracy interpolations of the polar trajectory for time intervals of 6–20 yrs and to perform forecasts for 1–3 yrs. The modified model satisfactorily describes the complex Chandler motion of the pole, which includes natural oscillations (free nutation) with a period $T_1 = 433$ sidereal days, annual oscillations (nutation forced by the solar gravitational moment) with a period of $T_h = 365$ sidereal days, and a slow, irregular drift (trend) of the Earth's figure axis at a rate of $\sim 0''.005$ per year in the direction of North America ($\sim 90°$ westward from Greenwich). In addition to these main components of the motion, the model admits future developments and refinements to take into account less important effects (both gravitational and geophysical). However, these developments may require more complex measuring systems and computations that are not currently justified.

2. FORMULATION OF THE PROBLEM. COORDINATE SYSTEMS

In order to adequately describe the rotational motion of the deformable Earth and the Chandler wobble of the pole, we use a simple mechanical model for a viscoelastic, rigid body. We assume that the planet

1063-7729/02/4610-0858$22.00 © 2002 MAIK "Nauka/Interperiodica"

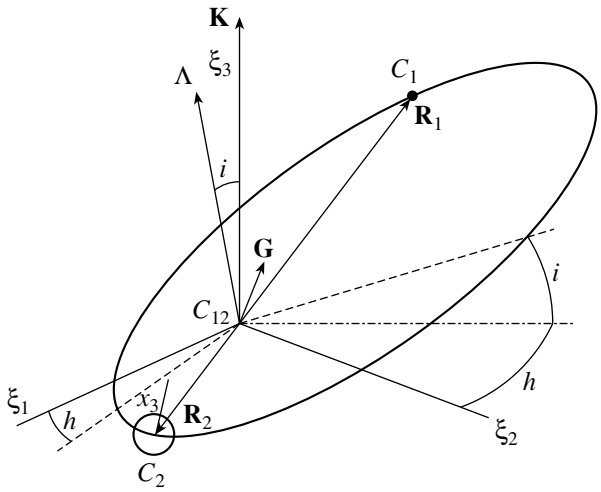

Fig. 1. Coordinate system for the two-body problem and directions of vectors.

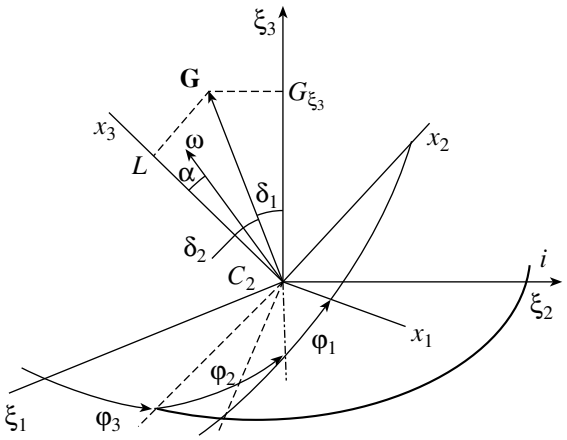

Fig. 2. Relative orientation of the coordinate system fixed to the deformable Earth and the reference system and Andoiyee variables.

consists of a perfectly rigid (spherical) core and viscoelastic mantle [5, 6]. There is no motion of the elastic medium at the inner boundary, while the outer boundary is free. We assume that this medium can be described using a linear theory of viscoelasticity, with deformations being quasi-static. These assumptions enable us to apply rigorous theorems of mechanics and perturbation and averaging methods [7–10] to construct a simple and reasonable model for the Earth's rotation.

In the initial stage of our investigation of the polar motion and its evolution under the action of perturbing moments, we consider the spatial two-body problem. We assume that a deformable planet (Earth) and point satellite (Moon) undergo mutual progressive and rotational motion about their center of mass (barycenter). Let us briefly indicate the main geometric properties of this problem [5, 6].

We introduce a fixed coordinate system $C_{12}\xi_1\xi_2\xi_3$ with its origin at the barycenter C_{12}. Let $\mathbf{G}$ be the proper angular momentum of the Earth and $\mathbf{\Lambda}$ be the orbital angular momentum of the centers of mass of the Moon (C_1) and Earth (C_2). Then, the axis of the total angular momentum of the entire system is stationary $\mathbf{K} = \mathbf{G} + \mathbf{\Lambda}$ in an inertial space and coincides with the $C_{12}\xi_3$ axis (Fig. 1), if external perturbations are neglected. In the $C_{12}\xi_1\xi_2\xi_3$ coordinate system, the radius vectors $\mathbf{R}_1$ and $\mathbf{R}_2$ of the points C_1 and C_2 take the form

