Astronomy Reports, Vol. 45, No. 11, 2001, pp. 841–853. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 963–976.
Original Russian Text Copyright © 2001 by Chernin, Kravtsova, Zasov, Arkhipova.

Galaxies with Rows

A. D. Chernin, A. S. Kravtsova, A. V. Zasov, and V. P. Arkhipova

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

Received March 16, 2001

Abstract—The results of a search for galaxies with straight structural elements, usually spiral-arm rows ("rows" in the terminology of Vorontsov-Vel'yaminov), are reported. The list of galaxies that possess (or probably possess) such rows includes about 200 objects, of which about 70% are brighter than 14^m. On the whole, galaxies with rows make up 6–8% of all spiral galaxies with well-developed spiral patterns. Most galaxies with rows are gas-rich Sbc–Scd spirals. The fraction of interacting galaxies among them is appreciably higher than among galaxies without rows. Earlier conclusions that, as a rule, the lengths of rows are similar to their galactocentric distances and that the angles between adjacent rows are concentrated near 120° are confirmed. It is concluded that the rows must be transient hydrodynamic structures that develop in normal galaxies. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The long, straight features found in some galaxies, which usually appear as straight spiral-arm rows and persist in spite of differential rotation of the galaxy disks, pose an intriguing problem. These features, first described by Vorontsov-Vel'yaminov (to whom we owe the term "rows") have long escaped the attention of researchers. They have attracted interest in recent years owing to a series of papers by Chernin *et al.*, who described straight structural rows in the nearby galaxies M101 and M51 (see [1] and references therein) and suggested a qualitative interpretation of these features. They linked the formation of such rows to the fact that powerful shock fronts forming along spiral arms have a tendency to straighten and become flat, so that sites of shock-triggered star formation can form linear structures. This raises the question as to what factors determine the conditions for the development of rows. To answer this question, we must analyze data for as many galaxies as possible whose spiral patterns exhibit these features.

Chernin *et al.* [1] described 15 spiral galaxies with rows, which they found primarily by inspecting the NASA Atlas of Galaxies [2]. It was shown that, as a rule, the row lengths are similar to their galactocentric distances and that these features are found in both two-armed (M51, NGC 4303) and multi-armed (M101, NGC 1232) galaxies, which appear to be quite normal as far as their integrated parameters are concerned. More extensive samples are required to estimate the rate of occurrence of galaxies with rows.

In the current paper, we study a sample of several thousand spiral galaxies whose Palomar sky atlas images were reproduced by Vorontsov-Vel'yaminov and analyze the general properties of these objects.

2. THE SAMPLE OF GALAXIES WITH STRAIGHT ROWS

To identify galaxies with rows, we inspected about 7000 photographs stored in the collection of the Sternberg Astronomical Institute, which consists of enlarged reproductions of nonoverexposed blue images of (mostly spiral) galaxies that are, with few exceptions, brighter than 15^m, adopted from the Palomar sky atlas. The collection contains about one-fourth of all galaxies north of $\delta = -45°$ catalogued in the five volumes of the Morphological Catalog of Galaxies of Vorontsov-Vel'yaminov [3]. It lacks, however, nearby galaxies with large angular diameters. Excluding nonspiral galaxies and galaxies with poorly defined morphological structure leaves us with 4200 objects, which make up the basic sample used below.

An inspection of the photographs enabled identification of about 200 *bona fide* galaxies possessing (or likely possessing) straight rows, which make up about 5% of all the galaxies considered. Figure 1 shows several galaxies with rows by way of example. In most cases, the rows appear as straight stretches of spiral arms (or, more rarely, of ring-shaped features), where they form regular hexagonal structures. However, the relationship between these straight rows and a regular pattern is by no means evident. Linear features such as bars are not considered to be "rows."

The mean magnitude of galaxies with rows is approximately the same as that of galaxies without them

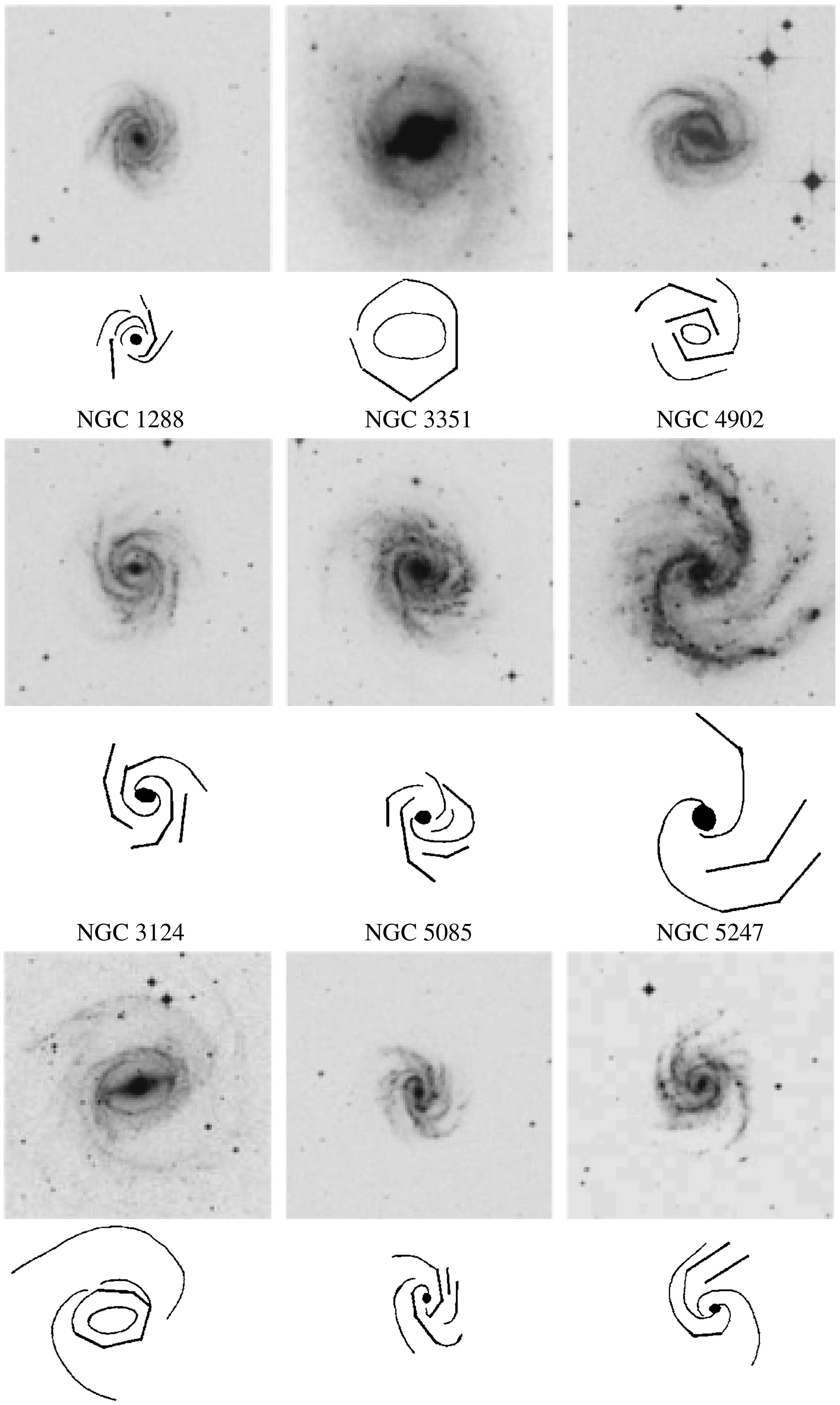

Fig. 1. The most typical examples of galaxies with rows.

($\sim 14^m$), although the scatter of the magnitude distribution is large in both cases. However, identifying galaxies with rows is undoubtedly more uncertain among faint systems. Some galaxies with rows could be overlooked due to insufficiently good seeing, especially in the case of faint galaxies. On the other hand, if identification is based on a visual inspection of photographs with insufficient resolution, this same factor could yield false detections of rows, which then contaminate the final sample. The subsample of galaxies brighter than 14^m contains about one thousand objects. Of these, we found 157 galaxies (about 15%) exhibiting rows. The sample size decreases substantially as we consider brighter galaxies. For example, the NASA Atlas of Galaxies [2] gives photos of about 200 nearby galaxies with well-defined spiral arms, of which about 15 objects [1] exhibit well-defined rows, corresponding to about 8% of the entire sample, although the statistical error of this fraction is rather high.

The table provides information on the entire sample of galaxies with rows. Its columns give (1) the name of the galaxy according to the PGC and NGC catalogs; (2) its morphological type; (3) its integrated magnitude B_0 corrected for Galactic extinction and the inclination of the galaxy to the line of sight; (4) its heliocentric radial velocity V_0; (5) its absolute magnitude M_B ($H_0 = 75$ km/s Mpc); (6) the HI mass-to-light ratio, $M(\mathrm{HI})/L_B$ (in solar units), inferred from the HI index of the RC3 catalog; (7) the number of rows detected in the galaxy; and (8) its VV number in the catalog of interacting galaxies [5]. We adopted the integrated parameters of the galaxies from the LEDA electronic database and RC3 reference catalog of galaxies [4]. Figures 2a and 2b show distributions of the morphological types of galaxies with and without rows. It is evident that the fraction of objects with rows is especially high among late Sbc–Scd galaxies. Figures 3a and 3b compare the absolute magnitudes of the two groups of galaxies. Galaxies with rows have absolute magnitudes ranging from -17^m to $-22\overset{m}{.}5$ with the mean and median over the entire sample equal to $M_B = -20\overset{m}{.}4$ and $-20\overset{m}{.}7$, respectively. Galaxies without rows exhibit the same general type of luminosity distribution.

The number of rows in a galaxy ranges from one to nine (Fig. 4), with the mean over the entire sample being close to three rows per galaxy. We found no correlation between the number of rows and the absolute magnitude of the galaxy.

The linear sizes L of rows vary over a wide range and can exceed 20 kpc in some cases (M101, Fig. 5). The mean and median row lengths are 4.9 and 4.0 kpc, respectively. We found the length of a row to be well correlated with the galactocentric distance d of the farthest tip of the row—the further from the galactic center, the longer the row. Figure 6 illustrates this correlation. The lengths and galactocentric distances d of the rows are expressed as fractions of the optical radii of their galaxies $D_{25}/2$. The solid line in Fig. 6 shows the relation $L = d$; the dotted lines, two regression fits to the observational data: $d = 0.87L + 0.11$ and $d = 1.13L - 0.25$. The standard deviation is 0.11.

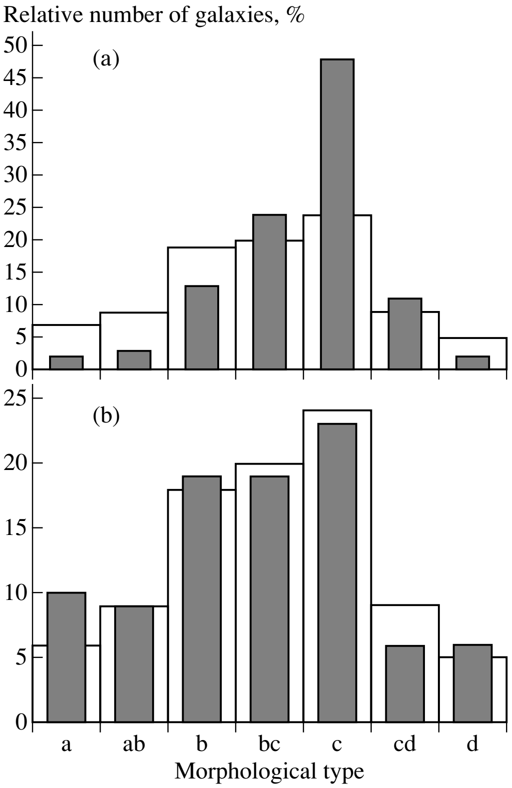

Fig. 2. Histograms of morphological types of galaxies (a) with and (b) without rows. Shaded entries correspond to interacting galaxies.

The angle between adjacent rows is usually close to 120°. Figure 7 shows the distribution of these angles. The median and mean are 122° and 125°, respectively.

The galactic rotation curves, which are available only for a small fraction of the galaxies in our sample, exhibit no apparent anomalies. For three galaxies with rows (M101, M51, and NGC 3631), the gas velocity fields were studied throughout the disk in great detail using both optical and radio methods. In each of these, the gas motions are dominated by regular differential rotation, although noncircular velocities (both local and ordered, associated with the spiral structure) are also present. This pattern is fairly typical of spiral galaxies. On the other hand, the galaxy with well-developed spiral arms NGC 157, in which the rotation is only slightly differential (the linear rotation velocity increases monotonically with galactocentric distance throughout the entire region of the spiral arms) and the wave nature of the spirals is indicated directly through analysis of the velocity field [6], fails to show even the slightest sign of rows.

Galaxies possessing (or possibly possessing) rows

PGC/NGC	Type	B_0	V_0, km/s	M_B	M_{HI}/L_B	Number of rows	VV	Comments
PGC								
120	SBc	13.6	4389	−21.1	0.21	5	254	1
2151	Scd	14.6	4929	−20.0		2	548	
2720	Sc	13.7	11 236	−22.8		5	527	
3011	Sb	15.3	5855	−19.9	0.13	2	554	
3846	SB(rs)bc	13.2	5639	−20.8		2		
3853	SAB(rs)d	12.1	1094	−18.6	0.7	2		
4189	SB(s)c	12.8	5708	−21.6	0.35	6		2
4227	SB(rs)c	12.9	6044	−21.6		4		
4672	SBc	13.6	5168	−21.2	0.39	3	*	
4915	SBc	14.0	5869	−21.4		2	951	
5227	SAB(rs)c	14.2	10 848	−21.6		5		
5362	SAB(r)bc	13.1	5538	−21.2		3		
5715	Sbc	15.4	12 435	−21.8		2	301	3
7359	Sbc	14.5	8054	−21.3	0.21	2	12	
7967	SBbc	13.6	5275	−20.7	0.29	3		
8360	SAB(s)cd	14.0	3256	−19.1	0.82	2		
8961	Sbc	13.4	7563	−22.4	0.3	2	323	
8974	SAB(s)cd	11.7	2289	−20.7	0.44	2		
9899	SAB(rs)cd	11.2	903	−19.5	0.63	4		
10 587	Scd	12.5	4044	−21.2		4		
11 691	SAB(r)c	13.2	4747	−20.7		2		
12 184	SA(s)c	14.9	4223	−19.6	0.72	4	1071	
15 429	SAB(rs)c	13.4	3541	−19.9	0.7	4		
15 821	SAB(rs)c	11.1	716	−18.6	0.38	3		
15 831	SAB(rs)c	13.1	4531	−20.7	0.27	2		
15 950	SBb	13.5	3487	−20.4		4	1110	
18 047	SA(rs)c	13.9	3121	−19.1	0.44	3		
19 682	SBbc	14.3	4645	−20.9		3	1168	
19 789	SA(s)cd	14.1	1442	−17.6	1.02	3		
20 911	Sb	13.3	4031	−21.0		1	528	
23 321	Sbc	14.2	5696	−21.1	0.54	1	1219	
23 935	SBc	14.6	11 159	−22.1	0.31	2	413	
23 936	SA(rs)c	13.5	5488	−20.8	0.24	3		
26 132	Sc	15.6	11 634	−21.1		1	155	
27 546	SBc	15.2	6825	−20.1		3	83	
28 888	SAB(rs)c	13.1	5144	−21.1	0.46	6		

(Contd.)

PGC/NGC	Type	B_0	V_0, km/s	M_B	$M_{\rm HI}/L_B$	Number of rows	VV	Comments
30 814	SAB(rs)c	12.7	3136	−20.2		3		
30 891	Sc	15.5	11 535			2	*	
31 551	SB(r)b	11.9	3706	−21.5		5		1, 2
32 495	SA(rs)c	13.0	3399	−20.4		3		
32 620	SBb	14.6	10 397	−21.6		2	233	
32 846	SA(s)c	14.4	8355	−20.8		2	466	
34 006	SAB(rs)c	13.7	7781	−21.4	0.43	1		
35 006	SA(rs)bc	12.5	5040	−21.6		2		
35 570	SBd	15.3				2	444	
37 400	SB(rs)bc	13.7	6972	−21.1	0.38	3		
37 444	SB(s)cd	13.2	1892	−18.8	0.52	2		
38 024	SA(rs)bc	13.6	6133	−21.0	0.54	4		
39 014	SBc	14.5	7118	−21.1		1	128	
39 483	SBc	13.8	733	−16.7		2	431	
41 011		14.7	7148			4	1541	
41 291	SB(r)bc	13.7	5462	−20.6		2		
44 032	SB(rs)bc	14.1	10 389	−21.7		3		
44 213	Sc	15.3	13 413	−21.8		2	1580	
44 810	SBbc	16.0				2	418	
45 921	SBb	14.8	10 571	−21.6		3	451	
46 041	SBc	14.9	2326	−18.4		2	438	2
46 114	Sab	14.4	9243	−21.6		1	250	
46 770		15.6	8834	−20.1		1	235	3
46 878	SB(s)c	12.4	2661	−20.3	0.23	4		
47 198	SB(rs)bc	12.9	3990	−20.8		2		
47 808	SBc	15.0	5016	−20.1	0.7	2	4	
47 867	Sbc	14.4	4962	−20.1	0.98	5	1635	
50 120	SB(s)c	12.6	3322	−20.7		3		
51 169	SA(rs)c	12.8	6711	−21.9		4		
51 456	SAB(s)cd	12.3	2683	−20.5		2		
52 853	SB(s)cd	13.8	2046	−18.4		2		
53 134	SBcd	13.2	1951	−17.6	0.92	2	815	3
54 776	SAB(s)c	12.2	2859	−20.7		3		
54 849	SB(r)bc	11.2	1480	−20.4	0.21	3		
55 750	SA(rs)c	14.1	4526	−19.8	0.32	3		
56 014	SAB(r)b	14.9	9310	−20.6	0.29	2		

(Contd.)

PGC/NGC	Type	B_0	V_0, km/s	M_B	$M_{\rm HI}/L_B$	Number of rows	VV	Comments
59 133	SB(rs)bc	13.3	7180	−21.6		3		
61 583	Sbc	15.3	4624	−19.7		4	1866	
66 076	Sbc	12.6	4881	−21.5	0.16	4		
66 333	Sb	14.6	8269	−21.9		4	748	1
68 344	SB(s)c	13.0	4304	−20.8		2		
68 543	SA(s)c	15.2	6412	−19.5	0.45	4		
69 439	SAB(s)bc	12.8	4695	−21.3	0.29	2		
70 020	SBb	13.7	7255	−21.2		2		
71 314	SB(r)c	13.3	6926	−21.5		2		
72 444	SB(r)bc	11.9	2961	−21.0	0.39	4		
NGC								
70	Sbc	14.2	7167	−21.3	0.12	3	166	
91	SBbc	14.5	5341	−20.7	0.49	2	881	
180	SB(rs)bc	13.0	5280	−21.3	0.44	3		
210	SAB(s)b	11.2	1634	−20.4	0.52	2		
266	SB(rs)ab	12.2	4657	−21.9	0.14	3		
309	SAB(r)c	12.2	5662	−22.2	0.23	3		2
341	SBb	13.8	4499	−20.4		2	361	
378	SB(r)c	13.4	9603	−22.1		3		
514	SAB(rs)c	11.9	2468	−20.7	0.3	2		
613	SBbc	11.0	1485	−20.7	0.12	4	824	
628	SA(s)c	9.7	656	−20.2	0.56	3		
642	SBc	13.6	5885	−21.5	3	419		
858	SB(rs)c	13.8	12 356	−22.3		2		
864	SAB(rs)c	11.1	1560	−20.5	0.5	4		
918	SAB(rs)c	12.4	1509	−19.2	0.35	4		
927	SB(r)c	13.9	8260	−21.3	0.28	5		
947	SBc	13.3	4938	−21.5		1	1028	
1042	SAB(rs)cd	11.2	1373	−20.0	0.26	2		
1097	(R)SB(r)b	9.6	1273	−21.2	0.25	1		
1144	S	13.8	8696	−22.2		2	331	
1179	SB(rs)d	12.3	1778	−19.4	0.61	2		
1232	SABc	10.7	1681	−21.3	0.29	8	*	
1241	SBb	13.2	4026	−21.0		1	334	
1255	SAB(rs)bc	11.3	1697	−20.5		2		
1288	SBc	12.6	4538	−21.3	0.12	3		2

(Contd.)