$$\mathbf{R}_j = R_j\mathbf{R}_j^0, \qquad R_j = c_j^* R_{21}, \qquad (1)$$
$$\mathbf{R}_2^0 = -\mathbf{R}_1^0, \qquad j = 1, 2,$$
$$\mathbf{R}_j^0 = \pm\Gamma_3(h)\Gamma_1(i)(\cos\vartheta, \sin\vartheta, 0)^T,$$
$$c_1^* = \frac{m_2}{m}, \quad c_2^* = \frac{m_1}{m}, \quad m = m_1 + m_2,$$
$$\Gamma_1(i) = \mathrm{diag}(1, \Pi_2(i)), \quad \Gamma_3(h) = \mathrm{diag}(\Pi_2(h), 1),$$

where $\Gamma_{1,3}$ are block-diagonal matrices and Π_2 is the planar-rotation matrix. The angular variables are the longitude of the ascending node, orbital inclination, and true anomaly. The radius vector $\mathbf{R}_{21}$ connecting the Earth's and Moon's centers of mass (i.e., C_2 and C_1) takes the form $\mathbf{R}_{21} = R_{21}\mathbf{R}_{21}^0$, so that $\mathbf{R}_{21}^0 = \mathbf{R}_1^0$.

We rigidly fix the $C_2x_1x_2x_3$ Cartesian coordinate system to the rigid core of the planet, with the axes directed along the main axes of inertia A, B, and C. In this coordinate system, the $\mathbf{R}_{21}^0$ unit vector is defined as follows:

$$S^{-1}\mathbf{R}_{21}^0 = (\gamma_1, \gamma_2, \gamma_3)^T, \qquad (2)$$
$$S^{-1} = \Gamma_3^{-1}(\varphi_1)\Gamma_1^{-1}(\delta_2)\Gamma_3^{-1}(\varphi_2)\Gamma_1^{-1}(\delta_1)\Gamma_3^{-1}(\varphi_3).$$

Here, the orthogonal matrix $S = S(t)$ in (2) ($S^{-1} = S^T$) specifies the transformation from the fixed to the inertial axes, and is expressed in terms of the canonical Andoiyee variables: the angular momenta L, G, and $G_{\xi 3}$ and the angular variables φ_1, φ_2, and φ_3. The angles δ_1 and δ_2 are determined by the relations (Fig. 2)

$$\cos\delta_1 = \frac{G_{\xi 3}}{G}, \qquad \cos\delta_2 = \frac{L}{G}, \qquad (3)$$

where G is the absolute value of the Earth's angular momentum and $G_{\xi 3}$ is its projection onto the ξ_3 inertial axis.

The mutual orbital motion of the centers of mass C_1 and C_2 is described by the canonical Delone variables Λ, H, ϑ, and h, where H is the projection of $\mathbf{\Lambda}$ onto the $C_{12}\xi_3$ axis; i.e., the axis of the total angular momentum $\mathbf{K}$. Λ and H are equal to $\Lambda = |\mathbf{\Lambda}|$ and $H = \Lambda\cos i$.

3. FREE NUTATION

To study free nutation (the Chandler component of the polar motion), it is convenient to substitute the Hamiltonian action-angle variables I_j and w_j ($j = 1, 2, 3$) for the Andoiyee variables. The main dynamical parameters κ and λ describe the Earth's rotational motions:

$$\kappa^2 = \frac{C^*(A^* - B^*)}{A^*(B^* - C^*)}, \qquad \lambda^2 = \kappa^2 \frac{2EC^* - I_2^2}{I_2^2 - 2EA^*}. \quad (4)$$

Here, A^*, B^*, and C^* are the effective main central moments, taking into account deformations of the "frozen" Earth due to the combined motion of its proper rotation and the motion with respect to the barycenter C_{12}. The constant E has the physical meaning of the integral of the kinetic energy of the intermediate motion (rotation) of the deformable Earth, and $I_2 = G$ (see below). The relations between the action-angle variables I_j, w_j ($j = 1, 2, 3$) and the canonical Andoiyee variables L, G, $G_{\xi 3}$, and φ_j are expressed via elliptic integrals and Jacoby functions. We obtain for the Earth's rotation the relations [6]

$$I_1 = \frac{2I_2}{\pi\kappa}\frac{\kappa_*}{\chi}[\chi^2 \Pi(\frac{\pi}{2}, \kappa^2, \lambda) - \lambda^2 \mathbf{K}(\lambda)], \quad (5)$$

$$\kappa_* = \sqrt{1+\kappa^2}, \quad \chi = \sqrt{\kappa^2 + \lambda^2}, \quad 0 < \lambda < 1,$$

$$I_2 = G, \quad I_3 = G_{\xi 3},$$

$$w_1 = \pm\frac{\pi}{2}\frac{F(\zeta, \lambda)}{\mathbf{K}(\lambda)}, \quad \tan\zeta = -\kappa_*^{-1}\mathrm{cotg}\varphi_1,$$

$$w_2 = \varphi_2 \pm \frac{\kappa_*\chi}{\kappa}\left[\Pi(\zeta, \kappa^2, \lambda) - \Pi(\frac{\pi}{2}, \kappa^2, \lambda)\frac{F(\zeta, \lambda)}{\mathbf{K}(\lambda)}\right],$$

$$\zeta = \pm am(\eta, \lambda), \quad \eta = \frac{2}{\pi}\mathbf{K}(\lambda)w_1, \qquad w_3 = \varphi_3.$$

Here, $F(\zeta, \lambda)$ and $\Pi(\zeta, \kappa^2, \lambda)$ are elliptical integrals of the first and third kind, respectively, and $\mathbf{K}(\lambda)$ and $\Pi(\pi/2, \kappa^2, \lambda)$ are the corresponding complete elliptical integrals.

We can obtain the Routh functional for the intermediate model problem and construct trajectories in the phase space I_j, w_j using the action-angle variables in accordance with (4) and (5). In particular, we find

$$R_0 = \frac{1}{2}\frac{I_2^2}{A^*}\left(1 - \frac{(C^* - A^*)}{C^*}\frac{\kappa^2}{\chi^2}\right), \quad (6)$$

$$I_i(t) = I_i^0 = \mathrm{const}, \qquad w_3(t) = w_3^0,$$

$$w_{1,2}(t) = n_{1,2}t + w_{1,2}^0.$$

Here, the angles (phases) w_1 and w_2 describe the Chandler wobble and diurnal rotation of the deformable Earth with the frequencies n_1 and n_2, respectively; we obtain in accordance with (6) the analytic expressions

$$n_1 = -\frac{\pi}{2}I_2\frac{C^* - A^*}{A^*C^*}\kappa\kappa_*^{-1}\chi^{-1}\mathbf{K}^{-1}(\lambda), \quad (7)$$

$$n_2 = \frac{I_2}{C^*}\left(1 + \frac{(C^* - A^*)}{A^*}\Pi(\frac{\pi}{2}, \kappa^2, \lambda)\mathbf{K}^{-1}(\lambda)\right).$$

Note that the parameter λ for the Earth is very small: $\lambda^2 \approx 1.36 \times 10^{-14}$, and we can simplify expressions (4)–(7). The corresponding calculations yield the period $T_1 = 2\pi/n_1 \approx 430$ sidereal days, in good agreement with observations [1].

Thus, the calculated frequencies of the Chandler wobble n_1 (7) and of the Earth's diurnal rotation n_2 agree well with the observed values. They are very stable over time intervals on the order of the precessional period (~ 26 thousand years). Apparently, substantial changes in the frequencies n_1 and n_2 due to the dissipative evolution of the dynamical characteristics (energy and angular momentum) of the viscoelastic planet can occur only very slowly, over time intervals considerably exceeding the precessional period [11].

We define the polar motion as the angular displacement of the rotational axis with respect to coordinates fixed to the body of the planet. The components of the angular-velocity vector are expressed in terms of the phase w_1 in the form of an expansion in the small parameter $\varepsilon = \lambda\kappa^{-1} \sim 10^{-6}$. With a relative error of $O(\varepsilon^4) \sim 10^{-24}$, the components of the angular velocity are

$$\omega_1 = \lambda I_2\frac{\mathrm{cn}(u, \lambda)}{A^*\chi} = \varepsilon\frac{I_2}{A^*}\left(\cos w_1 + \frac{\varepsilon^2}{16} \times [-(\kappa^2 + 8)\cos w_1 + \kappa^2 \cos 3w_1]\right) + O(\varepsilon^5), \quad (8)$$

$$\omega_2 = \lambda I_2\frac{\kappa_*\mathrm{sn}(\eta, \lambda)}{B^*\chi} = \varepsilon\sqrt{1+\kappa^2}\frac{I_2}{B^*}\left(\sin w_1 + \frac{\varepsilon^2}{16} \times [(\kappa^2 - 8)\sin w_1 + \kappa^2 \sin 3w_1]\right) + O(\varepsilon^5),$$

$$\omega_3 = \frac{I_2}{C^*}\frac{\kappa\mathrm{dn}(u, \lambda)}{\chi} = \frac{I_2}{C^*} \times \left(1 + \frac{\varepsilon^2}{4}[-2 - \kappa^2 + \kappa^2 \cos 2w_1]\right) + O(\varepsilon^4)],$$

$$u = \frac{2}{\pi}\mathbf{K}(\lambda)w_1, \qquad \varepsilon = 1.2 \times 10^{-6}.$$