PGC/NGC	Type	B_0	V_0, km/s	M_B	M_{HI}/L_B	Number of rows	VV	Comments
1313	SBc	9.7	461	−19.3	0.47	2	436	
1347	SBc	14.1	1760	−18.0		2	23	
1642	SA(rs)c	12.9	4633	−21.1		4		
1961	SBbc	11.8	3930	−22.9	0.33	5	1110	
2207	SBbc	12.1	2747	−21.6	0.2	5	1155	2
2222	Sb	14.4	6983	−21.7		2	315	
2223	SBbc	12.5	2721	−21.1	0.23	5		
2276	SBc	12.0	2405	−21.3	0.18	3	1189	
2326	SB(rs)b	12.7	5985	−21.9		3		
2342	Sc	13.1	5276	−22.2	0.17	4	1176	
2388	S? HII	14.2	4063	−19.5	0.61	3		
2445	Irr	13.9	3994	−20.4	0.62	3	117	
2523	SB(r)bc	12.0	3452	−21.5		2		
2535	Sc	13.3	4097	−21.3	0.41	2	9	
2565	SBbc	13.4	3589	−21.0	0.34	5	*	1
2595	SAB(rs)c	12.7	4330	−21.2	0.24	3		
2614	SA(r)c	13.2	3457	−20.3	0.27	6		2
2623	Sab	14.0	5532	−21.3		2	79	3
2633	SBab	12.9	2162	−20.3	0.27	2	519	
2642	SB(r)bc	12.9	4334	−20.9	0.19	3		
2750	Sc	12.8	2672	−20.4	0.31	1	541	
2763	SB(r)cd	12.4	1893	−19.7		3		
2817	SAB(rs)c	13.3	3534	−20.0		2		
2857	SA(s)c	12.9	4887	−21.6	0.28	4	*	
2889	SAB(rs)c	12.0	3365	−21.2	0.1	2		
2935	SAB(s)b	11.8	2277	−20.6		2		
2936	S	13.9	6989	−21.4		4	316	
2942	SAc	13.4	4425	−20.9	0.35	4		
2997	SAc	10.1	1087	−21.3	0.2	3		
3027	SBc	12.2	1059	−20.0	0.8	2	358	
3052	SAB(rs)c	12.1	3770	−21.3	0.18	4		
3124	SB(r)c	12.2	3561	−21.1	0.38	9		2
3184	SAB(rs)cd	10.1	593	−19.8	0.26	3		
3346	SB(rs)cd	12.4	1259	−18.8	0.22	4		
3351	SB()b	10.0	778	−20.2	0.13	3		1
3423	SA(s)cd	11.2	1010	−19.4	0.26	5		

(Contd.)

PGC/NGC	Type	B_0	V_0, km/s	M_B	M_{HI}/L_B	Number of rows	VV	Comments
3434	SA(r)b	14.2	3633	19.2	0.34	2		
3456	SB(rs)bc	12.7	4192	−21.0		2		
3486	SAB(r)c	10.6	681	−19.5	0.47	3		
3504	(R)SAB(s)ab	11.5	1539	−20.2	0.03	2		
3627	SBb	9.6	727	−21.1	0.03	3		2
3631	Sac	11.1	1158	−20.6	0.23	3	*	
3646	Sc	11.8	4249	−22.7	0.18	3	*	1
3672	SA(s)c	11.3	1862	−20.6	0.41	2		
3673	SB(r)b	11.8	1940	−20.1	0.14	3		
3786	Sba	13.4	2718	−20.0	0.36	2	228	
3861	Sb	13.5	5083	−21.3	0.2	3	1469	
3888	SBc	12.7	2407	−20.5	0.28	2	455	
3891	Sbc	12.9	6361	−21.8	0.29	4		
3897	SBbc	13.4	6411	−21.4	0.5	3		
3905	SB(rs)c	12.9	5768	−21.5	0.19	2		
3930	SBc	12.5	916	−18.7	0.82	4	1478	
3938	SAc	11.1	809	−19.8	0.32	5		
3963	SAB(rs)bc	12.2	3186	−21.1	0.27	2		
3968	SAB(rs)bc	12.8	6392	−21.8	0.37	4		
4137	SBc	14.9	9300	−21.2		1	454	
4152	SBc	12.7	2165	−20.1	0.32	2	1508	
4303	SABc	10.2	1575	−21.8	0.2	3		1
4321	SABbc	10.1	1579	−22.1	0.1	2		
4535	SABc	10.7	1958	−21.9	0.2	5		
4548	SB(rs)b	11.0	486	−21.0	0.04	3		
4561	SBc	13.0	1405	−19.1	0.53	2	571	
4676	SBa	14.1	6614	−21.3		1	224	3
4779	SB(rs)bc	12.8	2829	−20.1	0.21	3		
4902	SB(r)b	11.4	2625	−21.2	0.14	5		2
4939	SA(s)bc	11.1	3109	−22.0	0.29	1		
5000	SBbc	13.9	5608	−20.9	0.23	3	460	
5085	SA(s)c	12.8	1956	−19.1	0.44	6		2
5149	SBbc	13.1	5658	−21.3		2		
5161	SAc	12.0	2388	−21.5	0.43	3		
5194	Sc	9.0	461	−20.6	0.11	9	1	
5218	SBb	13.1	2878	−20.6		1	33	

(Contd.)

PGC/NGC	Type	B_0	V_0, km/s	M_B	M_{HI}/L_B	Number of rows	VV	Comments
5230	SBc	13.2	6852	−21.7	0.21	5		
5247	SBbc	10.7	1357	−20.5	0.14	5		2
5293	SA(r)c	13.4	5785	−21.1	0.26	4		
5339	SB(rs)a	12.5	2735	−20.3		3		
5350	SB(r)b	11.9	2312	−20.8	0.3	4		
5427	SBc	12.1	2619	−21.2	0.66	3	21	2
5457	Sbc	8.3	240	−20.8	0.49	8		
5468	SAB(rs)cd	12.4	2842	−20.5	0.47	4		
5579	Sc	14.3	3602	−19.8	0.45	3	142	
5595	SBc	12.7	2697	−20.9	0.29	4	530	
5679-	Sc	14.4	7483	−21.3		3	458	
5679H	Sb	14.4	7500	−21.4		2	458	2
5829	Sc	13.9	5697	−21.0	0.39	4	7	2
6118	SA(s)cd	11.1	1571	−20.6	0.2	5		
6412	SBc	12.3	1323	−19.8	0.25	3	*	
6872	SBb	12.6	4799	−22.8		2	*	2
6946	SABcd	9.7	37	−20.6	0.17	6	1892	
7137	SABc	13.1	1683	−19.5	0.16	3		
7298	SA(s)c	13.9	5035	−20.2	0.6	2		
7479	SB(s)c	11.3	2378	−21.3	0.18	2		
7610	SBc	13.5	3554	−20.5	1.08	2	*	
7685	SAB(s)c	13.5	5642	−20.9	0.26	4		
7741	SB(s)cd	11.5	755	−18.9	0.32	2		
7757	SA(rs)c	12.7	2955	−20.3	0.36	3	407	
7805	Sbc	14.3	4768	−20.3		2	226	

Asterisks in the VV column indicate interacting galaxies with no VV numbers. Comments: 1—ring galaxies with straight-line ring rows (hexagons); 2—most typical examples of galaxies with rows. 3—Galaxies with straight tails, purportedly of tidal origin.

About ten galaxies with rows that are not classed as interacting systems do nonetheless possess close companions. However, there are usually no radial velocity measurements for these companions, making it impossible, with few exceptions, to say anything certain about their physical association with the main galaxy.

3. GALAXIES WITH ROWS AMONG INTERACTING AND NONINTERACTING GALAXIES

The new version of the catalog of interacting galaxies of Vorontsov-Vel'yaminov *et al.* [5] lists a total of 2014 systems (about 4000 galaxies) detected in the Palomar Sky Atlas during the compilation of a five-volume morphological catalogue of galaxies. The

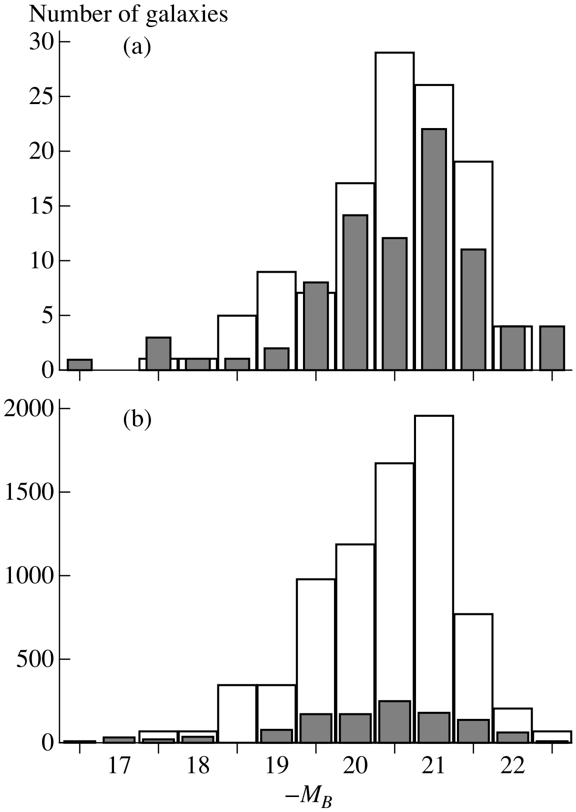

Fig. 3. Histograms of absolute magnitudes of galaxies (a) with and (b) without rows. Shaded entries correspond to interacting galaxies.

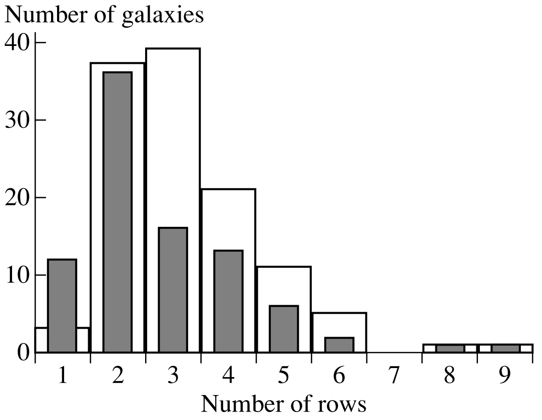

Fig. 4. Histograms of the number of rows in galaxies. Shaded entries correspond to interacting galaxies.

available images for 1200 of these galaxies enable determination of the shape of their spiral arms. We found rows in 85 of the 1200 galaxies considered ($7 \pm 0.7\%$ of all galaxies). These statistics include a number of galaxies with straight tidal tails. The fraction of galaxies with rows among noninteracting galaxies of our sample is substantially lower, $\sqrt{119}/3000 \approx 4 \pm 0.4\%$ of the entire sample. These proportions can be expressed in a different way. The fractions of interacting galaxies among galaxies without and with rows are $28 \pm 1\%$ and $43 \pm 5\%$, respectively. It follows that interacting galaxies occur almost twice as often among galaxies with rows. This is not due to differences in the luminosities of the two types of objects: interacting and noninteracting galaxies have similar absolute-magnitude distributions (Figs. 3a, 3b).

As a rule, rows are found in only one component of an interacting galaxy. The only system where rows were detected in both components is VV 458 = MCG 1-37-34/35.

In addition to galaxies with rows that are straight features of their internal structure, a few interacting galaxies have straight features that Vorontsov-Vel'yaminov referred to as bridges or tails. These features are the results of tidal interaction, and their nature probably differs from that of rows. However, it is not always possible to unambiguously distinguish these structural features, since tidal tails are often extensions of spiral arms.

Figure 4 shows the distribution of the number of rows in interacting galaxies. The mean number of rows in an interacting galaxy is 2.8, which is lower than that for noninteracting galaxies. The distribution of the number of rows in interacting galaxies peaks at two rows per galaxy, whereas the mean value for noninteracting galaxies is closer to three rows per galaxy (Fig. 4). In addition, the fraction of objects with only one row is appreciably higher among interacting galaxies. Of course, a role is played here by the contribution of straight tidal-induced features. However, galaxies with many (8–9) rows are roughly equally rare among interacting and noninteracting galaxies.

Figure 5 illustrates the distribution of row lengths in components of interacting and noninteracting galaxies. The range of row lengths is appreciably broader for interacting galaxies than for single objects. Rows in the external part of a spiral pattern are encountered much more frequently in interacting galaxies (mainly due to long bridges and tails); however, the conclusion that interacting galaxies have longer rows remains valid even if we exclude galaxies with tidal features: the fraction of short (2–4 kpc) rows in interacting galaxies is lower than the corresponding fraction in noninteracting galaxies. The close correlation between the size of rows and their galactocentric distance leads us to conclude that, in noninteracting galaxies, rows occur more often in outer regions, where the manifestations of interaction events are stronger.

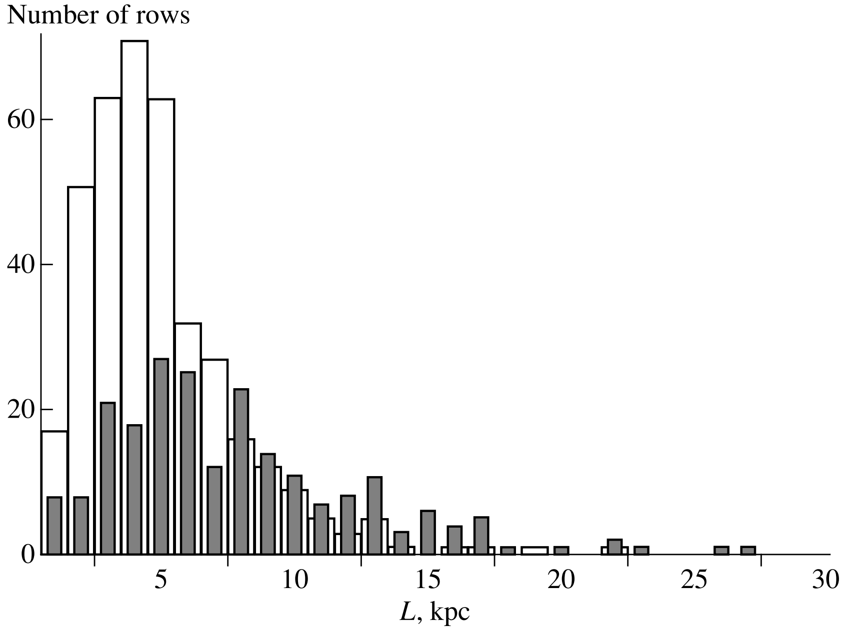

Fig. 5. Histogram of linear sizes of rows. Shaded entries correspond to interacting galaxies.

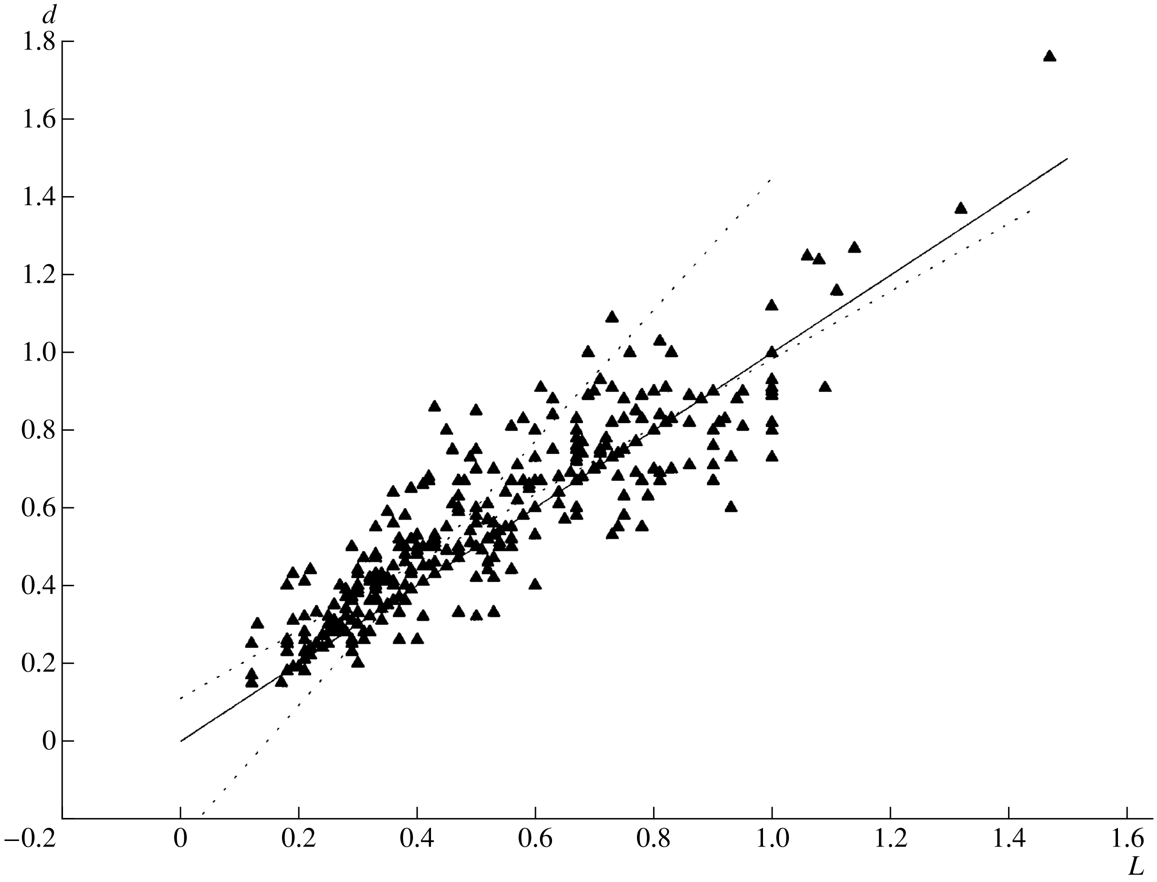

Fig. 6. The row length L as a function of galactocentric distance d (both parameters are normalized to the optical radius of the galaxy $D_{25}/2$). The dotted lines show regression fits; the solid line, the relation $L = d$.

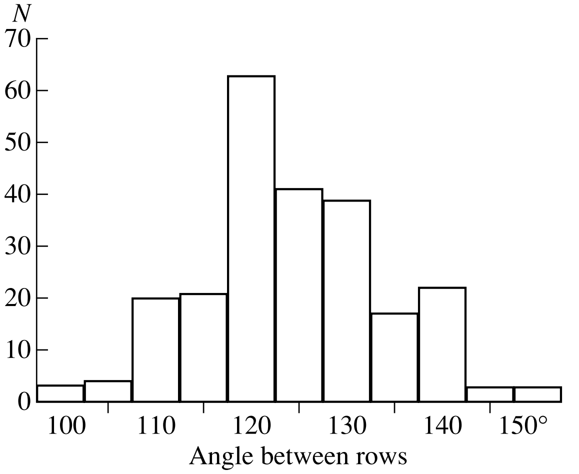

Fig. 7. Histogram of angles between adjacent rows.

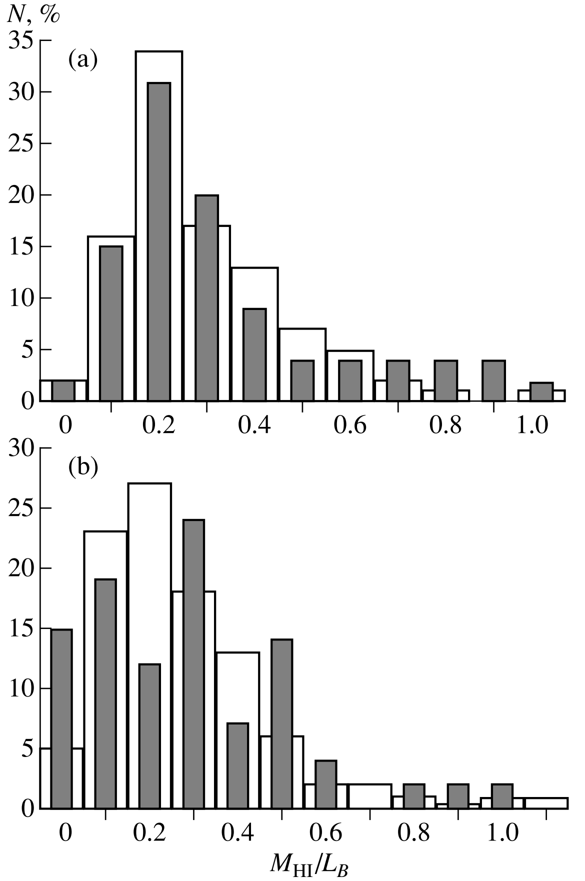

Fig. 8. Histogram of HI mass-to-light ratio $M_{\rm HI}/L_B$ (in solar units) for galaxies (a) with and (b) without rows. Shaded entries correspond to interacting galaxies

4. GAS CONTENT AND STAR FORMATION IN GALAXIES WITH ROWS

Information on the gas content in galaxies with rows is presented in Fig. 8a, which shows the distribution of the $M(\mathrm{HI})/L_B$ ratio (in solar units). This distribution peaks at 0.29, and the mean $M(\mathrm{HI})/L_B$ for the entire sample is virtually identical (0.30). These values are typical of late-type spirals. Galaxies with and without rows have similar distributions of gas content (Fig. 8b) (counting late-type galaxies exclusively in both cases). The somewhat increased fraction of interacting galaxies among gas-rich systems is a clear manifestation of observational selection effects: the dynamically cold disk component—a layer of diffuse medium—is more susceptible to gravitational perturbations and provides a favorable medium for the development of high-contrast tidal features, so that its brightness can be enhanced substantially by triggered star formation in perturbed regions.

UBV photometry is available for about half of all galaxies with rows. On a $(U-B)_0$–$(B-V)_0$ color–color diagram, they lie along the intrinsic galaxy color line and do not deviate appreciably from this line (Fig. 9). Although the rows themselves contain multiple sites of star formation, their contribution to the integrated luminosity is too weak to have any appreciable effect on the galaxy colors.

5. DISCUSSION AND CONCLUSIONS

It is evident from the distribution of morphological types (Figs. 2a, 2b) that galaxies with rows are found predominantly among late-type spirals, whereas, in the relatively small number of cases involving early-type (Sa–Sb) galaxies, most rows are found in interacting systems. Overall, there are nearly twice as many interacting systems as noninteracting systems among galaxies with rows. It is evident that interaction acts only as a factor stimulating the formation of rows, probably via the development of strong shocks associated with gas perturbations in the galactic disks. Of the factors considered here, only one apart from membership in an interacting system affects the fraction of galaxies with rows, the relative gas content (HI). Rows are virtually absent from galaxies with flocculent spiral patterns, even in those with a large number of star-forming regions. This applies to both interacting and noninteracting galaxies.

The high gas content and relatively high fraction of interacting systems among galaxies with rows is consistent with the idea that the formation of rows is a hydrodynamic process associated with density waves. If we consider only late-type galaxies, which usually have high gas contents, galaxies with rows do not differ systematically from galaxies with smooth spirals in their integrated properties (luminosity, color index, gas content, rotational velocity). Moreover, in some galaxies, a prominent row is found on one side of the disk, while a normal smoothly curved spiral is

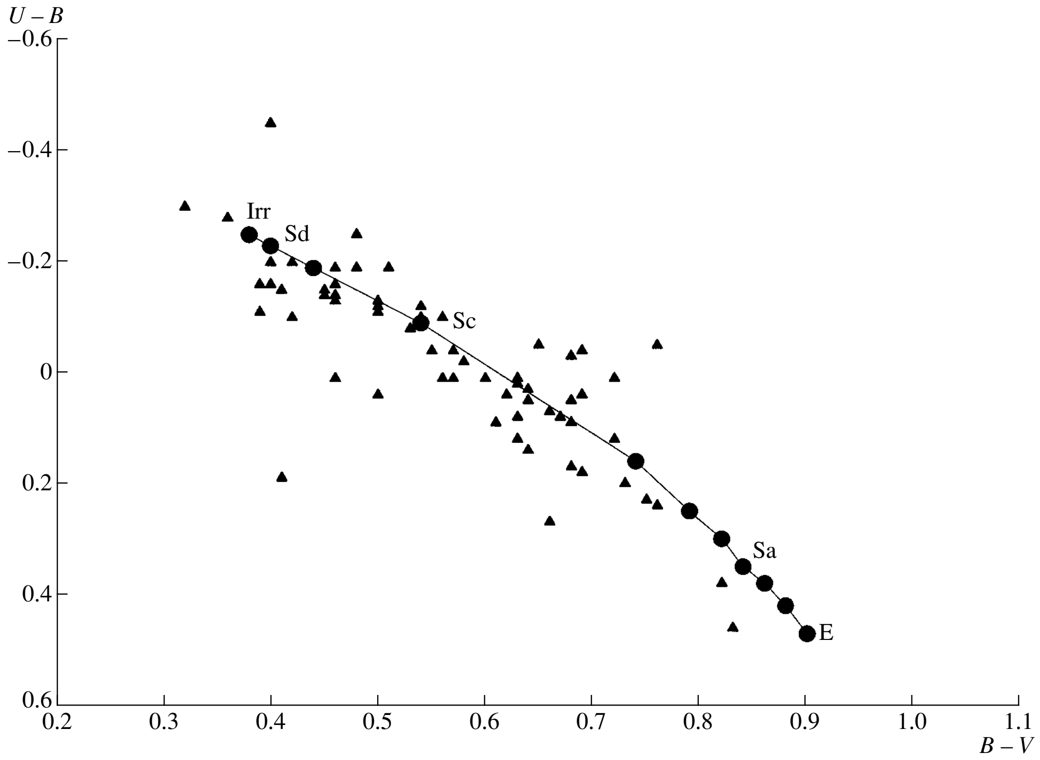

Fig. 9. Color–color diagram for galaxies with rows. The color indices are corrected for Galactic selective extinction and inclination to the line of sight (based on LEDA data).

located on the opposite side at the same galactocentric distance. This leads us to conclude that rows are transient features of the spiral structure in galaxies whose spiral arms are formed by large-scale density waves. The frequency of occurrence of galaxies with well-defined rows implies that, on average, a normal spiral galaxy has rows during no less than 6–10% of its lifetime. However, the most well-defined rows are usually observed in galaxies with well-defined (Grand Design) spiral structure, so that the fraction of galaxies with rows and, consequently, the tidal lifetime of the rows in these galaxies should exceed these estimates.

If each row were to be formed independently of other rows, adjacent straight features in galaxies would be an extremely rare phenomenon. The fact that many galaxies exhibit straight spiral-arm rows passing into each other indicates that adjacent rows form and disrupt synchronously. The necessary condition for this should be the presence of a large-scale density wave to synchronize star formation on scales ranging from several kpc to more than 10 kpc. In this case, the development of nonstationary processes associated with the external interaction stimulates the appearance of rows in interacting galaxies. Cases in which these conditions are fulfilled are illustrated by comparatively nearby galaxies with numerous rows belonging to gas-rich interacting systems (M51, M101).

6. ACKNOWLEDGMENTS

This work was supported by the program “Universities of Russia” and the Federal Science and Technology Program in Astronomy.

REFERENCES

1. A. D. Chernin, A. V. Zasov, V. P. Arkhipova, and A. S. Kravtsova, Pis’ma Astron. Zh. **26**, 342 (2000) [Astron. Lett. **26**, 285 (2000)].
2. A. Sandage and J. Bedke, *Atlas of Galaxies* (Washington, DC, 1988).
3. B. A. Vorontsov-Vel’yaminov, V. P. Arkhipova, and A. A. Krasnogorskaya, *Morphological Catalogue of Galaxies* [in Russian] (Mosk. Gos. Univ., Moscow, 1962,1963, 1964, 1968, 1974), Parts 1–5.
4. G. de Vaucouleurs, A. de Vaucouleurs, H. G. Corvin, *et al., Third Reference Catalogue of Bright Galaxies* (Springer-Verlag, New York, 1991).
5. B. A. Vorontsov-Vel’yaminov, R. I. Noskova, and V. P. Arhipova, Astron. Astrophys. Trans. **21** (2002) (in press); http://www.sai.msu.su/sn/VV/.
6. A. M. Fridman, O. V. Khoruzhii, V. V. Lyakhovich, *et al.,* Astron. Space Sci. **252**, 115 (1997).

Translated by A. Dambis

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 854–860. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 977–984.
Original Russian Text Copyright © 2001 by Wiebe, Tutukov, Shustov.

The Radial Gradient of Heavy-Element Abundances in Disk Galaxies

D. S. Wiebe, A. V. Tutukov, and B. M. Shustov

Institute of Astronomy, Moscow, Russia

Received May 17, 2001

Abstract—Various origins for the formation of the heavy-element abundance gradients observed in nearly all disk galaxies are analyzed in the framework of evolutionary models. In an isolated galaxy, there is a radial gradient of the abundance of heavy elements only early in its evolution (the first several billion years), which subsequently practically disappears. The gradients of chemical compositions of young objects and the interstellar gas require that typical disk galaxies be open systems (i.e., that they eject some heavy elements into circumgalactic space and/or accrete intergalactic gas) and that the rates of both processes be dependent on galactocentric distance. © *2001 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The chemical evolution of a disk galaxy is determined by star formation, stellar evolution, enrichment of the interstellar gas with heavy elements, the ejection of heavy elements into intergalactic space, and the accretion of intergalactic gas. Observations show that, in most disk galaxies, these processes interact to produce negative radial gradients of abundances of heavy elements—oxygen, iron, nitrogen, carbon, etc. This has been the subject of many observational and theoretical studies (see reviews [1, 2]).

In the disk of our own Galaxy, virtually all the most abundant elements (C, N, O, Ne, S, Fe, Ar, Al) show radial gradients. The gradients for many types of objects with ages from 10 to 10^4 Myr within $R \sim 5$–15 kpc of the Galactic center (open clusters, HII regions, planetary nebulae, B stars) have similar values of -0.05 to -0.08 dex kpc^{-1} (see, e.g. [3–6] and many other papers). However, the uncertainties are fairly high, even if we consider a single element and type of object. For example, estimates of the oxygen abundances in HII regions vary from -0.13 [7] to -0.05 [8]. At the same time, according to the data of Kaufer *et al.* [9], the relative oxygen abundances in B stars are virtually constant (and equal to zero) in the same interval of Galactocentric distances. In addition, four B stars in the central part of the Galaxy ($R \sim 2$ kpc) whose chemical composition was determined by Smartt *et al.* [10] were also found to have solar oxygen abundances.

Nevertheless, it can be considered well established that heavy-element abundances are a factor of two to five higher in the central part of the Galaxy than at its periphery. The similarity between the gradients inferred from young and old objects indicates that they do not depend significantly on age, at least during most of the Galaxy's lifetime. Observations of disk galaxies show that chemical composition gradients of about -0.03 to -0.1 dex kpc^{-1} are characteristic of such systems [11–15], so that the Milky Way is typical in this respect.

There have been many attempts to explain the origin of radial gradients of chemical composition. As a rule, "static" models (i.e., including only star formation and the return of gas to the interstellar medium by evolved stars) fail to reproduce the gradients: it is necessary to take into account gas motions in one way or another [16]. The dynamical factors involved can be subdivided into two groups : (1) accretion at a rate that depends on galactocentric distance and (2) radial gas flows in the galactic disk.

The first hypothesis is especially popular, since introducing accretion makes it possible to simultaneously resolve the "G-dwarf problem" [1]. To reproduce the observed gradients, we must assume that the time scale for accretion depends on galactocentric distance R, so that gas accumulates at the center of a galaxy and heavy elements are produced at a higher rate there [16–19].

The accreting gas could originate from (1) the galactic halo (in this case, accretion corresponds to the ongoing formation of the disk) and (2) matter swept up by the galaxy as it moves through the intergalactic medium. In their theoretical analysis of metallicity gradients in several spiral galaxies, Molla *et al.* [20] concluded that their results were inconsistent with sustained gas infall onto the disk, arguing in favor of the former scenario. On the other hand, Chiappini *et al.* [18] suggest that the formation of a disk from halo gas is inconsistent with the fact that these subsystems have very different angular-momentum distributions.

Kennicut [21] summarized arguments against sustained gas infall onto the Galactic disk. Modern observational data can be reconciled with the accretion scenario only if the accretion rate is much lower than the current star-formation rate. Direct observations of gas accretion onto the Galactic disk are available only for individual high-velocity clouds and say nothing about the role in the disk's evolution played by the periodic accretion of such clouds. Even if these clouds provide an infall of intergalactic gas onto the disk at a mean rate of $0.5\,M_\odot$/yr, as estimated by Blitz *et al.* [22], due to its episodic nature, this process cannot produce a regular and sustained heavy-element distribution across the disk. Moreover, the typical sizes of high-velocity intergalactic clouds (up to 25 kpc [22]) substantially exceed the interval of Galactocentric distances where the gradient is observed (about 10 kpc).

Another argument against a radial dependence of the accretion rate is that the accumulation of gas by a galaxy can be a self-regulating process. An enhanced accretion rate in the central region of a galaxy should lead to more active star formation and, consequently, to more intense energy release by young stars. This energy (e.g., in the form of galactic fountains and wind) can act against the infall of gas onto the disk, weakening the dependence of the accretion rate on R.

A possible alternative scenario to radially dependent accretion could be radial gas flows in the galactic disk. Lacey and Fall [23] identified three main origins for the development of radial gas motions in a galactic disk: (1) infall onto the disk of a material with low angular momentum; (2) viscosity of the gaseous disk; (3) gravitational interaction between the gas and spiral density waves. It has been shown that taking these processes into account together with radially dependent accretion can explain the development of chemical-composition gradients (see e.g. [24] and [25] and references therein). However, thus far, radial gas flows lack a sound theoretical basis and must be artificially added to models. Moreover, to equalize the radial gas distribution and solve the G-dwarf problem, a model must include radially dependent accretion even if radial inflows are present [25].

Our model for the evolution of a disk galaxy can reproduce the distribution of stellar metallicities in the solar neighborhood even without assuming accretion of gas with low Z [26, 27]. Thus, the main incentive for introducing accretion into the model disappears. The aim of this paper is to elucidate whether this model can also reproduce the radial gradients of heavy-element abundances. Section 2 describes our disk-galaxy model; Section 3, the results of modeling the evolution of the Galaxy using a closed model, which we use as a reference solution. Section 4 gives the results of introducing various factors describing matter exchange with the intergalactic medium into our standard model. In Section 5, we discuss our results and formulate our conclusions.

2. THE MODEL

As a basis, we used a single-zone, evolutionary model for a disk galaxy, for which the underlying ideas were formulated by Firmani and Tutukov [28]. A detailed description of this model can be found in [27]. On the whole, the model follows Tinsley's [29] formalism with one important difference: the thickness of the gaseous disk of the galaxy is not fixed and is determined by two competing processes, energy input into the interstellar medium as a result of supernova explosions and dissipation of the kinetic energy of gas in interstellar-cloud collisions. Thus, the main quantitative parameter of the interstellar gas is not its surface density, as in many other models, but its volume density ρ_{g}.

We take the star-formation rate ψ in the Galaxy to be proportional to the square of the gas volume density:

$$\psi \propto \rho_{\mathrm{g}}^2. \tag{1}$$

This relation follows from the assumption that star formation is self-regulated by the ionizing radiation of young, massive stars [30, 31]. The evolution of the gas mass in the galaxy is described by the equation

$$\frac{\mathrm{d}M_g}{\mathrm{d}t} = -\psi(t) \tag{2}$$

$$+\int\limits_{M_{\min}}^{M_{\max}} psi(t-\tau_{\mathrm{M}})(M-M_{\mathrm{r}})\phi(M)\mathrm{d}M - \dot{M}_{\mathrm{g}}^{\mathrm{out}} + \dot{M}_{\mathrm{g}}^{\mathrm{in}},$$

where M_{g} is the mass of gas in the galaxy, $\psi(t)$ is the star-formation rate, $\phi(M)$ is the initial mass function (IMF), $M_{\min}$ and $M_{\max}$ are the limiting masses of forming stars, τ_{M} is the lifetime of a star of mass M, M_{r} is the mass of a stellar remnant, $\dot{M}_{\mathrm{g}}^{\mathrm{out}}$ is the rate of mass ejection into circumgalactic space, and $\dot{M}_{\mathrm{g}}^{\mathrm{in}}$ is the rate of accretion of intergalactic gas. We will use a Salpeter IMF with mass limits $M_{\min} = 0.1\,M_\odot$ and $M_{\max} = 100\,M_\odot$.

The evolution of the mass of chemical element i is described by the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(Z_i M_{\mathrm{g}}) \tag{3}$$

$$= \int\limits_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} \psi(t-\tau_{\mathrm{M}})\,[Z_i(t-\tau_{\mathrm{M}})(M-M_{\mathrm{r}})$$

$$+P_i(M)]\,\phi(M)\mathrm{d}M - Z_i(t)\psi(t) - \dot{M}_i^{\mathrm{out}} + \dot{M}_i^{\mathrm{in}}.$$

Here, $Z_i(t)$ is the relative abundance of element i at time t and $P_i(M)$ is the mass of element i synthesized in a star of mass M. The last two terms describe mass exchange with the intergalactic medium. We will analyze the evolution of the abundance of oxygen—neglecting, for the sake of simplicity, the production and destruction of oxygen in intermediate-mass stars—and of nitrogen (an element produced largely in long-lived stars). We adopted the oxygen synthesis data from [32]. When determining nitrogen abundances, we used the results of computations reported in [33].

In our previous papers [26, 27, 34, 35], we showed that our model can successfully reproduce a number of fundamental properties of disk galaxies: the current star-formation rate, distribution of stellar heavy-element abundances, dependence of Z on mass (luminosity) of a disk galaxy, the [O/Fe]–[Fe/H] correlation, etc.. Unlike other similar models, our model can also reproduce the distribution of heavy elements perpendicular to the galactic disk (i. e., the vertical gradient of chemical composition).

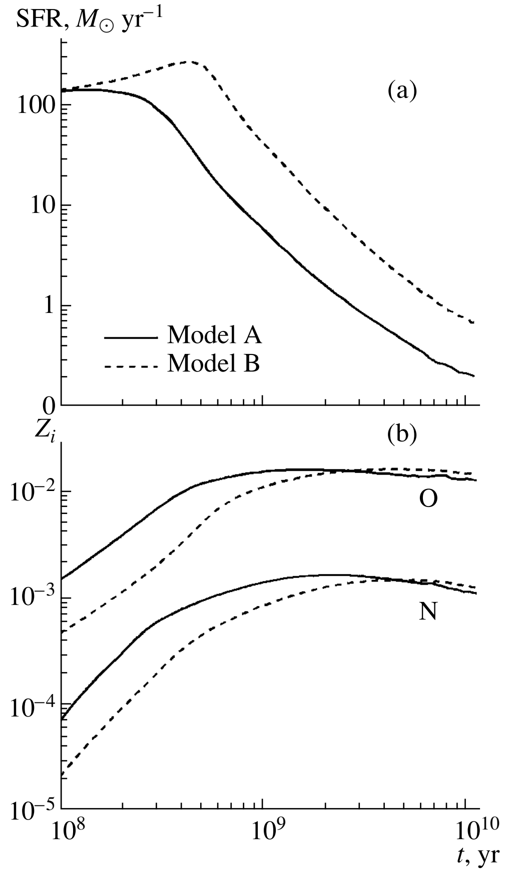

Fig. 1. Results for the standard model.

We will analyze the radial gradient of chemical composition in the Galaxy; thus, single-zone models are no longer adequate. We chose to model the radial structure of the Galaxy using a superposition of two models: one describing the inner part of the galaxy (Model A) and the other describing the outer part (Model B). Increasing the number of models (to improve the radial resolution) does not significantly change the overall results.