The angular coordinates describing the free nutation (Chandler component x_c, y_c), corresponding to the angle α between the rotational axis and the figure axis, and the linear coordinates in a plane tangential

to the geoid, are approximately equal to ([6, 11] and see (8))

$$x_c = \frac{\omega_x}{\omega} \approx \varepsilon \frac{C^*}{A^*} \cos w_1, \quad y_c = -\varepsilon \sqrt{1+\kappa^2} \frac{C^*}{B^*} \sin w_1, \tag{9}$$

$$X_c = Rx_c, \quad Y_c = Ry_c;$$

$$\cos\alpha = \frac{\omega_3}{\omega}$$

$$\approx 1 - \frac{\varepsilon}{2}\left(\frac{C^{*2}}{A^{*2}}\cos^2 w_1 + (1+\kappa^2)\frac{C^{*2}}{B^{*2}}\sin^2 w_1\right),$$

$$\max|X_c|, |Y_c| \approx 7.5 \text{ m}, \quad R = 6.38 \times 10^6 \text{ m}.$$

In a first approximation in ε, the polhode (free nutation) becomes an ellipse with a very small eccentricity $e \approx 0.005$. The theoretical estimates (8), (9) are in agreement with the IERS data [1].

4. THEORETICAL MODEL FOR THE EARTH'S ROTATION

The Chandler wobble (x_c, y_c) makes one of the main contributions to the complex polar trajectory, which resembles a spiral with a moving center. The data detected and processed by the IERS in 1980–2000 clearly detect oscillations with the annual period $T_h = 365$ sidereal days and amplitudes of the order of $0''.07$–$0''.08$. There is also an irregular drift at a rate of about $0''.005$ per year that has shifted the Earth's figure axis toward North America by approximately 15 m over the observation period of ~ 110 yrs. These components of the polar trajectory are also among the most important, and, together with the Chandler wobble, give rise to a rather chaotic pattern of irregular beating with a period about six years. Some doubts have been expressed in the literature about prospects for mathematical modeling and prediction of the polar trajectory [1–4].

We propose a comparatively simple mathematical model for the polar motion in terms of theoretical and celestial mechanics. The model is based mainly on a simplified analysis of the orbital motion of the Earth's center of mass (or the barycenter of the Earth and Moon) and estimates of the quasi-static deformation of the inertia tensor and diurnal tide in the Earth's mantle due to the gravitation forces of the Sun and Moon. Neglecting irregularities in the Earth's proper rotation, we can conduct an asymptotic analysis of the precession and nutation. For this, it is convenient to use the Euler–Liouville equations and kinematic equations in orbital coordinates, taking into account small additional perturbations in the inertia tensor and the dynamical asymmetry of the Earth's "frozen" body. A spectral analysis of the oscillations [1–3] demonstrates that the Chandler peak is fairly extended, while the annual peak is narrow and cannot be explained by seasonal phenomena.

The asymptotic methods of mechanics enable us to establish that the Earth's annual nutational oscillations relative to the center of mass are due to the solar gravitational moment, orbital motion of the rotating Earth, and diurnal tides in the mantle. Estimates of the gravitational moment provide evidence for this mechanism for perturbing the oscillations, since it requires very small relative perturbations in the inertia tensor ($\sim 5 \times 10^{-5}$), which are approximately a factor of 20 smaller than the "equatorial bulge" introduced as a basis for the Chandler period. The amplitudes of the semiannual and more rapid nutational oscillations are considerably smaller (by a factor of 10^2–10^3). The comparatively strong annual effect can be explained by the closeness of the frequencies of the natural and forced oscillations (Chandler and annual [2], see below).

Monthly nutation oscillations are virtually absent, and can be attributed to measurement errors. It is important that the lunar gravitational moment is stronger than the solar moment by a factor of two to three. However, due to the substantial difference between the natural and forced frequencies, the effect of the Moon gives rise to monthly amplitudes that are a factor of 15–20 smaller than the annual variations. Measurements with an interval of $\Delta\theta = 1$ day indicate that the amplitude of these oscillations is of the order of 10^{-3} angular seconds.

Similarly to the precession (free nutation), the drift of the Earth's figure axis (trend) can be explained by quasi-continuous changes in the inertia tensor. In particular, the inertia tensor components J_{pr} and J_{qr}, which have amplitudes of the order of $10^{-8}A^*$, lead to the observed trend of $0''.5$ along the y-axis (90° westward from Greenwich). Since the relative errors in the properties of the Earth's inertia (mass and inertia tensor) are considerably higher (10^{-6}–10^{-5}), we can use the measured trend to refine the Earth's inertia tensor. Changes in the trend itself can be explained by slow movements of the mantle on time scales of the order of 10^3 yrs, as well as by variations in the dynamical symmetry, which similarly gives rise to a slow modulation of the Chandler oscillations.