In both models, the system is a cylinder with a fixed mass M and radius R and a time-dependent height. We assume that $M(R) \propto R$ in the Galaxy, as follows from its flat rotation curve. Therefore, the parameters of our models are Model A—$M = 0.5 \times 10^{11}\,M_\odot$, $R = 5$ kpc and Model B—$M = 1.5 \times 10^{11}\,M_\odot$, $R = 15$ kpc. In both models, the initial half-thickness of the disk H was taken to be equal to 15 kpc. We considered both closed and open models. In the closed model, the Galaxy evolves as an isolated system without mass exchange with circumgalactic space: ($\dot{M}_{\mathrm{g}}^{\mathrm{in}} = \dot{M}_{\mathrm{g}}^{\mathrm{out}} = \dot{M}_i^{\mathrm{in}} = \dot{M}_i^{\mathrm{out}} = 0$). In the open model, the Galaxy ejects heavy elements or accretes intergalactic gas with $Z = 0$. We integrated the model equations to a time $t = 1.2 \times 10^{10}$ yr.

3. RESULTS FOR THE CLOSED MODEL

We used the closed model, in which the Galaxy does not exchange matter with the surrounding space, as our reference solution. Figure 1 shows its main characteristics: star formation rate (SFR) and abundances of oxygen Z_{O} and nitrogen Z_{N} as functions of time. The solid and dashed curves correspond to Model A and Model B, respectively. At the initial stage of the Galaxy's life, the gradient of the oxygen and nitrogen abundances is about 0.05 kpc^{-1}. This gradient comes about because, for the adopted mass distribution in the Galaxy ($M(R) \propto R$), the initial density in the inner regions is a factor of three higher than in the outer regions. Combined with the adopted quadratic dependence of the SFR on the gas density, this means that the time scale for star formation at the Galactic center is shorter than at the periphery, implying a faster accumulation of heavy elements at the center. Therefore, the closed model can adequately explain the chemical-composition gradient for old objects. However, the gradient Z_{O} disappears by the time the Galaxy reaches an age of $t \sim 3 \times 10^9$ yr. The situation is similar for nitrogen; however, since this element is synthesized in long-lived, intermediate-mass stars, the nitrogen abundances in the outer and inner regions equalize at a later stage, when the Galaxy is $t \sim 5 \times 10^9$ yr old. This absence of a gradient is typical of models in which the star-formation parameters (and the rate of gas accretion onto the disk) are independent of R [16, 36]. A decrease in the gradient with time was also noted in models with radially varying accretion, such as those in [19, 20].

The absence of any significant dependence of $Z_{\rm O}$ and $Z_{\rm N}$ on t (Fig. 1b) in our model is due to the fact that we take into account the finite lifetime of stars. If the star-formation rate decreases with time, the enrichment of the ISM with heavy elements synthesized in massive stars at late stages of the Galaxy's evolution is compensated by the supply to the ISM of gas with low Z ejected by long-lived, low-mass stars formed during the main episode of star formation and whose lifetimes are shorter than the age of the Galaxy. As a result, Z is maintained at an approximately constant level. The "asymptotic" values $Z_{\rm O}$ and $Z_{\rm N}$ are determined by fixed parameters of stellar evolution and the shape of the IMF and are, accordingly, virtually the same in outer and inner regions.

3.1. Variation of the Duration of the "Initial Accretion Phase"

In our earlier paper [34], we assumed that a galaxy does not form instantaneously and introduced an "initial accretion phase," during which the mass of the galaxy increases linearly. By the end of this phase, the initial accretion ceases and the galaxy evolves as a closed system from this time onward. We assume here that the Galaxy forms instantaneously, to avoid introducing an additional parameter into the standard model, i.e., the difference between the durations of the initial accretion phase in outer and inner regions of the Galaxy (the meaning of this parameter is similar to that of the time scale for accretion of unprocessed halo gas in the models noted in the introduction).

It is evident from Fig. 1 that the development of an abundance gradient in the early evolution of the Galaxy is due to the fact that star formation does not proceed simultaneously in the outer and inner regions of the Galaxy. If the Galaxy forms from away its center to its periphery, the initial accretion is likely to last longer at the periphery than at the center. Allowance for this fact shifts the peak of the star-formation rate in the outer regions further into the future, prolonging the existence of the gradient. Note, however, that the oxygen abundance gradient has already disappeared 1–2 Gyr after the end of accretion. Therefore, for young objects to exhibit the gradient associated with the delay of primary star formation at the periphery of the Galaxy, this delay should be very large, comparable with the age of the Galaxy.

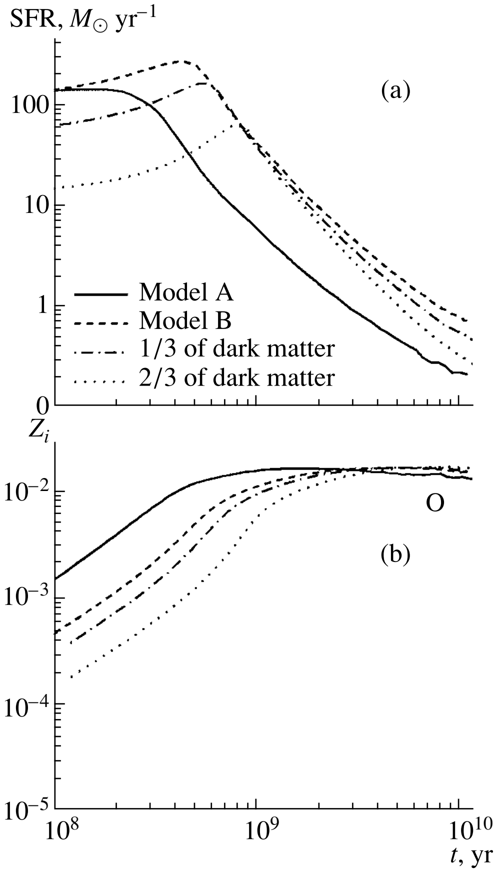

Fig. 2. Results for the standard model with dark matter.

3.2. The Role of Dark Matter

Our assumption that $M(R) \propto R$ in the Galaxy was based on its flat rotation curve. The shape of the rotation curve depends on the distribution of gravitating mass, which can consist partly of dark matter. We already pointed out that the main episode of star formation in the outer regions of the Galaxy is delayed relative to the central region, due to the low initial density there. If we suppose that the mass of the outer regions is partly contained in dark matter, we can adopt even lower gas densities, further delaying the onset of mass star formation. Together with curves for the standard model, Fig. 2 shows curves computed for models whose outer regions contain some mass fraction of dark matter. We considered two cases, with the dark-mass fractions equal to one-third and two-thirds of the initial mass of the outer region (corresponding to initial gas masses of $10^{11}\,M_\odot$ and $5 \times 10^{10}\,M_\odot$, respectively). It is evident from the figure that the gradient increases during the initial evolution of the Galaxy; however, even in this case, it eventually levels off. Note that the adopted dark-mass fractions for the disk are probably strongly overesti-

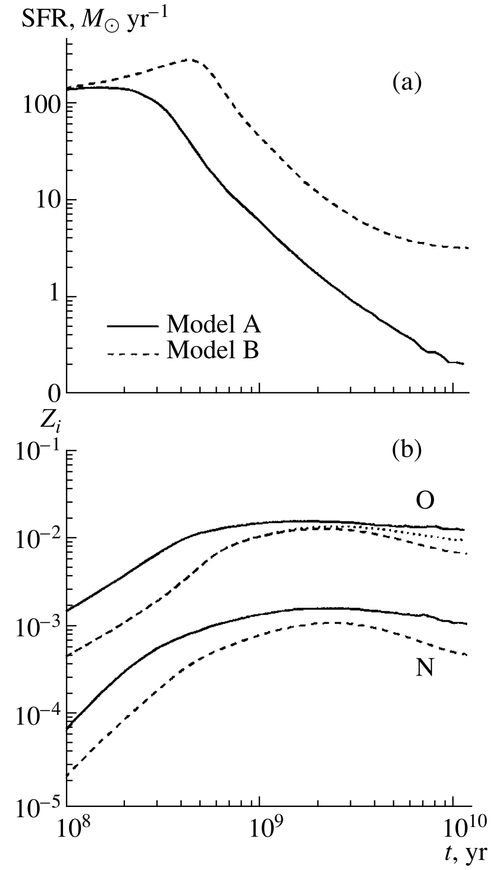

Fig. 3. Results for the model with accretion.

mated and that, in reality, dark matter accounts for a much smaller fraction of the disk mass [37].

4. RESULTS FOR THE OPEN MODEL

As we saw above, a closed model of the Galaxy fails to reproduce the chemical-composition gradients for young objects and the ISM, as in other studies [16, 23, 36]. This means that we must abandon the idea that the Galaxy is a closed system and suppose that it accretes intergalactic matter with low heavy-element contents and/or loses heavy elements, e.g., via Galactic wind and expulsion of dust by the pressure of stellar radiation [34, 38].

Note that, in this case, we are dealing with the accretion of intergalactic matter, not of gas ejected by halo stars. Accretion of processed halo matter is already implicitly included in the standard model and leads to constancy of the abundance gradients.

4.1. Accretion of Intergalactic Gas

The development of gradients can be affected not only by accretion with a rate that decreases with time, specified in our model via the duration of the initial accretion phase, but also by accretion of ambient gas, again, provided that its rate depends on Galactocentric distance. In Fig. 3, the evolutionary curves for the inner region are computed for the standard closed model, while those for the outer region are computed assuming accretion of gas with $Z_{(\mathrm{O,N})} = 0$ at a constant rate of $2\,M_\odot$ yr^{-1}. Such accretion rates are expected for a galaxy moving at a velocity of 100–300 km/s through intergalactic space of density 10^{-3}–10^{-4} cm^{-3}. In this case, the negative abundance gradients of both oxygen and nitrogen vary little throughout the evolution of the Galaxy, in agreement with observations.

However, there are two arguments against this possibility. First, it is not obvious why accretion of intergalactic gas should be more efficient at the periphery than at the center of the Galaxy. Observational evidence of the infall of intergalactic clouds onto the Galactic disk has been found both far from the center [39] and in the solar neighborhood [40], and even near the Galactic center [41]. Second, the metallicities of intergalactic gas (and of high-velocity clouds) are lower than the solar value but are not equal to zero. The dashed curve in Fig. 3 corresponds to the case when the value of $Z_{(\mathrm{O,N})}$ in the accreting gas is one-third of the current oxygen abundance in the ISM at any time (according to X-ray data, Z_{O} in the intergalactic medium is about one-third of the solar value [42]). We can see that the gradient levels off with time, even if accretion is present.

4.2. Open Model with Ejection

In our previous papers [26, 38], we showed that a spiral galaxy loses heavy elements into the surrounding space via at least two mechanisms: (1) the expulsion of dust by stellar radiation pressure and (2) Galactic wind. In both processes, the fraction of ejected heavy elements depends on the mass of the galaxy: low-mass galaxies eject matter more efficiently. It seems logical to assume that a similar law works in our Galaxy—it is easier for heavy elements to leave the periphery than the center.

In Fig. 4, the curves for Model A again correspond to the standard case, while Model B assumes that 30% of the heavy elements synthesized in the disk of the Galaxy are ejected into the surrounding space via Galactic wind (with an ejection efficiency that is somewhat higher than our estimates for the Galaxy as a whole [26]) and the expulsion of dust by stellar radiation pressure. It is evident that these models are characterized by very constant gradients. This is due to the fact that a fixed fraction of heavy elements is ejected from the Galaxy at any given time.

Of course, the actual situation is more complicated: as the Galaxy evolves, its shape and, therefore, its gravitational potential change. As a result, the fraction of ejected heavy elements should be a function of both radius and time. Moreover, as we already noted in the Introduction, the ejected matter interacts with infalling matter, affecting the parameters of both ejection and accretion. Nevertheless, this mechanism seems to be the most natural one.

5. DISCUSSION AND CONCLUSIONS

The distributions of heavy elements in both our own and other disk galaxies exhibit negative radial gradients of Z of about $\mathrm{d}\log Z/\mathrm{d}R \approx -0.03$ to $-0.1\ \mathrm{kpc}^{-1}$. Numerical simulations of the evolution of disk galaxies show that the development of these gradients cannot be explained by "static" models, in which individual portions of the galactic disk evolve in isolation from each other and from the circumgalactic medium. It is necessary to introduce gas motions within the disk and/or matter exchange with the surrounding space.

Of the various mechanisms invoked to explain the abundance gradient, the most popular is accretion at a rate that depends on galactocentric radius. This makes it possible to understand the observed gradients and also to solve the G-dwarf problem. Our model [27] was able to resolve the G-dwarf problem without invoking accretion. In the current paper, we have answered the question as to whether this same model is capable of reproducing the chemical-composition gradients in our Galaxy.

Our analysis shows that, in a closed model of the Galaxy, a gradient similar to that observed is present during the first several billion years, due to the time lag between the main star-formation episodes at the center ($R < 5$ kpc) and periphery ($R < 15$ kpc) of the Galaxy. The gradient then levels off. A modification of the closed model—introducing dark matter and delaying the onset of mass star formation at the periphery of the Galaxy—does not solve the problem of the decrease in the gradient with time. The difference in chemical composition between the center and periphery of the Galaxy that develops early in the Galaxy's evolution can persist until the present time only if the formation of the outer regions of the Galaxy is not yet completed; i.e., the main episode of star formation there has not yet ended.

To explain the observed gradient (in our model), we must suppose that the Galaxy actively exchanges matter with intergalactic space. Introducing accretion of intergalactic matter makes it possible to reproduce the observed gradient only if (1) gas accretes mainly at the periphery of the Galaxy and (2) the abundance of heavy elements in this gas is $\ll Z_\odot/3$. We believe that it is more natural to suppose that the ejection of heavy elements from the Galaxy occurs at a rate that depends on R.

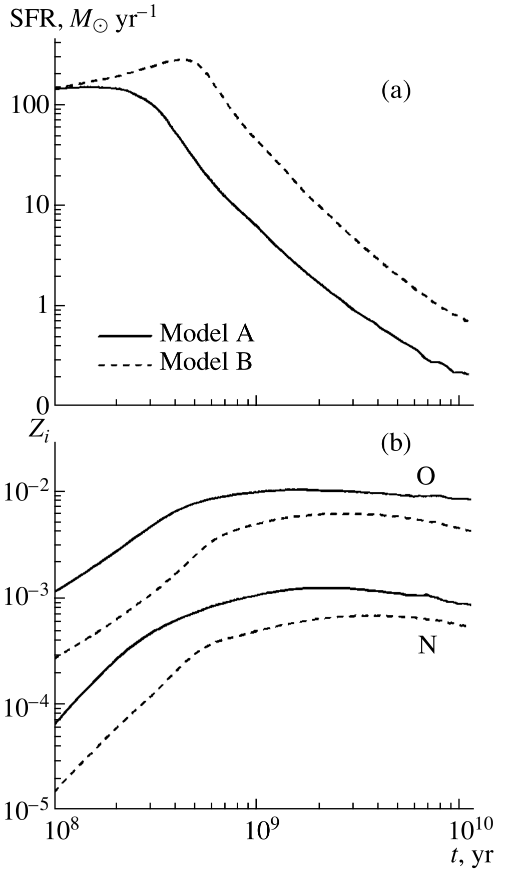

Fig. 4. Results for the model with ejection.

Possible explanations of the negative gradient of chemical composition are not confined to accretion of gas and loss of heavy elements by the Galaxy. A similar gradient also develops if radially dependent star-formation parameters are introduced into the model. For example, the coefficient of proportionality in (1) could depend on R if we allow for stimulation of star formation by spiral arms. The minimum mass of forming stars could also depend on the Galactocentric radius, due to the decrease in the mean blackbody temperature of the interstellar gas from the center toward the periphery. An increase in $M_{\min}$ increases the number of supernovae per unit mass of gas converted into stars; i.e., it increases the efficiency of converting gas into heavy elements. However, issues connected with quantitative estimates of these parameters are not sufficiently well understood to be included in realistic evolutionary models of the Galaxy and we, accordingly, do not discuss them further here.

6. ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-16037) and grant no. 00-15-96722 of the President of the Russian Federation. D. Wiebe acknowledges support from the Russian Foundation for Basic Research (project no. 01-02-06080).

REFERENCES

1. B. E. J. Pagel, *Nucleosynthesis and Chemical Evolution of Galaxies* (Cambridge Univ. Press, Cambridge, 1997).
2. R. B. C. Henry and G. Worthey, Publ. Astron. Soc. Pac. **111**, 919 (1999).
3. P. A. Shaver, R. X. McGee, L. M. Newton, *et al.*, Mon. Not. R. Astron. Soc. **204**, 53 (1983).
4. B. A. Twarog, K. M. Ashman, and B. J. Antony-Twarog, Astron. J. **114**, 2556 (1997).
5. G. Carraro, Y. K. Ng, and L. Portinari, Mon. Not. R. Astron. Soc. **296**, 1045 (1998).
6. W. R. J. Rolleston, S. J. Smartt, P. L. Dufton, and R. S. I. Ryans, Astron. Astrophys. **363**, 537 (2000).
7. M. Peimbert, J. F. Rayo, and S. Torres-Peimbert, Astrophys. J. **220**, 516 (1978).
8. A. Afflerbach, E. Churchwell, and M. W. Werner, Astrophys. J. **478**, 190 (1997).
9. A. Kaufer, Th. Szeifert, R. Krenzin, *et al.*, Astron. Astrophys. **289**, 740 (1994).
10. S. J. Smartt, K. A. Venn, P. L. Dufton, *et al.*, Astron. Astrophys. **367**, 86 (2001).
11. D. Zaritski, R. C. Kennicutt, Jr., and J. P. Huchra, Astrophys. J. **420**, 87 (1994).
12. R. C. Kennicutt and D. R. Garnett, Astrophys. J. **456**, 504 (1996).
13. E. D. Skillman, R. C. Kennicutt, G. A. Shields, and D. Zaritski, Astrophys. J. **462**, 147 (1996).
14. D. R. Garnett, G. A. Shields, E. D. Skillman, *et al.*, Astrophys. J. **489**, 63 (1997).
15. L. van Zee, J. J. Salzer, M. P. Haynes, *et al.*, Astron. J. **116**, 2805 (1998).
16. L. Portinari and C. Chiosi, Astron. Astrophys. **350**, 827 (1999).
17. F. Matteucci and P. François, Mon. Not. R. Astron. Soc. **239**, 885 (1989).
18. C. Chiappini, F. Matteucci, and R. Gratton, Astrophys. J. **477**, 765 (1997).
19. N. Prantzos and S. Boissier, Mon. Not. R. Astron. Soc. **313**, 338 (2000).
20. M. Mollá, F. Ferrini, and A. I. Díaz, Astrophys. J. **475**, 519 (1997).
21. R. C. Kennicutt, in *New Light of Galaxy Evolution (IAU Symposium 171)*, Ed. by R. Bender and R. L. Davies (Kluwer, Dordrecht, 1996), p. 11.
22. L. Blitz, D. N. Spergel, P. J. Teuben, *et al.*, Astrophys. J. **514**, 818 (1999).
23. C. G. Lacey and S. M. Fall, Astrophys. J. **290**, 154 (1985).
24. S. G. Simakov, Astron. Zh. **70**, 277 (1993) [Astron. Rep. **37**, 144 (1993)].
25. L. Portinari and C. Chiosi, Astron. Astrophys. **355**, 929 (2000).
26. B. M. Shustov, D. S. Wiebe, and A. V. Tutukov, Astron. Astrophys. **345**, 93 (1999).
27. D. S. Wiebe, A. V. Tutukov, and B. M. Shustov, Astron. Zh. **75**, 3 (1998) [Astron. Rep. **42**, 1 (1998)].
28. C. Firmani and A. V. Tutukov, Astron. Astrophys. **264**, 37 (1992).
29. B. Tinsley, Fundam. Cosm. Phys. **5**, 287 (1980).
30. A. Tutukov and É. Krugel, Astron. Zh. **57**, 942 (1980) [Sov. Astron. **24**, 539 (1980)].
31. D. P. Cox, Astrophys. J. Lett. **265**, L61 (1983).
32. A. Maeder, Astron. Astrophys. **264**, 105 (1992).
33. P. Marigo, Astron. Astrophys. **370**, 194 (2001).
34. D. S. Wiebe, B. M. Shustov, and A. V. Tutukov, Astron. Astrophys. **345**, 93 (1999).
35. A. V. Tutukov, B. M. Shustov, and D. S. Wiebe, Astron. Zh. **77**, 803 (2000) [Astron. Rep. **44**, 711 (2000)].
36. M. G. Edmunds and R. M. Greenhow, Mon. Not. R. Astron. Soc. **272**, 241 (1995).
37. D. Romano, F. Matteucci, P. Salucci, and C. Chiappini, Astrophys. J. **539**, 235 (2000).
38. B. M. Shustov and D. S. Wiebe, Astron. Zh. **75**, 3 (1995) [Astron. Rep. **39**, 1 (1995)].
39. B. P. Wakker, J. C. Howk, B. D. Savage, *et al.*, Nature **402**, 388 (1999).
40. D. M. Meyer, M. Jura, I. Hawkins, and J. A. Cardelli, Astrophys. J. Lett. **437**, L59 (1994).
41. D. A. Lubowich, J. M. Pasachoff, Th. J. Balonek, *et al.*, Nature **405**, 1025 (2000).
42. Y. Fukazawa, K. Makishima, T. Tamura, *et al.*, Mon. Not. R. Astron. Soc. **313**, 21 (2000).

Translated by A. Dambis

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 861–864. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 985–989.

Do the Galaxies NGC 936 and NGC 3198 Possess Massive Spheroidal Subsystems?

A. V. Khoperskov

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

Received March 5, 2001

Abstract—The condition for gravitational stability of the stellar disks of the galaxies NGC 936 and NGC 3198 makes maximum disk models unacceptable. We present mass estimates for these objects' spheroidal components. The mass of the dark halo of NGC 3198, within four disk radial scale lengths, exceeds its disk mass by a factor of 1.6 to 2. The masses of the disk and spheroidal subsystem (halo + bulge), within four radial scale lengths, are approximately the same for NGC 936. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The possible existence of an invisible massive component in S galaxies is one of the most important problems in the physics of galaxies. The presence of extended regions in the gas rotation curves showing no decrease with radius beyond the optical radius are among the features that provide evidence for the presence of massive halos, whose mass may appreciably exceed that of the visible matter in the disk, M_d. Within the optical radius, the halo mass M_h can be comparable to and even exceed M_d. Among other signatures, this is suggested by the efficient stabilization of the global bar mode by the massive spheroidal subsystem [1–3], but here the situation is more complex: the asymmetrical bar forming in the disk leads to a gravitational interaction with the halo matter, resulting in a transfer of angular momentum from the disk to the spheroidal subsystem [4]. Analyzing the results of dynamical modeling of interacting disk/bar/halo subsystems, Dabattista and Sellwood [5] concluded that the halo mass within the optical radius had to be small compared to the disk mass and, in particular, that maximum disk models (MDMs) were suitable for the galaxies NGC 936 and NGC 3198. The MDM is based on the observed rotation curve of the gaseous subsystem, V_{gas}, and the condition that the disk mass be the maximum possible for the case when the mass-to-luminosity ratio in the disk population is approximately constant and the circular velocity in the galaxy's equatorial plane V_c is equal to V_{gas}.

The observed radial distribution of stellar velocity dispersions constrains the value of $\mu \equiv M_s/M_d$, where M_s is the mass of the spheroidal halo + bulge + core subsystem. For the stellar disk to be gravitationally stable, the system must not be cool: $c_r > c_r^{\text{crit}}$, where c_r is the stellar radial velocity dispersion [6–9]. The velocity dispersions in the stellar disks of NGC 936 and NGC 3198 are known. Earlier [10], we considered a model that satisfactorily described the velocity curve and stellar velocity dispersion in NGC 3198. Here, we show that the gravitational stability condition for the stellar disks makes maximum disk models unacceptable for NGC 936 and NGC 3198 and leads to underestimated masses for their spheroidal subsystems.

2. GRAVITATIONAL STABILITY OF A STELLAR DISK

We assume that there is an exponential distribution of matter in the disk, $\varrho(r,z) = \varrho_0 \operatorname{sech}^2(z/h)\exp(-r/L)$, so that the disk subsystem is governed by three parameters: the disk's radial scale length L, vertical scale length h, and central surface density $\sigma_0 = 2h\varrho_0$.

We will describe the spheroidal components (halo, bulge, core) with the masses M_h, M_b, M_c and scale lengths a, b, c, respectively. The matter of the bulge and core will be assumed to be confined to radii $(r_b)_{\max}$ and $(r_c)_{\max}$, respectively. We will not consider the density distributions of the individual components but instead restrict our treatment to the integrated density $\mu = M_s/M_d = (M_h + M_b + M_c)/M_d$.

For the disk to be gravitationally stable to nonaxially symmetrical perturbations, it is necessary that $c_r \geq c_r^{\text{crit}}$. The minimum stellar radial velocity dispersion required for stability, c_r^{crit}, is traditionally described using Toomre's parameter $Q_T = c_r^{\text{crit}}/c_T$,

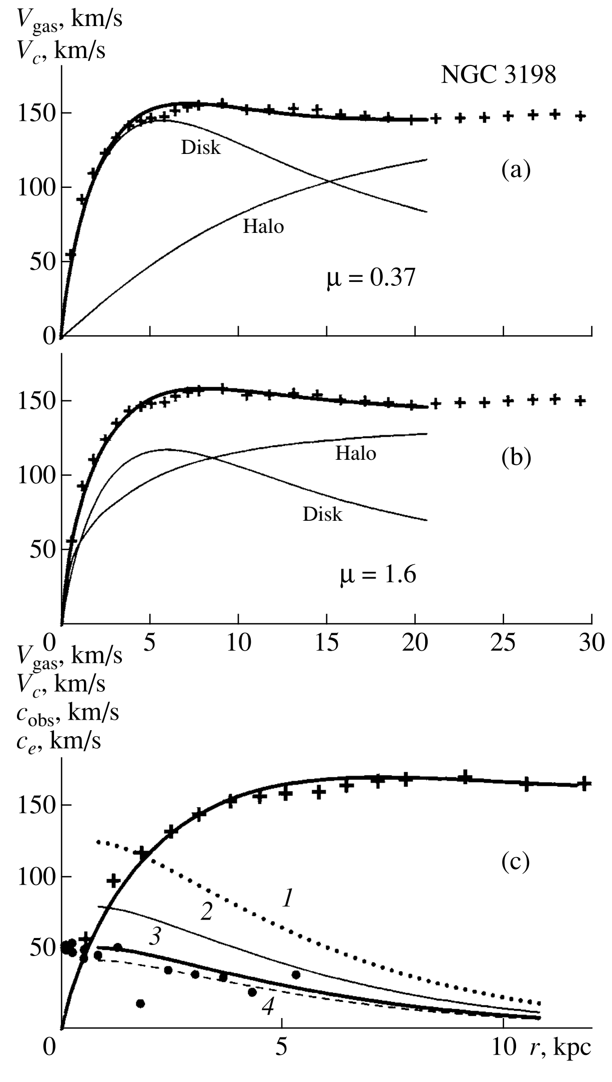

Fig. 1. Radial dependences of the gas rotational velocity $V_{\rm gas}$ (crosses), model circular velocity V_c (thick solid curve), and observed $c_{\rm obs}$ (dots) and theoretical c_e (curves *1–4*, see text) stellar velocity dispersions for various models for NGC 3198 (a, b, c).

where $c_T = 3.36 G\sigma/\varkappa$, σ is the local surface density, and $\varkappa = 2\Omega \times \sqrt{1 + r\,d\Omega/(2\Omega\,dr)}$ is the epicyclic frequency, which we will determine from the circular rotational velocity of the system as a whole, $V_c = r\Omega$. The local gravitational stability of an infinitely thin, uniform stellar disk rotating as a rigid body requires that $Q_T \geq 1$ [6]. This condition was derived for axially symmetrical perturbations. If the rotation is differential, as in the case of spiral waves, Q_T is increased. The value of Toomre's parameter depends on the character of the rotation curve. Values $Q_T \simeq 3$ are required for a flat rotation curve; lower values, $Q_T \simeq 2$, are needed in regions of rigid-body rotation [8]. Inhomogeneity of the disk, described by the radial dependence of the surface density and stellar velocity dispersion, leads to higher $c_r^{\rm crit}$ values. Dynamical modeling with Kuzmin–Toomre disks led to the conclusion that the condition $Q_T \geq 2.2$–2.4 must be fulfilled to ensure gravitational stability of a stellar disk [11]. The condition $Q_T \simeq 2.3$ was obtained for dynamical models neglecting vertical motions [11]. In some cases, it may be important to take into account the disk's finite thickness [12]. This factor has a stabilizing influence on gravitational instability, leading to lower radial velocity dispersions. Our three-dimensional dynamical models with exponential disks show that, as a rule, within two radial scale lengths ($r < 2L$), the disk is stable if $Q_T \gtrsim 1.5$ [10]. Strictly speaking, this parameter varies with radius (increases towards the center), but this only strengthens our conclusions.

Spectroscopic observations with a long slit positioned along the major axis of a galaxy yield the measured stellar velocity dispersion $c_{\rm obs}$ to be comparable to $c_e = \sqrt{c_\varphi^2 \sin^2 i + c_z^2 \cos^2 i}$. We will assume for the dispersion of azimuthal velocities, that $c_\varphi = \varkappa c_r/2\Omega$ and for the dispersion of vertical velocities, that $c_z = 0.5\,c_r$, corresponding to a lower limit of $c_z/c_r = 0.5$–0.6 for the dispersion ratio in the solar neighborhood [13, 14].

3. NGC 3198

The gas rotation curve $V_{\rm gas}$ has a plateau extending to at least 30 kpc. The radial scale length of the stellar disk is $L = 2.7$ kpc, and the galaxy's inclination is $i = 72°$ [15]. We will use the data of Bottema [16] on the radial distribution of $c_{\rm obs}$. The vertical scale length h is not known from observations; we will assume $h/L = 300\,{\rm pc}/2.7\,{\rm kpc} = 0.11$. The uncertainty in h will not influence our final conclusions. Because of the large inclination of NGC 3198, the uncertainty in the ratio c_z/c_r also has little influence on the conclusions.

The rotation curve for the MDM is shown in Fig. 1a. The contribution of the halo to the system's mass within $r = 4L$ does not exceed 37%. In this case, the central surface density of the disk is $\sigma_0 = 764\ M_\odot \cdot {\rm pc}^{-2}$. Figure 1b shows a model with a more massive spheroidal subsystem, with $\mu = 1.6$ ($\sigma_0 = 501\ M_\odot \cdot {\rm pc}^{-2}$). Figure 1c presents the radial distributions of the observed velocity dispersion, $c_{\rm obs}$ [16], and the $c_e(r)$ dependence for the central part of the disk, assuming $Q_T = 2.3$ in the MDM (curve *1*). The condition $c_{\rm obs} < c_e$ appears to be valid everywhere, with an ample margin. Since the old stellar disk of NGC 3198 should be gravitationally stable, the MDM is not appropriate for this galaxy if we assume $Q_T \geq 2.3$. The condition $Q_T = 2.3$ leads to dispersions so high that we will not further discuss this case, bearing in mind that the finite thickness of the disk plays a significant role in determining the gravitational stability of real galaxies [12].

In the MDM case, the criterion $Q_T = 1.5$ also means that $c_{\rm obs} < c_e$ (curve *2* in Fig. 1c). Only in the case of a more massive halo, with $\mu = 1.6, 2$ (curves *3*, *4*, Fig. 1c), is the disk stable in the region $r \leq 2L$. If we do not take into account data on the velocity dispersion, we can approximate the rotation curve $V_{\rm gas}$ for NGC 3198 with V_c over a wide range of μ values. For this galaxy, it is possible to develop models with $0.3 \lesssim \mu \lesssim \infty$. The case of $\mu = \infty$ was considered in [15]. However, taking into account the observational velocity dispersion data places strong constraints on this freedom. For NGC 3198, a massive disk with a low-mass halo is gravitationally unstable and the stellar disk satisfies the stability criterion only for $\mu \gtrsim 1.6$. Refinement of the minimum halo mass within the radius $r = 4L = 10.72$ kpc needed for stability requires the development of a dynamical model for the galaxy, which yields $\mu = 2.1$ [10]. This conclusion is completely compatible with the current paper's estimates.

4. NGC 936

NGC 936 is an SB0 galaxy containing a central bar and a disk with a radial scale length $L = 3.7$ kpc [17]. The rotation curve outside 8 kpc is unknown; thus, in the MDM, we will not be concerned with achieving a plateau at the edge of the stellar disk.

The orbital velocities of stars are known only within two disk radial scale lengths, and the gas rotational velocity is known only near $r \simeq 2L$. Therefore, when constructing the circular velocity, we will adopt $V_* \leqslant V_c$ for the velocity of the stars. This approach gives a lower limit for the disk mass. In addition to the thick disk, the central region's important components are the core and the bulge/bar needed to describe the kinematics within $r < L$.

In the MDM, the relative contribution of the spheroidal component is small, $\mu = 0.16$ (Fig. 2a). When $Q_T = \text{const} = 1.5$, we formally have $c_{\rm obs} < c_e$ in the region $r > 0.6L$ (curve *1* in Fig. 2c). We should take into account the radial dependence of Q_T needed for stability. At the stability limit, the ratio of the central radial velocity dispersion $c_r(r = 0)$ to the maximum circular velocity in a system with no halo is $\simeq 0.7$–0.8. Adopting the value 0.7, the condition $c_{\rm obs} < c_e$ is fulfilled everywhere (curve *2* in Fig. 2c). However, the margin is not wide and a component with moderate mass is sufficient to bring the $c_{\rm obs}$ and c_e values into agreement (Figs. 2b, 2c). The disk is clearly gravitationally stable for $\mu \simeq 1$. The color index of NGC 936 corrected for Galactic and intergalactic reddening is $B - V = 0.95$ [18]. Obviously, the old population of the stellar disk should have some margin of stability ($c_{\rm obs} > c_e$), such that our estimates of the characteristics of the spheroidal subsystem should be considered lower limits. In addition, the galaxy's observed central velocity dispersion $c_{\rm obs}$ includes a contribution from high-velocity bulge stars.

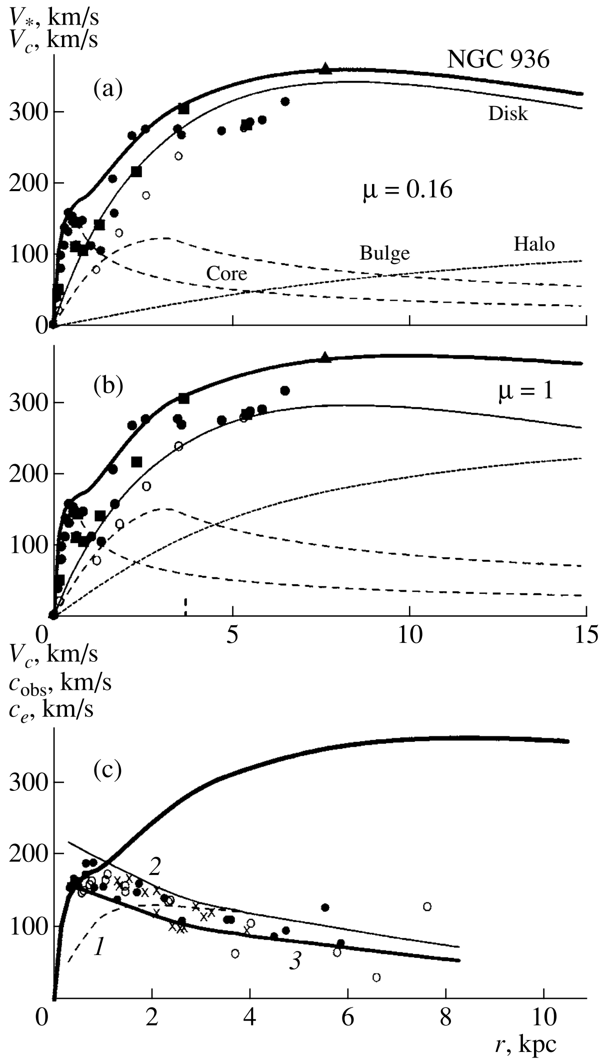

Fig. 2. (a, b) Radial dependences of the rotational velocities of the stars V_* (various symbols) and model circular velocity V_c (thick solid curves); (c) radial dependences of V_c (thick curve) and the observed $c_{\rm obs}$ (various symbols) and theoretical c_e (curves *1*–*3*, see text) stellar velocity dispersions for different models for NGC 936. The symbols for $c_{\rm obs}$ and V_* are the same as in [17].

The stellar velocity curve of NGC 936 is well known, but the gas rotational velocity is known only near $r \simeq 8$ kpc. The difference between $V_{\rm gas}$ and V_* is not large: $V_*/V_{\rm gas} = 320/360 = 0.89$. This provides additional evidence of a relatively massive halo, since dynamical models with no halo give $V_*/V_{\rm gas} \simeq 0.7$.

The resulting total mass estimate for the spheroidal subsystem of NGC 936 and the character of the mass distribution in this subsystem (a fairly rarified halo with $a/L = 4.64$ kpc/3.7 kpc $= 1.3$ and the absence of a very compact, massive core, $c/L = 0.143$ kpc/3.7 kpc $= 0.039$) permit the formation of a bar via the development of a global bar mode.

5. DISCUSSION

Debattista and Sellwood's [5] conclusion that there was no massive dark halo within the optical radius of the stellar disk is based on dynamical models reflecting the interaction between the bar and halo. The bar leads to appreciable rotation of the halo at on average about 25% of the disk rotational velocity. Note that inferred rotation of the halo should not be taken as evidence that it does not exist. Elliptical galaxies rotate at typical speeds of 20–60 km/s, permitting halo rotation in SB galaxies. The galaxy NGC 3198 probably has no significant bar, and this galaxy's rather massive halo may be nonrotating.

The conclusion of Debattista and Sellwood [5] that the MDM is applicable to NGC 936 and NGC 3198 contradicts the observational data on the stellar velocity dispersions in the disks of these objects. The gravitational stability condition places limits on the masses of the stellar disks. As a result, the mass of the spheroidal subsystem (halo + bulge) of NGC 936 may be comparable to the disk mass within the optical radius, whereas the mass of the dark halo of NGC 3198 may exceed the mass of the disk by a factor of 1.5–2.

ACKNOWLEDGMENTS

The author is grateful to Prof. A.V. Zasov for discussions of this work. Our models of the rotation curves used the GR code developed by A.N. Burlak. The author wishes to thank the referee for useful remarks. This study was financially supported by grants from the Russian Foundation for Basic Research (project no. 01-02-17597) and the Federal Scientific and Technological Program Astronomy.

REFERENCES

1. J. P. Ostriker and P. J. E. Peebles, Astrophys. J. **186**, 467 (1973).
2. R. A. James and J. A. Sellwood, Mon. Not. R. Astron. Soc. **182**, 331 (1978).
3. G. Efstathion, G. Lake, and J. Negroponte, Mon. Not. R. Astron. Soc. **199**, 1069 (1982).
4. B. Little and R. G. Carlberg, Mon. Not. R. Astron. Soc. **250**, 161 (1991).
5. V. P. Debattista and J. A. Sellwood, Astrophys. J. **543**, 704 (2000).
6. A. Toomre, Astrophys. J. **139**, 1217 (1964).
7. A. G. Morozov, Astron. Zh. **58**, 734 (1981) [Sov. Astron. **25**, 421 (1981)].
8. V. L. Polyachenko, E. V. Polyachenko, and A. V. Strel'nikov, Pis'ma Astron. Zh. **23**, 598 (1997) [Astron. Lett. **23**, 525 (1997)].
9. E. Griv, B. Rosenstein, M. Gedalin, and D. Eichler, Astron. Astrophys. **347**, 821 (1999).
10. A. V. Khoperskov, A. V. Zasov, and N. V. Tyurina, Astron. Zh. **78**, 213 (2001) [Astron. Rep. **45**, 180 (2001)].
11. E. Athanassoula and J. A. Sellwood, Mon. Not. R. Astron. Soc. **221**, 213 (1986).
12. N. N. Gor'kavyĭ and A. M. Fridman, *Physics of Planetary Rings* [in Russian] (Nauka, Moscow, 1994).
13. R. Wielen, Astron. Astrophys. **60**, 263 (1977).
14. R. Wielen and B. Fuchs, in *Kinematics, Dynamics and Structure of the Milky Way* (D. Reidel, Dordrecht, 1983), p. 81.
15. T. S. van Albada, J. N. Bahcall, K. Begeman, and R. Sanscisi, Astrophys. J. **295**, 305 (1985).
16. R. Bottema, Astron. Astrophys. **275**, 16 (1993).
17. J. Kormendy, Astrophys. J. **286**, 116 (1984).
18. G. de Vaucouleurs, A. de Vaucouleurs, H. Corwin, *et al., Third Reference Catalogue of Galaxies* (Springer-Verlag, New York, 1991).