In order to construct a simplified mathematical model in a first approximation, we assume that small deformations of the Earth are mainly radial. Then, the equations governing the rotation relative to the center of mass take the form of the classical Euler equations with a variable inertia tensor J [2, 3]

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} = \mathbf{M}, \quad \boldsymbol{\omega} = (p, q, r)^T, \tag{10}$$
$$J = J^* + \delta J, \quad J^* = \text{const},$$

$$J^* = \mathrm{diag}(A^*, B^*, C^*), \delta J = \delta J(t), ||\delta J|| \ll ||J^*||.$$

Here, $\boldsymbol{\omega}$ is the angular velocity defined in some (reference) coordinate system fixed to the Earth and approximately coinciding with the main central inertia axes J^* of the "frozen" Earth, taking into account the equatorial bulge [2–4]. It is thought that small variations δJ of the inertia tensor can contain various harmonic components due to the solar and lunar diurnal tides, and possibly others as well (annual, semiannual, monthly, semidiurnal, etc). We assume that gravitational effects are the source of the main perturbing external force moments $\mathbf{M}$ responsible for the nutational oscillations. The possible presence of a term of the form $\dot{J}\boldsymbol{\omega}$ does not give rise to any improvement of the first-approximation model.

The kinematic Euler equations governing the orientation of the fixed axes with respect to the orbital coordinates take the form

$$\dot{\theta} = p\cos\varphi - q\sin\varphi - \omega_0(\nu)\sin\psi, \qquad (11)$$
$$\dot{\nu} = \omega_0(\nu) = \omega_*(1 + e\cos\nu)^2,$$
$$\dot{\psi} = \frac{p\sin\varphi + q\cos\varphi}{\sin\theta} - \omega_0(\nu)\,\mathrm{ctg}\,\theta\cos\psi, \quad e = 0.0167,$$
$$\dot{\varphi} = r - (p\sin\varphi + q\cos\varphi)\,\mathrm{ctg}\,\theta + \omega_0(\nu)\frac{\cos\psi}{\sin\theta}.$$

Here, $\nu(t)$ is the true anomaly, e the orbital eccentricity, and ω_* a constant determined by the gravitational and focal parameters. When solving Eqs. (10) and (11) for the polar motion, terms proportional to ω_0 in (11) turn out to be considerably larger (by approximately a factor of 300) than p and q, and become the main terms for $\dot{\theta}$ and $\dot{\varphi}$. This important property has not been noted earlier in the literature, and these terms were unjustifiably neglected.

The expressions for the components of the solar gravitational force moment take the form

$$M_q = 3\omega^2[(A^* + \delta A - (C^* + \delta C))\gamma_r\gamma_p \qquad (12)$$
$$+ \delta J_{pq}\gamma_r\gamma_q + \delta J_{pr}(\gamma_r^2 - \gamma_p^2) - \delta J_{rp}\gamma_p\gamma_q],$$
$$\omega = \omega_*(1 + e\cos\nu)^{3/2}, \quad \gamma_p = \sin\theta\sin\varphi,$$
$$\gamma_q = \sin\theta\cos\varphi, \quad \gamma_r = \cos\theta.$$

A cyclic permutation of the indices p, q, and r can be used to compute $M_{p,r}$. The annual component of the polar oscillations can be ascribed to terms containing products of the direction cosines $\gamma_p\gamma_r$ and $\gamma_q\gamma_r$. For their computation in a first approximation, we integrate Eqs. (11):

$$r = r^0, \quad \varphi \approx rt + \varphi^0, \quad \nu \approx \omega_* t + \nu^0, \qquad (13)$$
$$\cos\theta(\nu) = a(\theta^0, \psi^0)\cos\nu, \quad \theta(0) = \theta^0 = 66^\circ 33',$$
$$0.4 \le a \le 1, \quad 0 \le \psi^0 \le 2\pi,$$
$$\cos\theta\sin\theta = b(\theta^0, \psi^0)\cos\nu + d\cos 3\nu + \ldots,$$
$$0.4 \le b \le \frac{4}{3}\pi, \quad |d| \ll 1.$$

The second and higher harmonics in ν give rise to terms that are a factor of 10^2–10^3 smaller than the main terms, and can be neglected. The quantity B^*-A^* is also considerably smaller than C^*-A^* (by approximately a factor of 160). After averaging over the rapid phase φ, the estimates of the terms for p and q in (10) simplify the model to the simple analytical form

$$\dot{p} + N_p q = \kappa_q r^2 + 3b\omega_*^2\chi_p\cos\nu, \qquad (14)$$
$$N_{p,q} \approx N = \frac{2\pi}{T_1} \approx 0.84\omega_*,$$
$$\dot{q} - N_q p = -\kappa_p r^2 - 3b\omega_*^2\chi_q\cos\nu,$$
$$p(0) = p^0, \qquad q(0) = q^0.$$

Here, κ_p and κ_q are the mean $\delta J_{pr}/B^*$ and $\delta J_{qr}/A^*$, which can be slow functions. The quantities χ_p and χ_q result from averaging the coefficients of $\cos\nu$ (in the components of the solar gravitational force moment) over φ; as noted, they are due to diurnal tides. The lunar gravitational force moments are neglected due to their weak effect on the nutational oscillations.