Translated by N. Samus'

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 865–871. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 990–996.
Original Russian Text Copyright © 2001 by Malov.

Pulsed Optical Emission by Radio Pulsars

I. F. Malov

Astro–Space Center, Lebedev Physical Institute, Leninskiĭ pr. 53, Moscow, 117924 Russia

Received March 6, 2001

Abstract—The luminosity L of radio pulsars due to synchrotron radiation by the primary beam at the magnetosphere periphery is derived. There is a strong correlation between the observed optical luminosities of radio pulsars and the parameter $\dot{P}/P^4$ (where P is the pulsar period). This correlation predicts appreciable optical emission from several dozen pulsars, in particular, from all those with $P <$ 0.1 s. Agreement with optical observations can be achieved for Lorentz factors of the secondary plasma $\gamma_p = 2$–13. Plasma with such energies can be produced only when the magnetic-field structure near the neutron-star surface deviates substantially from a dipolar field. The peak frequency of the synchrotron spectrum should shift toward higher values as the pulsar period P decreases; this is, in agreement with observational data for 27 radio pulsars for which emission has been detected outside the radio band. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

After the discovery of pulsed optical emission from the radio pulsar PSR 0531+21 in the Crab Nebula, this emission was thought to be synchrotron radiation [1–3]. In 1972, Zheleznyakov and Shaposhnikov [4] presented a detailed model for the synchrotron radiation of PSR 0531+21 from optical to gamma-ray energies for the case of a power-law or piecewise power-law energy distribution for the emitting electrons, with electron pitch angles $\psi \sim 1$. Pacini and Salvati [5, 6] estimated the optical luminosity that could result from synchrotron radiation of the secondary electron–positron plasma in short-period pulsars. Another formula relating the luminosity of a radio pulsar in hard-energy bands (from optical to gamma-ray) to other parameters of these objects was obtained in [7]. One basic parameter critically affecting both the spectrum and character of the emission is the pitch angle ψ of the emitting charges. Our analysis of cyclotron instability in a pulsar magnetosphere and the kinetics of processes resulting in both on increase and decrease in the transverse momentum of the electrons enabled us to calculate the electron pitch-angle distribution. In the case of PSR 0656+14, which has been most thoroughly studied in the optical, the spectral peak corresponds to optical frequencies $(3-9)\times10^{14}$ Hz [8]. As was shown later [9], under certain circumstances, the peaks of the synchrotron spectra for other pulsars possessing pulsed optical emission could also be in the optical.

It is the aim of the present paper to investigate the possibilities of obtaining agreement between theoretical results and the available optical data for radio pulsars.

2. DESCRIPTION OF THE MODEL

Our studies [7, 10] used a model with an emitting torus localized near the light cylinder (Fig. 1). Relativistic electrons filling this torus are responsible for the synchrotron luminosity, which can be compared with the observed luminosities of pulsars. We shall use here a formula derived in [7], changing the factor related to the average energy of the emitting electrons in the cone of open field lines. The dependence of the electron Lorentz factor in the beam γ_b on the angle θ in the model of Fawley *et al.* [11] can be written [12] as

$$\gamma = \gamma_b f(\theta) = \gamma_b[1 - 0.944\left(\frac{\theta}{\theta_c}\right)^2 + 1.063\left(\frac{\theta}{\theta_c}\right)^4 - 1.119\left(\frac{\theta}{\theta_c}\right)^6], \quad (1)$$

where θ_c is the angular radius of the cone. Then, we can calculate the value of γ/γ_b using the formula

$$f(\theta) = \frac{\int_0^{\theta_c} f(\theta)d\theta}{\theta_c}, \quad (2)$$

which turns out to be 0.74. Using the model developed in [7], the total synchrotron power of the incoherent electron emission in the torus L takes the form

$$L \approx \frac{16\pi^8 e^4 R_*^6 I}{m^3 c^{11}} \frac{B_0^2 \gamma_b \dot{P} \sin^4\beta \sin^2\psi}{P^8}. \quad (3)$$

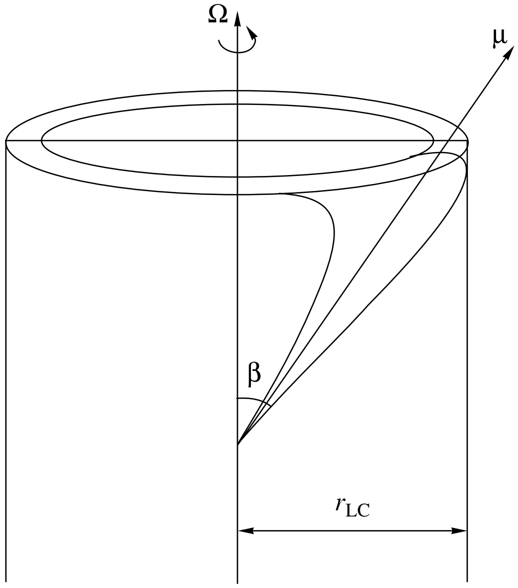

Fig. 1. Schematic model of the magnetosphere.

Here, R_* is the radius of the neutron star, I is its moment of inertia, $B_0 = B_s \sin\beta$ is the magnetic field at the neutron-star surface, β is the angle between the axis of rotation of the neutron star and the magnetic moment $\vec{\mu}$, P is the pulsar period, and $\dot{P}$ is its time derivative. We assume in (3) that the emission is generated near a light cylinder with radius

$$r_{LC} = cP/2\pi \tag{4}$$

at a distance from the neutron-star surface

$$r = \frac{r_{LC}}{\sin\beta} \tag{5}$$

and that the magnetic field has a dipolar structure:

$$B = B_s \left(\frac{R_*}{r}\right)^3. \tag{6}$$

3. PITCH ANGLES OF THE EMITTING ELECTRONS

Solving the kinetic equation describing the momentum distribution of the electrons in a beam at the magnetosphere periphery, we obtain for the pitch-angle distribution of the particles [8]

$$\xi(\psi) = Ce^{-A\psi^2}, \tag{7}$$

where $C = \mathrm{const}$,

$$A = \frac{2m^2c^4\gamma_b^2(\omega_B/\omega_p)^2}{\pi e^2\rho\gamma_p|E(k)|^2}, \tag{8}$$

γ_p is the Lorentz factor for the electrons of the secondary plasma, ρ is the radius of curvature of the field lines, $\omega_b = eB/mc$, $\omega_p = 4\pi n_p e^2/m$, n_p is the density of the secondary plasma, and $|E(k)|^2$ is the energy density of oscillations at the frequency corresponding to the resonance wavenumber k.

Using the Goldreich–Julian formula [13] for the density of particles in the beam

$$n_b = B/Pce \tag{9}$$

and the estimates

$$\rho = r_{LC} = cP/2\pi, \tag{10}$$

$$|E(k)|^2 = mc^2 n_b\gamma_b c/2\omega \tag{11}$$

for the average pitch angle of an emitting electron, we obtain [9]

$$\bar{\psi} = \frac{\int\limits_0^\infty \psi e^{-A\psi^2}d\psi}{\int\limits_0^\infty e^{-A\psi^2}d\psi} = \frac{1}{\sqrt{\pi A}} = \frac{\sqrt{\pi}mc\gamma_b}{2\sqrt{2}ePB\gamma_p}. \tag{12}$$

This angle is measured in the rest frame of the plasma. In the observer's frame,

$$\bar{\psi}_H \approx \frac{\psi}{2\gamma_p} = \frac{\sqrt{\pi}mc\gamma_b}{4\sqrt{2}ePB\gamma_p^2} \tag{13}$$

$$= \frac{mc^4\gamma_b P^2}{32\sqrt{2}\pi^{5/2}eR_*^3B_s\gamma_p^2\sin^3\beta} = 1.9 \times 10^3\frac{\gamma_b P^2}{\gamma_p^2 B_0\sin^2\beta}.$$

The Lorentz factor of the electrons in the beam is $\gamma_b = 10^6$–10^7. The value of γ_p depends substantially on the magnetic-field structure near the neutron-star surface in the region of formation of the secondary plasma. In particular, $\gamma_p \sim 10^3$ for a dipolar magnetic field and $\gamma_p = 2$–10 for a multipolar structure with $\rho \sim R_* \sim 10^6$ cm [14]. In the dipolar case, we obtain for a pulsar with $P \sim 1$ s and $B_0 \sim 10^{12}$ G

$$\bar{\psi}_H \approx 2 \cdot (10^{-9}\text{–}10^{-8})\sin^{-2}\beta, \tag{14}$$

whereas in the multipolar case,

$$\bar{\psi}_H \approx (2 \cdot 10^{-5}\text{–}5 \cdot 10^{-3})\sin^{-2}\beta. \tag{15}$$

We shall assume below that the magnetic-field structure just near the surface differs considerably from a dipolar field and that $\gamma_p \leq 10$. Then, as follows from (15), $\gamma_b\psi_H \gg 1$ and we can use well-known formulas of synchrotron radiation (for example, from [15]).

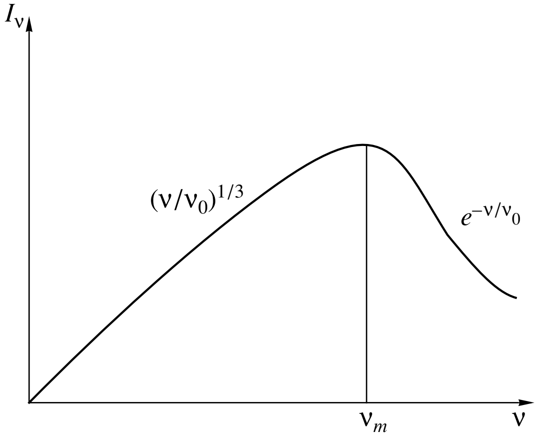

Fig. 2. The spectrum of synchrotron radiation by a single electron.

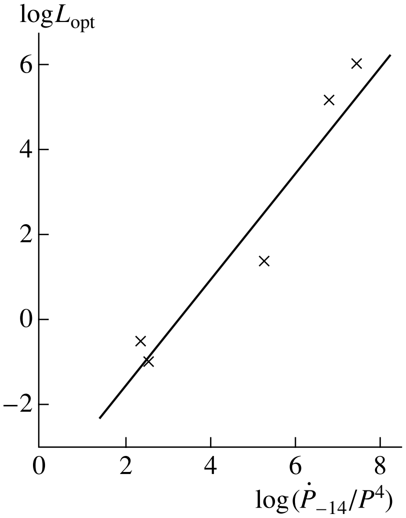

Fig. 3. Observed relation between the optical luminosities of radio pulsars $L_{\rm opt}$ (in arbitrary units, see text) and $\dot P/P^4$ (where the period P is in seconds).

4. SYNCHROTRON RADIATION BY THE PRIMARY ELECTRONS

The spectrum of synchrotron radiation by relativistic electrons depends substantially on their energy distribution. In the case of the primary beam, this distribution can be taken to be monoenergetic [7]:

$$N(\gamma) = N_0\delta(\gamma - \gamma_b). \qquad (16)$$

The synchrotron spectrum for this distribution will be similar to the radiation spectrum for a single electron (Fig. 2), with a maximum at the frequency

$$\nu_m \approx 0.29\nu_0, \qquad (17)$$

where $\nu_0 = \frac{3}{4\pi}\frac{eB\sin\psi}{mc}\gamma_b^2$. At $\psi \ll 1$, we obtain from (13)

$$\nu_m = \frac{0.87\gamma_b^3}{16\sqrt{2\pi}P\gamma_p^2}. \qquad (18)$$

Taking the Lorentz factors $\gamma_b = 10^6$ and $\gamma_p = 10$, we obtain for five radio pulsars with pulsed optical emission [16] peak frequencies in the range from 5.6×10^{14} Hz to 6.6×10^{15} Hz (Table 1).

The peak for PSR 0656+14 corresponds to the optical, as was already noted in [8]. Since the intensity decreases with decreasing frequency as $\nu^{1/3}$ in the range $\nu < \nu_m$, then, if the emission at the peak is fairly strong, this emission should also be detectable at optical frequencies for all the pulsars listed in Table 1. The ν_m values for pulsars with $P \sim 1$ s correspond to the infrared range. Note that, if $\gamma_b = 10^7$, the peak frequency increases by three orders of magnitude; however, the intensity in the optical for this value of γ_b will be only an order of magnitude less than at the maximum. Moreover, if we take the minimum value $\gamma_p = 2$, we obtain an additional increase in ν_m by a factor of 25. The maximum values of ν_m are indicated in column 6 in Table 1. These frequencies correspond to peaks at gamma-ray energies, but the intensity in the optical could still be appreciable.

The magnetic-field magnitudes in [17] were calculated using the relation

$$|I\Omega\dot\Omega| = \frac{2m_\perp^2\Omega^4}{3c^3} \qquad (19)$$

for the magnetic moment $m = BR_*^3$. In (19), Ω is the angular velocity of the neutron star. However, in fact (see, for example, [18]),

$$m = BR_*^3/2; \qquad (20)$$

thus, the calculated magnetic field is decreased by a factor of two. Precisely these corrected values for the field are presented for $\log B_0$ in the right-hand subcolumn of column 5 in Table 1.

Let us now estimate the total synchrotron luminosity of the electrons in the beam. Using (3) and (13), we obtain

$$L = \frac{\pi^3 e^2 I\gamma_b^3\dot P}{128mc^3P^4\gamma_p^4} = 2.27\times10^7\frac{\gamma_b^3\dot P_{-14}}{\gamma_p^4P^4}, \qquad (21)$$

where we have adopted $I = 10^{45}$ g cm^2 and $\dot P_{-14} = \dot P/10^{-14}$. The luminosities that can be obtained in the model are presented in column 7 of Table 1. The left-hand subcolumn corresponds to Lorentz factors $\gamma_b = 10^6$ and $\gamma_p = 10$ and the right-hand subcolumn corresponds to $\gamma_b = 10^7$ and $\gamma_p = 2$. The calculated values of L are similar to the observed luminosities [16]; thus, the theoretical and observational values

Table 1.

No.	PSR	P, ms	$\dot{P}_{-14}$	$\log B_0$ (G)		ν_m, Hz	$\log L$ (erg/s)		$L_{\rm opt}$
1	2	3	4	5		6	7		8
1	B 0531+21	33	42	12.68	12.88	6.58×10^{15}–1.64×10^{20}	28.91	34.70	10^6
2	B 0540-69	50	48	12.70	13.00	4.34×10^{15}–1.08×10^{20}	28.24	34.03	1.4×10^5
3	J 0633+1746	237	1.1	12.21	12.51	9.16×10^{14}–2.29×10^{19}	23.90	29.69	0.1
4	B 0656+14	385	5.5	12.67	12.97	5.64×10^{14}–1.41×10^{19}	23.75	29.55	0.3
5	B 0833-45	89	12	12.53	12.83	2.44×10^{15}–6.10×10^{19}	26.64	32.43	21

$L_{\rm opt}$ is expressed in arbitrary units, such that the luminosity of PSR B0531+21 is equal to 10^6.

can be brought into agreement through appropriate choice of γ_p and γ_b.

Note that the magnitude of the magnetic field B does not appear in (21). This fortunate circumstance enables us to avoid having to consider the field structure in outer regions of the pulsar magnetosphere and to use for the estimates only the observed parameters P and $\dot{P}$ together with the values of γ_b and γ_p, which are nearly constant over the magnetosphere.

Shearer *et al.* [16] found a strong correlation between the optical luminosity and the magnetic field at the light cylinder. This is surely a strong (and, probably, decisive) argument in favor of the idea that the optical emission is generated at the magnetosphere periphery. In this case, $L \propto (\dot{P}^{1/2}/P^{5/2})^3$. On the other hand, (21) leads to a slightly different functional dependence: $L \propto \dot{P}/P^4$. If we use the data of [16], in which the optical luminosities are presented in units of 10^{-6} of the luminosity of PSR 0531+21 (Table 1), we obtain the following relation between the observed values of $L_{\rm opt}$ and the ratio $\dot{P}/P^4$ (Fig. 3):

$$\log L_{\rm opt} = (1.30 \pm 0.19) \log \left(\dot{P}_{-14}/P^4\right) - 4.21 \pm 1.02, \qquad (22)$$

with the correlation coefficient being $\rho = 0.97 \pm 0.14$. This strong correlation shows that our model can describe the observed optical emission of radio pulsars. Using (21) and (22) together with the optical luminosity of PSR 0531+21 $\log L$ (erg/s) = 34.1, we obtain the following mean value for the five pulsars:

$$\gamma_b^3/\gamma_p^4 = 3.42 \times 10^{16}. \qquad (23)$$

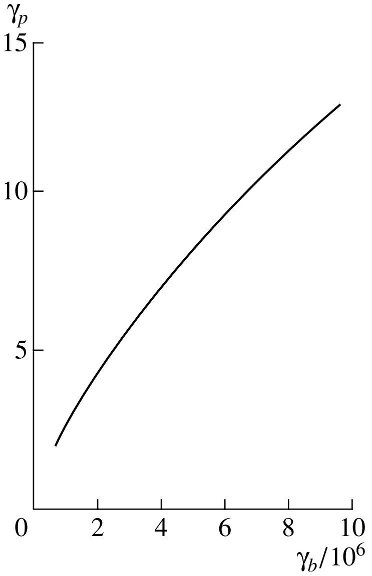

Fig. 4. Dependence of the Lorentz factor of the secondary plasma γ_p on the Lorentz factor of the beam γ_b from relation (23).

The corresponding γ_p and γ_b values are plotted in Fig. 4, where γ_b varies from 10^6 to 10^7. As follows from this dependence, γ_p takes on values from 2.3 to 13.1 for the above range of γ_b. In order for secondary plasma with these Lorentz factors to be formed, the magnetic field near the neutron-star surface must differ substantially from a dipolar field. We are in agreement with Shearer *et al.* [16] that the optical emission is formed at the periphery of the pulsar magnetosphere; the agreement between (21) and the observational data supports this point of view. We emphasize that (22) was derived using only five objects; therefore, more reliable conclusions should be obtained if the number of known pulsars with optical emission can be increased.

However, we can already predict that a radio pulsar will produce appreciable optical radiation if $\dot{P}_{-14}/P^4$ is sufficiently large. This parameter takes on values from 250 to 3.55×10^7 for the five objects in Table 1. The mean value $\log(\dot{P}_{-14}/P^4) = 3.90$ for 38 pulsars with $P < 0.1$ s from [17] is in the same

Table 2.