The right-hand terms of (5) explicitly contain an annual harmonic effect that explains the nutational oscillations detected in the IERS observations. Although the sensitivity of the coefficients $\kappa_{p,q}$ exceeds that of $\chi_{p,q}$ by five orders of magnitude, any explicit annual (force-moment) effects with the required amplitude $M_h \sim 10^{20}$ kg m^2 s^{-2} associated with internal geophysical phenomena appear implausible from a mechanical point of view.

5. INTERPOLATING AND FORECASTING THE POLAR MOTION

The approximate solution of the Euler–Liouville and kinematic Euler equations yields expressions comprising the basis of the approximate theoretical model:

$$x(\theta) = c_x^0 + c_x^1\theta - a_x^c\cos 2\pi\Omega\theta + a_x^s\sin 2\pi\Omega\theta \qquad (15)$$
$$- \frac{\Omega}{1-\Omega^2}d_x^c\cos 2\pi\theta - \frac{1}{1-\Omega^2}d_x^s\sin 2\pi\theta$$
$$- b_x^c\theta\cos 2\pi\Omega\theta - b_x^s\theta\sin 2\pi\Omega\theta,$$
$$y(\theta) = c_y^0 + c_y^1\theta + a_y^c\cos 2\pi\Omega\theta + a_y^s\sin 2\pi\Omega\theta$$
$$- \frac{\Omega}{1-\Omega^2}d_y^c\cos 2\pi\theta + \frac{1}{1-\Omega^2}d_y^s\sin 2\pi\theta$$
$$- b_y^c\theta\cos 2\pi\Omega\theta + b_y^s\theta\sin 2\pi\Omega\theta,$$
$$\Omega = 0.840 - 0.845.$$

The secular terms in (15) are responsible for the possible slow evolution of the Chandler period, i.e.,

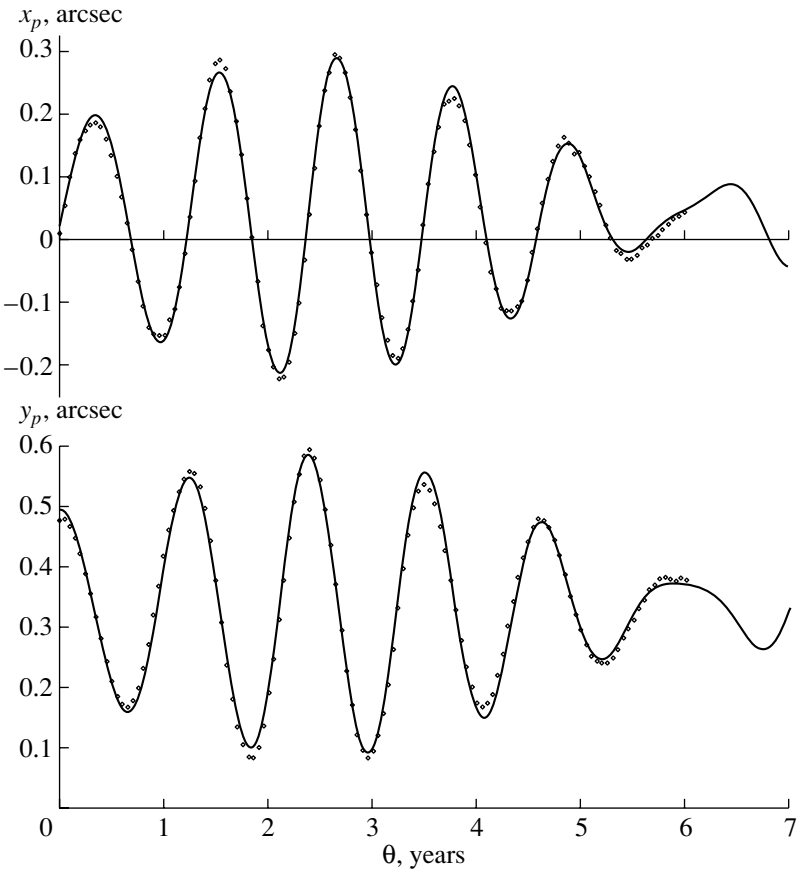

Fig. 3. Polar oscillations x_p and y_p for 1994–2000 (theoretical curves) based on observations in 1994–1999 (dots); the forecast for 2000.