No.	PSR	$\dot{P}_{-14}/P^4$	$\log L^*$ (erg/s)	d, kpc	No.	PSR	$\dot{P}_{-14}/P^4$	$\log L^*$ (erg/s)	d, kpc
1	2	3	4	5	1	2	3	4	5
1	B 0114+58	5.62×10^3	27.64	2.14	25	1634-45	1579	27.09	4.84
2	0138+57	195	26.18	2.89	26	1643-43	3894	27.48	6.85
3	0355+54	742	26.76	2.07	27	1706-44*	8.59×10^4	28.82	1.82
4	0456-69	315	26.39	49.40	28	1715-40	148	26.06	6.29
5	J 0538+2817	876	26.83	1.77	29	1718-35	410	26.50	6.36
6	0540+23	421	26.51	2.54	30	1719-37	350	26.43	2.51
7	0611+22	479	26.57	4.72	31	1727-33*	2.28×10^4	28.25	4.24
8	J 0631+1036	1520	27.07	6.56	32	1737-30	343	26.42	3.28
9	0740-28	2160	27.22	1.89	33	1754-24	434	26.53	3.50
10	0906-49	1.16×10^4	27.95	6.57	34	1757-24*	5.24×10^4	28.61	4.61
11	1046-58*	4.06×10^4	28.49	2.98	35	1800-21*	4.22×10^4	28.51	3.94
12	1055-52	387	26.48	1.53	36	1821-19	409	26.50	5.19
13	1221-63	230	26.25	2.29	37	1822-14	374	26.46	5.42
14	1317-53	150	26.07	4.23	38	1823-13*	7.08×10^4	28.74	4.12
15	1338-62*	1.82×10^4	28.15	8.66	39	1828-10	223	26.24	3.63
16	1356-60	2399	27.27	5.91	40	1832-06	463	26.56	6.34
17	1449-64	268	26.32	1.84	41	1838-04	534	26.62	5.16
18	1508-57	2474	27.28	12.69	42	1841-05	226	26.24	6.16
19	1509-58*	2.96×10^5	29.36	4.40	43	1842-04	548	26.63	4.72
20	1535-56	139	26.03	4.01	44	1853+01	4101	27.50	3.30
21	1556-57	368	26.46	6.30	45	1855+02	134	26.02	8.58
22	1557-50	148	26.06	5.34	46	1915+13	502	26.59	4.07
23	1607-52	474	26.57	3.34	47	2011+38	316	26.39	13.09
24	1610-50*	1.76×10^4	28.14	7.26	48	2334+61	170	26.12	2.47

The period P in column 3 is in seconds. The asterisks denote pulsars whose optical luminosity exceeds 10^{28} erg/s.

interval.[1] Therefore, if the distance to a pulsar is not extremely large, it is worthwhile to search for optical emission from this object, since its mean luminosity is of the order of 10^{28} erg/s. The mean value for most of the pulsars from [17] is $\log(\dot{P}_{-14}/P^4) \approx 0.14$ (i.e., $\dot{P}_{-14}/P^4 \approx 1$) for 497 sources with $P > 0.1$ s and measured values of $\dot{P}$. According to (21), their optical luminosity should be extremely small. Nevertheless, for several objects (Table 2), their values of $\dot{P}_{-14}/P^4$ could result in high optical luminosities L^*, as calculated for γ_b^3/γ_p^4 from (23). Pulsars with $\log L^* > 26$ are listed in Table 2. This luminosity was chosen arbitrarily based on the lower integrated radio luminosity of the pulsars [20]. We emphasize that it will be quite difficult to detect optical emission from most of these objects, due to their considerable distances from the Earth. Nevertheless, the pulsars with luminosities $> 10^{28}$ erg/s, marked by asterisks in Table 2, could be appreciable optical emitters. Indeed, optical flux was detected from two pulsars not listed in Table 1, PSR 1055-52 and PSR 1509-58 [19]. PSR 1509-58 possesses the maximum luminosity among the sources from Table 2, while PSR 1055-52 is the most

---

[1] The total number of pulsars with $P < 0.1$ s in this catalog is 41. When calculating the mean value $\log(\dot{P}_{-14}/P^4)$, we excluded the three pulsars that were already included in Table. 1.

Table 3.

No.	PSR	$\dot{P}_{-14}/P^4$	$\log L^*$ (erg/s)	d, kpc
1	0823+26	2.15	24.22	0.38
2	0950+08	5.61	24.64	0.12
3	1929+10	43.69	25.30	0.17
4	J0454+5543	17.50	25.13	0.79
5	0906-17	2.57	24.30	0.63
6	1133+16	46.20	25.55	0.27
7	1451-68	2.07	24.21	0.45
8	J1908+0734	40.84	25.50	0.58
9	J1918+1541	13.42	25.02	0.68

nearby. PSR 1706-44, which has a high expected luminosity and a fairly small distance, was detected as an X-ray and gamma-ray source. The peak of its synchrotron spectrum probably occurs in a band that is harder than the optical.

Another argument in favor of the proposed model is the fact that six of the nine radio pulsars marked by asterisks in Table 2 also possess emission at X-ray or gamma-ray energies. Moreover, PSR 0355+54, PSR 2334+61, and PSR J0538+28 were identified in [19] as hard emitters. This may be explained by their rather small distances. As is noted above, optical (and probably harder) emission should be expected from pulsars with short periods ($P < 0.1$ s). X-ray and gamma-ray emission from nine such pulsars—1951+32 ($P = 40$ ms), 1259−63 (48 ms), J0437−47 (6 ms), 1821−24 (3 ms), J2124−33 (5 ms), 1957+20 (1.6 ms), J1012+53 (5 ms), J0218+42 (2 ms), and J0751+18 (3.5 ms)—was reported in [19].

There is soft X-ray emission from three other pulsars with $\log L^* = 24.22 - 25.30$ [19] (Table 3). These objects are very nearby and, therefore, may have detectable fluxes. In addition, Table 3 contains pulsars with $d < 1$ kpc and $\dot{P}_{-14}/P^4 > 1$, which could also be used for searching for pulsed nonthermal radiation.

It follows from the above considerations that the proposed synchrotron model can explain the available observations of radio pulsars in hard bands and predict for which sources the probability of detecting hard emission is high.

5. CONCLUSIONS

(1) We have derived a refined formula for the synchrotron power emitted by the primary-beam particles in radio pulsars when the emission is generated at the magnetosphere periphery. The corresponding power L depends on the pulsar period P and its derivative $\dot{P}$, as well as on the Lorentz factors of the beam γ_b and secondary plasma γ_p, and does not depend on the magnitude of the magnetic field in the region in which the radiation is generated.

(2) In accordance with this dependence, we found a strong correlation between the observed optical luminosity $L_{\rm opt}$ of pulsars and the ratio $\dot{P}/P^4$, with a correlation coefficient close to unity.

(3) The predicted and observed luminosities can be brought into agreement if $\gamma_p = 2{-}13$. In order for secondary plasma with such Lorentz factors to be formed, the magnetic-field structure near the neutron-star surface must strongly deviate from a dipolar field, as was proposed, for example, in the classic paper by Ruderman and Sutherland [21].

(4) We predict appreciable optical emission to be observed from radio pulsars with large values of $\dot{P}_{-14}/P^4$. All pulsars with $P < 0.1$ s ($\overline{\log \dot{P}_{-14}/P^4} =$ 3.90) belong to this category. In addition, we expect strongsynchrotronradiationfrom57 other pulsars (Tables 2,3) with $P > 0.1$ s ($\dot{P}_{-14}/P^4 = 139{-}2.96\times10^5$). All 27 pulsars that have been detected at optical to gamma-ray energies [19] are included in our list of expected emitters. This result demonstrates that our model describes the essence of the observed phenomenon.

(5) It follows from (18) that the shorter the period of a pulsar, the harder the energy at which the peak of its synchrotron radiation is observed. As a result, our list of possible synchrotron radio pulsars includes not only optical but also X-ray and gamma-ray objects. Among the 27 radio pulsars whose emission has been quantitatively measured outside the radio band [19], we can see a tendency of the period to decrease as the hardness of the peak radiation increases: 7 optical objects have $\bar{P} = 163$ ms, 27 X-ray emitters have $\bar{P} = 139$ ms, and 6 gamma-ray pulsars have $\bar{P} = 116$ ms. This represents further evidence in favor of the proposed model. As noted above, if the peak frequency in a pulsar spectrum is located at X-ray or even gamma-ray energies, such a pulsar can still emit appreciable optical emission, since the intensity decreases with decreasing frequency as $\nu^{1/3}$.

ACKNOWLEDGMENTS

This work was supported by the Federal Scientific Technical Program Astronomy and the Russian Foundation for Basic Research (project no. 00-02-17850). We are grateful to L. B. Potapova for help in preparation of the manuscript.

REFERENCES

1. V. L. Ginzburg, V. V. Zheleznyakov, and V. V. Zaĭtsev, Usp. Fiz. Nauk **98**, 201 (1969).
2. I. S. Shklovsky, Astrophys. J. Lett. **159**, L77 (1970).
3. F. Pacini, Astrophys. J. Lett. **163**, L17 (1971).
4. V. V. Zheleznyakov and V. E. Shaposhnikov, Astrophys. Space Sci. **18**, 141 (1972).
5. F. Pacini and M. Salvati, Astrophys. J. **274**, 369 (1983).
6. F. Pacini and M. Salvati, Astrophys. J. **321**, 447 (1987).
7. I. F. Malov, Astron. Zh. **76**, 825 (1999) [Astron. Rep. **43**, 727 (1999)].
8. I. F. Malov and G. Z. Machabeli, Astron. Zh. **76**, 788 (1999) [Astron. Rep. **43**, 691 (1999)].
9. I. F. Malov and G. Z. Machabeli, Astrophys. J. **554**, 587 (2001).
10. I. F. Malov, Astron. Zh. **75**, 281 (1998) [Astron. Rep. **42**, 246 (1998)].
11. W. M. Fawley, J. Arons, and E. T. Scharlemann, Astrophys. J. **217**, 227 (1977).
12. I. F. Malov, Astrofizika **19**, 161 (1983).
13. P. Goldreich and W. H. Julian, Astrophys. J. **157**, 869 (1969).
14. G. Z. Machabeli and V. V. Usov, Pis'ma Astron. Zh. **15**, 910 (1989) [Sov. Astron. Lett. **15**, 393 (1989)].
15. A. Pacholczyk, *Radio Astrophysics. Nonthermal Processes in Galactic and Extragalactic Sources* (Freeman, San Francisco, 1970; Mir, Moscow, 1973).
16. A. Shearer, A. Golden, and J. Beskin, Astron. Soc. Pac. Conf. Ser. **202**, 307 (2000).
17. J. H. Taylor, R. N. Manchester, A. J. Lyne, and F. Camilo, unpublished work (1995).
18. S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars: the Physics of Compact Objects* (Wiley, New York, 1983; Mir, Moscow, 1985), Vol. 2.
19. W. Becker and J. Trumper, Astron. Astrophys. **326**, 682 (1997).
20. I. F. Malov and O. I. Malov, Astron. Zh. **72**, 574 (1995) [Astron. Rep. **39**, 510 (1995)].
21. M. A. Ruderman and P. G. Sutherland, Astrophys. J. **196**, 51 (1975).

Translated by Yu. Dumin

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 872–881. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 997–1007.
Original Russian Text Copyright © 2001 by Kuznetsov, Bisikalo, Boyarchuk, Khruzina, Cherepashchuk.

Synthetic Doppler Tomograms of Gas Flows in the Binary System IP Peg

O. A. Kuznetsov[1, 2], **D. V. Bisikalo**[2], **A. A. Boyarchuk**[2],
T. S. Khruzina[3], and **A. M. Cherepashchuk**[3]

[1] *Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya pl. 4, Moscow, 125047 Russia*
[2] *Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia*
[3] *Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119899, Russia*

Received March 20, 2001

Abstract—We have synthesized Doppler tomograms of gas flows in the binary system IP Peg using the results of three-dimensional gas-dynamical computations. Gas-dynamical modeling in combination with Doppler tomography enables identification of the key elements of flows in Doppler maps without solution of an ill-posed inverse problem. A comparison of the synthetic tomograms with observations shows that, in the quiescent state of the system, the most luminous components are (1) the shock wave induced by interaction between the circumbinary envelope and the stream from the Lagrange point L_1 (the "hot line") and (2) the gas condensation at the apogee of the quasi-elliptical disk. Both the single spiral shock wave arm in the gas-dynamical solution and the stream from L_1 contribute little to the luminosity. In the active state of the system, when the stream from L_1 does not play an appreciable role and the disk dominates, both areas of enhanced luminosity in the observational tomograms are associated with the two arms of the spiral shock wave in the disk. *© 2001 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

As a rule, observations of binaries are photometric or spectrographic. The first technique provides information on the temporal dependence of the star's brightness in some waveband $I(t)$, while the second yields the temporal dependence of the wavelength for a given line $\lambda(t)$ via its Doppler shift. Given the ephemeris $\phi(t)$, the dependences $I(t)$ and $\lambda(t)$ can be transformed into a light curve $I(\phi)$ and radial-velocity curve $V_R(\phi)$.

Recently, observations of binaries in the form of consecutive (trailed) spectrograms in some emission line $I(\lambda, t)$ (or, in other words, $I(V_R, \phi)$) have become widespread. The most suitable method for the analysis of such spectrograms is Doppler tomography [1]. In this technique, the orbital variability of emission-line intensity is translated into a luminosity map in two-dimensional velocity space. A Doppler tomogram is constructed by transforming a number of spectrograms for consecutive times $I(V_R, t)$ (or, equivalently, consecutive orbital phases $I(V_R, \phi)$) into a luminosity distribution in the (V_x, V_y) plane. In order to transform the distribution $I(V_R, \phi)$ into the Doppler tomogram $I(V_x, V_y)$, the radial velocity is expressed in terms of the projection of the velocity vector onto the line of sight $V_R = -V_x \cos(2\pi\phi) + V_y \sin(2\pi\phi)$ (here, $V_z \sim 0$; the minus sign before V_x is dictated by the choice of coordinate system); the inverse problem described by the following integral equation is solved (see Appendix A in [1]):

$$I(V_R, \phi) = \int\int I(V_x, V_y) g(V_R + V_x \cos(2\pi\phi) - V_y \sin(2\pi\phi)) dV_x dV_y, \tag{1}$$

where $g(V)$ is the normalized local line profile (for example, a δ function) and the integrals are taken from $-\infty$ to $+\infty$. This inverse problem is ill-posed; its solution requires special regularization (for example, using the maximum entropy method [2, 1], the method of filtered inverse projection [3], the rapid maximum entropy method [4], etc; see also [5, 6]).

The solution of this inverse problem yields a Doppler map, i.e., the distribution of the radiation intensity at the frequencies of some emission line in velocity space. In a number of cases, Doppler maps are easier to interpret than the initial spectrograms; in addition, tomograms may indicate (or, at least, give hints of) some peculiarities in the flow structure. In particular, a line with a double-peaked profile corresponding to circular motion (for example, in the accretion disk) becomes a blurred ring in the Doppler map. In other words, binary components are resolved in velocity space, despite the fact that they cannot be spatially resolved directly in observations. This makes Doppler tomography a powerful technique for the study of binaries.

Unfortunately, the problem of reconstructing the spatial distribution for an emission-line intensity based on a Doppler map is, in general, not solvable, since points separated by large distances can have the same velocity and contribute to the same location in the Doppler map. Therefore, the transformation $I(V_x, V_y) \to I(x, y)$ is impossible without some *a priori* assumptions being made as about the structure of the velocity field.

The situation changes dramatically when gas-dynamical computations are used together with Doppler tomography. In this case, there is no need to solve the inverse problem, since the Doppler tomogram can be constructed directly: $\rho(x,y)\,\&\,T(x,y) \to I(x,y)$ and $I(x,y)$ & $V_x(x,y)$ & $V_y(x,y) \to I(V_x,V_y) \to I(V_R,\phi)$. Difficulties arise only when transforming the spatial distributions for the density and temperature $\rho(x,y)$, $T(x,y)$ into the intensity distribution for some emission line $I(x,y)$. The profiles of optically thick lines should be calculated using the radiative transfer equation (see, for example, [7]). Here, when deriving the synthetic Doppler tomograms, we assumed that the medium was optically thin and the intensity of the recombination line considered was $I \sim \rho^2 T^{1/2}$ [8, 9].

2. PARAMETERS OF THE SYSTEM

The variable star IP Peg was discovered by Lipovetskiĭ and Stepanian [10] in 1981. In 1985, Goranskij *et al.* [11] showed that IP Peg was an eclipsing dwarf Nova (orbital period $3^{\mathrm{h}}.79$) with a deep eclipse and a peak in its light curve.

The binary is fully determined by a set of parameters consisting of the orbital period P, masses of the components M_1 and M_2, distance between the centers of the components A, and the orbital inclination i. Of all these parameters, only P can be determined directly from observation, while M_1, M_2, A, and i (or alternatively M, q, A, and i, where $M = M_1 + M_2$ is the total mass of the system and $q = M_2/M_1$ is the mass ratio of the components) must be derived from observational data. These four parameters are related via Kepler's third law, $\Omega^2 A^3 = GM$ ($\Omega = 2\pi/P$ is the orbital angular velocity and G is the gravitational constant); therefore, another three independent relationships must be found if they are all to be determined.

Several methods for deriving relationships between the parameters of a binary from observations are known.

(a) We can derive $q = K_1/K_2$ from the radial-velocity amplitudes

$$K_1 = A\Omega\sin i\frac{M_2}{M} = A\Omega\sin i\frac{q}{1+q},$$
$$K_2 = A\Omega\sin i\frac{M_1}{M} = A\Omega\sin i\frac{1}{1+q}$$

for the accretor and donor, and, taking into account the angular velocity of rotation for the binary previously determined from observations, establish the relationship between A and i.

(b) We can find the relationship between q and i from the phase width $\Delta\phi$ of the eclipse of the white dwarf by the donor star (in eclipsing systems).

Determining K_1, K_2, and $\Delta\phi$ from observations enables calculation of the desired binary parameters. Other methods for establishing relationships between binary parameters exist. For example, the velocity $V_{\mathrm{rot}} = R_{RL}(q, A)\Omega\sin i$ corresponding to rotational broadening of the absorption lines of the donor star (R_{RL} is the effective radius of the Roche lobe) can be used to establish the relationship between q, A, and i. The desired parameters can also be refined using supplementary relations.[1]

The half-amplitude of the radial velocity of the white dwarf in the IP Peg system was determined in [13–15, 19]: Marsh [15] obtained $K_1 = 175 \pm 15$ km/s; Shafter [14], $K_1 = 164$ km/s; and Wood and Crawford [13], $K_1 = 141$ km/s. The half-amplitude of the radial velocity of the donor was determined in [15–17, 18]: Marsh [15] obtained $K_2 = 305 \pm 15$ km/s; Martin *et al.* [16, 17], $K_2 = 288$–298 km/s; and Beekman *et al.* [18], $K_2 = 331.3 \pm 5.8$ km/s. Here, we adopt the radial-velocity half-amplitudes from [13, 19]: $K_1 = 148$ km/s and $K_2 = 301$ km/s.

Having fixed the values for K_1 and K_2, let us now use the second relationship between the desired parameters. For this purpose, we will use the width of the eclipse of the white-dwarf by the donor star, which has been measured fairly accurately: Wood and Crawford [13] obtained $\Delta\phi = 0^{\mathrm{P}}.0863$ and Marsh [15] obtained $\Delta\phi = 0^{\mathrm{P}}.0858$. The width of the eclipse of the white-dwarf by the donor yields a relationship between the orbital inclination i and mass ratio q in the form (see, for example, [20, 21])

$$\left(\frac{A\tan(\pi\Delta\phi)\sin i}{Y_{RL}}\right)^2 + \left(\frac{A\cos i}{Z_{RL}}\right)^2 = 1. \quad (2)$$

---

[1] Here, we do not consider methods for determining binary parameters based on eclipse of the hot spot [12, 13], since we consider a substantially different model for the flow; we also do not consider methods based on the radius–mass dependence for the donor star, since this is derived for single stars.

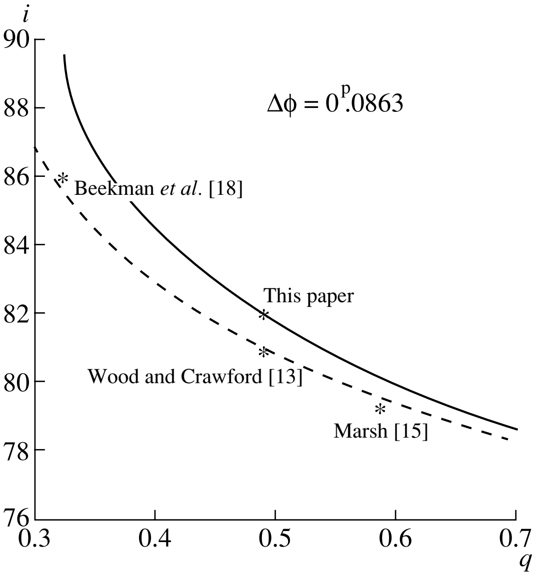

Fig. 1. Relationship between the orbital inclination i and the mass ratio of the binary components $q = M_2/M_1$ for the width of the eclipse of the white-dwarf by the donor star $\Delta\phi = 0\overset{p}{.}0863$. The dashed curve represents calculations of the Roche-lobe radius according to (3). The solid curve corresponds to exact calculations of the Roche-lobe size in the Y and Z directions. We also show i and q obtained in [13, 15, 18], as well as the values adopted in our computations: $i = 82\overset{\circ}{.}0$ and $q = 0.49$.