Ω. The argument θ denotes the time in years. The parameter Ω describes the ratio of the annual period T_h to the Chandler period T_1. The constants $a^{c,s}_{x,y}$, $c^{0,1}_{x,y}$, $d^{c,s}_{x,y}$, and $b^{c,s}_{x,y}$ must be calculated in accordance with the model (15) using the least squares method [12]. The eight parameters for the functions x and y can be determined independently. Note that, theoretically, the constructed first-approximation model obeys the laws

$$a^c_x = a^s_y, \quad a^s_x = a^c_y, \quad d^c_x = d^s_y, \qquad (16)$$

$$d^s_x = d^c_y, \quad b^c_x = b^s_y, \quad b^s_x = b^c_y.$$

The vector (x, y) contains a total of ten independent parameters. However, it is more convenient to determine independently these eight parameters (sixteen numbers) from measurements, then compare the corresponding coefficients in accordance with (16) as a verification of the model (15). Let us express the variables $x(\theta)$ and $y(\theta)$ in the compact form

$$x(\theta) = (\xi, f(\theta)), \quad y(\theta) = (\eta, f(\theta)), \quad \xi = (\xi_1, \xi_2, \ldots, \xi_8)^T, \quad \eta = (\eta_1, \eta_2, \ldots, \eta_8)^T, \qquad (17)$$
$$f(\theta) = (1, \theta, \cos 2\pi\Omega\theta, \sin 2\pi\Omega\theta, \cos 2\pi\theta, \sin 2\pi\theta, \theta\cos 2\pi\Omega\theta, \theta\sin 2\pi\Omega\theta)^T.$$

There is an unambiguous correspondence between the components of the vectors ξ, η (17), and the model coefficients (15). The quantities x and y are known at discrete times θ_i [1]

$$x_i = x(\theta_i), \quad y_i = y(\theta_i), \quad \theta_i = \frac{i}{20}, \qquad (18)$$
$$i = 0, 1, 2, \ldots, N, \quad N = 20\Theta,$$

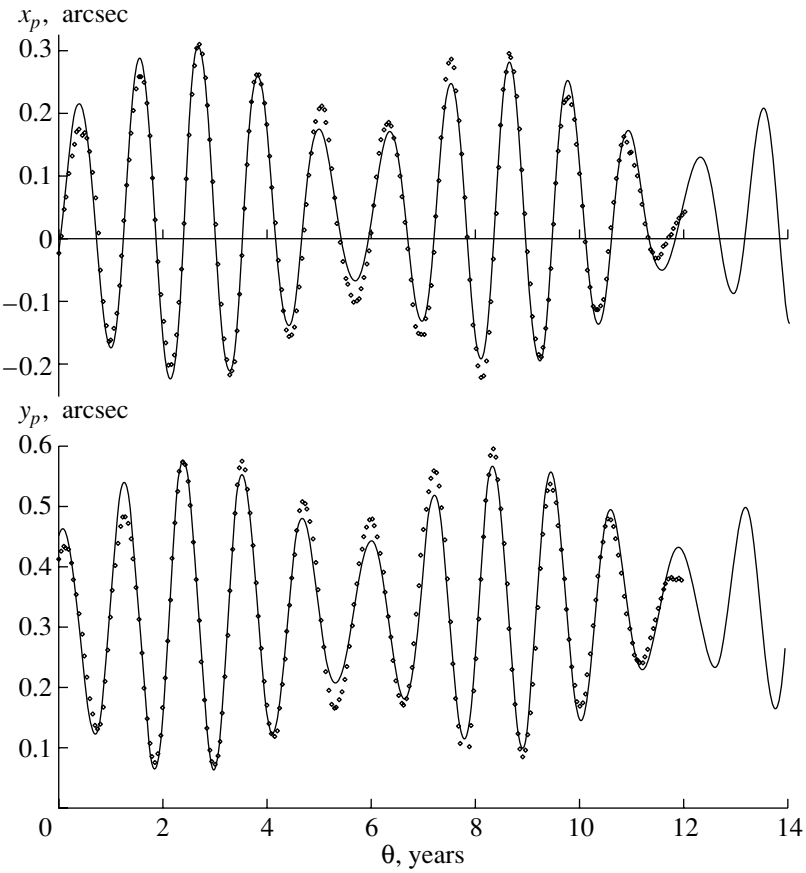

Fig. 4. Polar oscillations x_p and y_p for 1988–2001 (theoretical curves) based on observations in 1988–1999 (dots); the forecast for 2000–2001.

where the quantity Θ corresponds to the number of years. The best-fit coefficients ξ^* and η^* can be calculated and interpolated using the modified theoretical model (17) and observational data. This provides the ability to forecast beyond the interval used. We have considered intervals of 6, 12, and 20 yrs.

We will present the results of our interpolation and forecasting for two observational intervals: a comparatively short interval with $\Theta = 6$ yrs (121 points) and a longer interval with $\Theta = 12$ yrs (241 points); see (18). Expressions (17) for the first interval, corresponding to the observational data of 1994–1999 [1], take the form

$$\begin{aligned} x^*(\theta) &= -0.03221 - 0.00113\theta \qquad (19)\\ &\quad - 0.01885\cos 2\pi\Omega\theta - 0.22648\sin 2\pi\Omega\theta\\ &\quad + 0.030060\cos 2\pi\theta + 0.07306\sin 2\pi\theta\\ &\quad + 0.0156\theta\cos 2\pi\Omega\theta + 0.01952\theta\sin 2\pi\Omega\theta,\\ y^*(\theta) &= 0.33763 - 0.000870\theta\\ &\quad + 0.22566\cos 2\pi\Omega\theta - 0.24660\sin 2\pi\Omega\theta\\ &\quad - 0.06904\cos 2\pi\theta - 0.02544\sin 2\pi\theta\\ &\quad - 0.191660\theta\cos 2\pi\Omega\theta + 0.0033194\theta\sin 2\pi\Omega\theta. \end{aligned}$$

Figure 3 shows the data (121 points) and theoretical curves $x^*(\theta)$, $y^*(\theta)$ for the polar trajectory over the six years indicated, together with the forecast for 2000. The coefficients of (19) satisfy the conditions (16) quite convincingly. The graphical data for 2000 presented by the IERS on their internet web site agree well with the curves in Fig. 3.

For the longer interval with $\Theta = 12$ yrs [1], expressions (17) for 1988–1999 take the form

$$\begin{aligned} x(\theta) &= -0.03935 + 0.000075\theta \qquad (20)\\ &\quad + 0.03047\cos 2\pi\Omega\theta - 0.20456\sin 2\pi\Omega\theta\\ &\quad + 0.030469\cos 2\pi\theta + 0.06669\sin 2\pi\theta\\ &\quad - 0.000496\theta\cos 2\pi\Omega\theta + 0.004203\theta\sin 2\pi\Omega\theta,\\ y(\theta) &= 0.30798 + 0.002772\theta\\ &\quad + 0.19990\cos 2\pi\Omega\theta + 0.02753\sin 2\pi\Omega\theta\\ &\quad - 0.06302\cos 2\pi\theta + 0.02766\sin 2\pi\theta\\ &\quad - 0.003854\theta\cos 2\pi\Omega\theta + 0.000133\theta\sin 2\pi\Omega\theta. \end{aligned}$$

Figure 4 presents the theoretical curves (20) and data (241 points). They testify to the good accuracy of the interpolation and a good agreement with conditions (16). The forecast curves are also presented,

and are in satisfactory agreement with the processed digital IERS data.

6. CONCLUSIONS

The proposed model can also be applied to longer (20-year) and shorter (2-year) interpolation intervals. The IERS data [1] and least-squares results indicate that the constructed mathematical model we have derived using a celestial-mechanics approach can adequately describe the complex oscillations of the Earth's pole over intervals of 20–30 yrs.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-02-17250) and the State Science and Technology Program in Astronomy (project 1.8.1.2).

REFERENCES

1. *IERS Annual Reports, 1990 July 1991 to 2000 July 2001* (Central Bureau of the IERS, Observatoire de Paris, 2001).
2. W. Munk and G. Macdonald, *The Rotation of the Earth* (Cambridge Univ. Press, Cambridge, 1960; Mir, Moscow, 1964).
3. H. Moritz and I. I. Mueller, *Earth Rotation: Theory and Observation* (Ungar, New York, 1987; Naukova Dumka, Kiev, 1992).
4. Yu. N. Avsyuk, *Tidal Forces and Natural Processes* [in Russian] (Ob"edin. Inst. Fiz. Zemli Ross. Akad. Nauk, Moscow, 1996).
5. L. I. Konkina, Yu. G. Markov, A. M. Mikisha, and L. V. Rykhlova, Dokl. Akad. Nauk **375** (5), 619 (2000) [Dokl. Phys. **45**, 672 (2000)].
6. L. D. Akulenko, S. A. Kumakshev, Yu. G. Markov, and L. V. Rykhlova, Astron. Zh. **79**, 81 (2002) [Astron. Rep. **46**, 74 (2002)].
7. A. A. Il'yushin, *Mechanics of Continuous Media* [in Russian] (Mosk. Gos. Univ., Moscow, 1990).
8. F. L. Chernous'ko, Izv. Akad. Nauk SSSR, Mekh. Tverd. Tela, No. 1, 22 (1980).
9. N. N. Bogolyubov and Yu. A. Mitropol'skii, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Nauka, Moscow, 1974; Gordon and Breach, New York, 1962).
10. V. V. Beletskiĭ, *Motion of a Satellite Relative to the Center of Mass in a Gravitational Field* [in Russian] (Mosk. Gos. Univ., Moscow, 1975).
11. L. D. Akulenko, Yu. G. Markov, and L. V. Rykhlova, Dokl. Akad. Nauk **377** (5), 618 (2001) [Dokl. Phys. **46**, 261 (2001)].
12. Yu. V. Linnik, *The Least-Squares Method and the Foundations of Mathematical-Statistical Theory of Observation Processing* [in Russian] (Fizmatgiz, Moscow, 1962).

Translated by V. Badin