Here, $Z_{RL}(q)$ and $Y_{RL}(q)$ are the dimensions of the donor Roche lobe in the Z (normal to the orbital plane) and Y (along the orbital motion of the donor) directions. The functions $Z_{RL}(q)$ and $Y_{RL}(q)$ are usually presented in an approximate form [22]:

$$Z_{RL}(q)/A = Y_{RL}(q)/A \tag{3}$$

$$= R_{RL}/A \approx 0.462\left(\frac{q}{1+q}\right)^{1/3}.$$

Figure 1 presents the derived dependence between i and q for the adopted value of $\Delta\phi = 0\overset{p}{.}0863$ (dashed curve). However, using the approximate formula (3) can result in some loss of accuracy. Therefore, we calculated the relations $Z_{RL}(q)$ and $Y_{RL}(q)$ in the interval $0.3 \le q \le 0.7$ exactly and approximated them with errors $< 0.1\%$ using the formulas

$$Y_{RL}(q)/A \approx -0.14858q^2 + 0.32031q + 0.18991,$$

$$Z_{RL}(q)/A \approx -0.14466q^2 + 0.30559q + 0.18320.$$

The derived dependence $i(q)$ is also presented in Fig. 1 (solid curve). We also show the i and q values adopted by Wood and Crawford [13], Marsh [15], and Beekman *et al.* [18].

Using the above K_1 and K_2 values and the refined dependence $i(q)$, we will adopt for further computations the following parameters for the IP Peg system: $q = 0.49$, $M_1 = 1.02M_\odot$, $M_2 = 0.5M_\odot$, and $i = 82\overset{\circ}{.}0$. The inner Lagrange point L_1 is located at a distance $D = 0.573A = 0.812R_\odot$, while the center of mass of the system is at a distance $0.329A = 0.466R_\odot$ from the accretor.

The adopted parameters imply a velocity for the rotational broadening of the donor's absorption lines of $V_{\rm rot} = 143$ km/s. Observational values for $V_{\rm rot}$ are $V_{\rm rot} = 146$ km/s [23] and $V_{\rm rot} = 125$ km/s [24]. The consistency between the computed and observed parameters in the supplementary relation suggests that the values we have for the parameters of IP Peg are correct.

Note also two factors that were not considered in our model. There have been numerous suggestions [16, 18] that the orbits of the components may deviate from being circular (with eccentricities $e = 0.05$–0.075). In addition, some observational peculiarities of IP Peg have been interpreted as evidence of the presence of a third body in the system [25, 26]: a star with mass 0.05–$0.1M_\odot$. Here, we assume that the system contains only two stellar components rotating in circular orbits.

3. THE MODEL

3.1. Gas-Dynamical Model

A complete description of the three-dimensional gas-dynamical model is presented in [27], and we note only key elements of the model here. We used three-dimensional Euler equations in Cartesian coordinates to describe the gas flows in the binary system. This system of gas-dynamical equations was closed with the equation of state for an ideal gas with adiabatic index γ. To take into account radiation loss in the system, we assumed that γ was close to unity ($\gamma = 1.01$), which corresponds to a case that is close to isothermal [28, 29].

The computations were carried out in a noninertial reference frame rotating with the binary. The results of the gas-dynamical computations and corresponding Doppler tomograms are presented in a coordinate system in which the origin is at the center of the accretor, the X axis is directed from the accretor to the donor, the Y axis is directed along the orbital motion of the donor, and the Z axis is normal to the equatorial plane and oriented such that the coordinate system is right-handed (upper panel in Fig. 2). Figure 2 also indicates the phase angles of the observer, the accretor, the donor, the Roche lobe of the accretor, and the ballistic trajectory from the point L_1. The lower panel in Fig. 2 presents the coordinate system used for the Doppler tomograms. It is very easy to transform the donor shape from spatial coordinates to coordinates in velocity space, since the donor star is fixed in the rotating coordinate system. A point $\mathbf{r}$ that is fixed in the rotating coordinate system moves with velocity $\mathbf{\Omega} \times \mathbf{r}$ in the laboratory system. Since this expression is linear in the distance to the rotational axis, the projection of the donor's shape onto the orbital plane is preserved in velocity coordinates; since the velocity vector for each point of the donor is normal to the radius vector, all points of the donor rotate by 90° counterclockwise (shaded areas in the upper and lower panels of Fig. 2). The location of the accretor in velocity space is specified by the point $(0, -K_1)$. Figure 2 also presents concentric circles corresponding to various disk radii, as well as their images in the velocity plane (for Keplerian rotation). The inner circle has larger velocities and forms the outer circle on the Doppler tomogram.

We used a higher-order Roe–Osher difference scheme [30, 31] when deriving the numerical solution of the gas-dynamical equations. The computation domain had the form of a parallelepiped $[-D \ldots D] \times [-D \ldots D] \times [0 \ldots {}^1/_2 D]$; due to the symmetry of the problem, the computations were carried out only in the upper half-space. We cut a sphere of radius ${}^1/_{100}D$ from the computation domain to represent the accretor. We adopted free-outflow conditions for the boundary conditions at the accretor and outer boundary. Matter with $\rho = \rho(L_1)$, $V_x = c(L_1)$, and $V_y = V_z = 0$ was injected into the system at the grid node corresponding to L_1, where $c(L_1)$ is the sound speed at the donor-star surface. Since the system of equations is scalable in ρ and P, the value for $\rho(L_1)$ can be chosen arbitrarily. We adopted the value $\rho(L_1) = 1$ and took the sound velocity at the inner Lagrange point to be 5.5 km/s, which roughly corresponds to $T(L_1) = 3500$ K. We adopted rarified gas with $\rho_0 = 10^{-5} \cdot \rho(L_1)$, $P_0 = 10^{-4}\rho(L_1)c^2(L_1)/\gamma$, and $\mathbf{V}_0 = 0$ for the initial conditions.

3.2. Method for Constructing the Synthetic Doppler Tomograms

Before considering the synthetic Doppler maps for IP Peg, we should emphasize the complexity of analyzing Doppler tomograms for eclipsing systems (see, for example, [32]). Synthetic Doppler maps imply that each point of a geometrical area contributes to a corresponding point in the velocity plane and that this point is visible at any orbital phase. Indeed, it follows from (1) that

$$I(V_R, \phi + \pi) = I(-V_R, \phi)$$

and Doppler maps can be constructed only for sets of consecutive spectrograms $I(V_R, \phi)$ that can be mapped symmetrically when the system is viewed "from the opposite side," i.e., when there are no eclipses and no elements of the binary are obscured by others. It is obvious that eclipsing systems are inconsistent with this assumption. When a Doppler map is constructed from observations, eclipsed sections of the consecutive spectrograms are simply excluded from the input data set. The resulting Doppler map corresponds to the case without eclipses (in other words, to a "transparent" donor star). In our case, our transformation of the results of the gas-dynamical

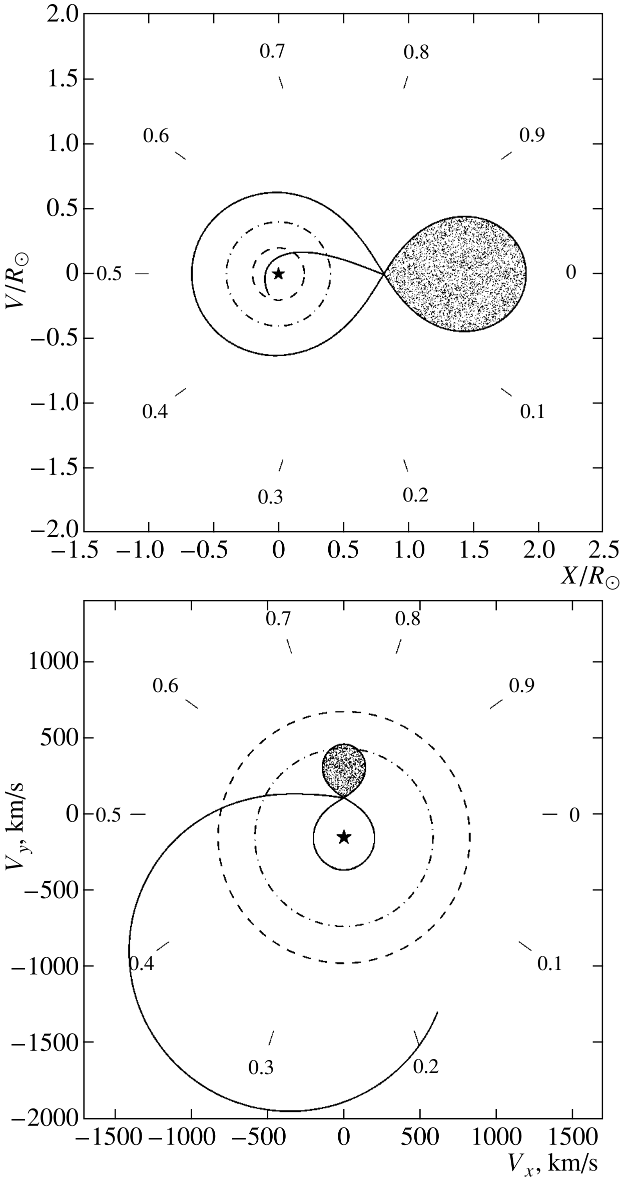

Fig. 2. Top: the coordinate system used with the phase angles of the observer. The orbital rotation of the binary is counterclockwise. The star denotes the accretor. The filled Roche lobe of the donor star is shaded. The Roche lobe of the accretor and the ballistic trajectory from L_1 are marked with solid curves. The dashed and dot–dash curves indicate concentric circles corresponding to various disk radii. Bottom: the coordinate system in velocity space. The notation is the same.

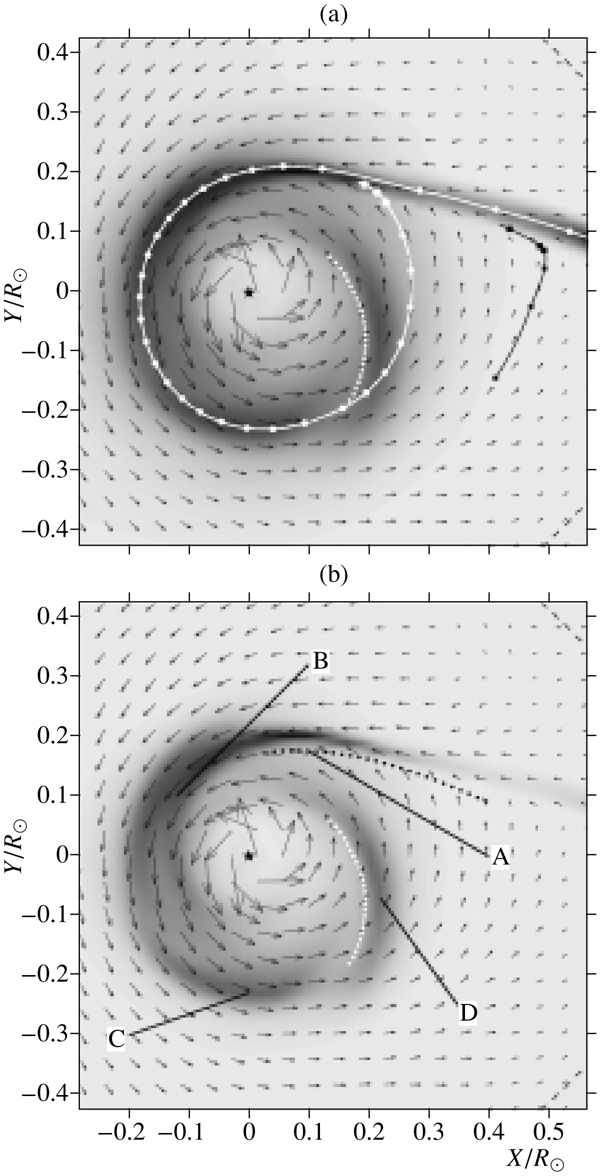

Fig. 3. (a) Density distribution in the equatorial plane of the system. The arrows denote velocity vectors in the laboratory coordinate system. The star marks the accretor. The dashed curve indicates the accretor's Roche lobe. The white dotted curve represents a tidal spiral shock. A gas-dynamical trajectory of a particle ejected from L_1 is shown by the white curve with circles. Another gas-dynamical trajectory passing through the shock at the edge of the stream is marked by the black curve with squares (see also Fig. 4). (b) Distribution of $\rho^2 T^{1/2}$ in the equatorial plane of the system. The black dotted curve indicates the shock wave along the edge of the stream (the hot line). A, B, C, and D denote zones with the highest intensity of radiation. The remaining notation is the same.

computations assumes that all data have been used, which also corresponds to a transparent donor.

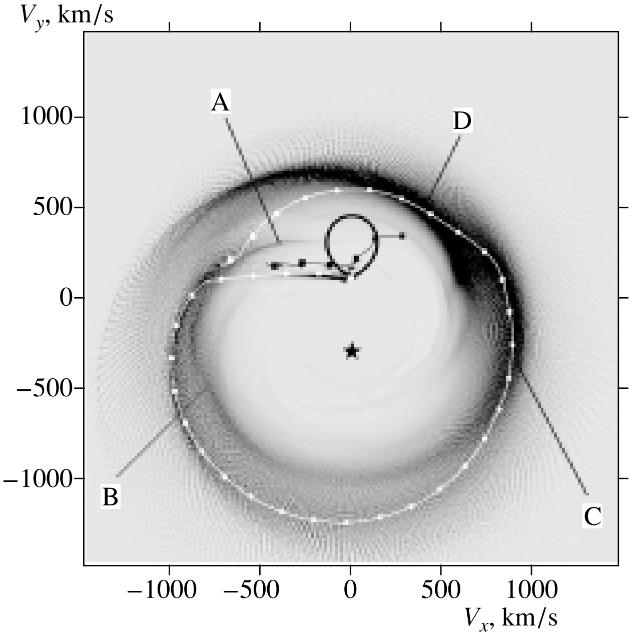

Fig. 4. Synthetic Doppler tomogram for $I \sim \rho^2 T^{1/2}$. The Roche lobe of the donor star (thick black curve) and the accretor (star) are indicated. The white curve with circles and black curve with squares display gas-dynamical trajectories in velocity space (see also Fig. 3a). A, B, C, and D denote the same zones of maximum radiation intensity as in Fig. 3b.

Doppler tomograms display luminosity in the velocity coordinate plane. Let us denote the velocity vector field in the laboratory coordinate system (the observer's frame) $\mathbf{U} = (U_x, U_y, U_z)$. When the observer is situated in the orbital plane of the binary, the coordinates of the Doppler tomogram (V_x, V_y) coincide with U_x and U_y. When the system is inclined, these coordinates are specified using the projection $\mathbf{U}$ onto the plane formed by the vectors $\mathbf{n}$ and $\mathbf{n} \times \mathbf{\Omega}$, where $\mathbf{n}$ is the direction from the observer to the binary.

The intensity of a line in velocity space is calculated as follows:

$$I(V_x, V_y) \sim \int\limits_{\mathcal{V}} \int\limits_{U_x} \int\limits_{U_y} I(x,y,z)\delta(U_x(x,y,z) \tag{4}$$

$$\times \sin i + U_z(x,y,z)\cos i - V_x)$$

$$\times \delta(U_y(x,y,z)\sin i + U_z(x,y,z)\cos i - V_y) d\mathcal{V} dU_x dU_y,$$

where $d\mathcal{V} = dxdydz$ and i is the orbital inclination.

As was noted above, when constructing the synthetic Doppler tomograms, the intensity was expressed in the form $I \sim \rho^2 T^{1/2}$.

4. RESULTS FOR THE QUIESCENT STATE OF IP Peg

To carry out three-dimensional gas-dynamical modeling for IP Peg in its quiescent state, we used

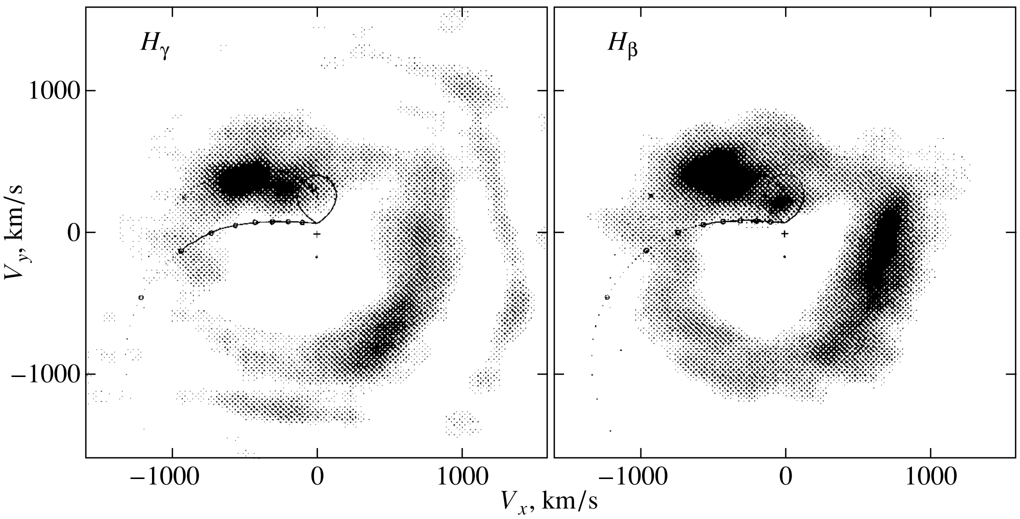

Fig. 5. Doppler map of IP Peg in H_γ and H_β for the quiescent state of the system [40]. Figure reproduced by the courtesy of S. Wolf and A. Bobinger.

the model presented in Section 3.1, continuing the computations until a steady-state flow regime was attained. Figure 3a presents the morphology of the gas flows in the binary together with distributions of the density and velocity vectors in the equatorial plane of the system (the XY plane). We also can see in Fig. 3a the gas-dynamical trajectory of a particle ejected from L_1 (white curve with circles) and a gas-dynamical trajectory passing through the shock wave at the edge of the stream from L_1 (black curve with squares, see also Fig. 4 below). These results, as well as our previous studies [33, 34], indicate that the rarified gas of the circumbinary envelope substantially affects the structure of gas flows in semidetached binaries. The circumbinary envelope interacts with and deflects the stream from L_1. In particular, this results in a shockless interaction between the stream and the accretion disk, and, consequently, in the absence of a hot spot at the edge of the disk. At the same time, the interaction between the circumbinary envelope and the stream forms an extended shock along the edge of the stream (a hot line). The hot line model is consistent with observations [35, 36]. In addition, Fig. 3a displays a spiral shock (white dotted curve), which originates in the accretion disk due to the tidal effect of the donor star. The existence of double-arm spiral shocks was discovered in [28, 37, 38]. In our computations, only a single-arm spiral shock is visible, whereas the flow structure in the area where the second arm should be located is specified by the stream from L_1, which apparently inhibits the formation of the second arm of the tidal spiral shock. The flow structure beyond the equatorial plane indicates that some of the circumbinary envelope interacts with the flow, passing around it from above and below, resulting in the formation of a "halo." It is suggested in [27] that this halo is formed by matter that (1) rotates around the accretor (is gravitationally bound with it), (2) does not belong to the disk, (3) interacts with the stream (collides or bypasses), and (4) is either accreted or leaves the system after the interaction.

Figure 3b presents the $\rho^2 T^{1/2}$ distribution in the equatorial plane of the system. As in Fig. 3a, the white dotted curve in Fig. 3b indicates the spiral shock wave. The black dotted curve displays the shock along the edge of the stream, the hot line. Since the quantity displayed in Fig. 3b is proportional to the intensity of a recombination line, its distribution can be used to identify the features of the flow that make the largest contributions to the radiation of the system. We can see that four areas of the flow, denoted A, B, C, and D, radiate the most. The letter A denotes the shock along the edge of the stream (hot line) due to the collision of the circumbinary envelope gas with the stream. The letter B denotes the stream from L_1 or, more precisely, the part of this stream with maximum I, where the gas density is still fairly high and the temperature increases appreciably due to dissipation. The letter C marks the area of the disk near the apogee. We can see that the disk has a quasi-elliptical shape, so that a condensation forms in the disk during the approach toward the apogee, when the velocity of the matter decreases, resulting in a zone of enhanced radiation. The letter D denotes a condensation behind the spiral shock.

Figure 4 presents a synthetic Doppler tomogram constructed from our gas-dynamical computations. The basic features of the flow structure were analyzed in the equatorial plane of the system (Figs. 3a, 3b);

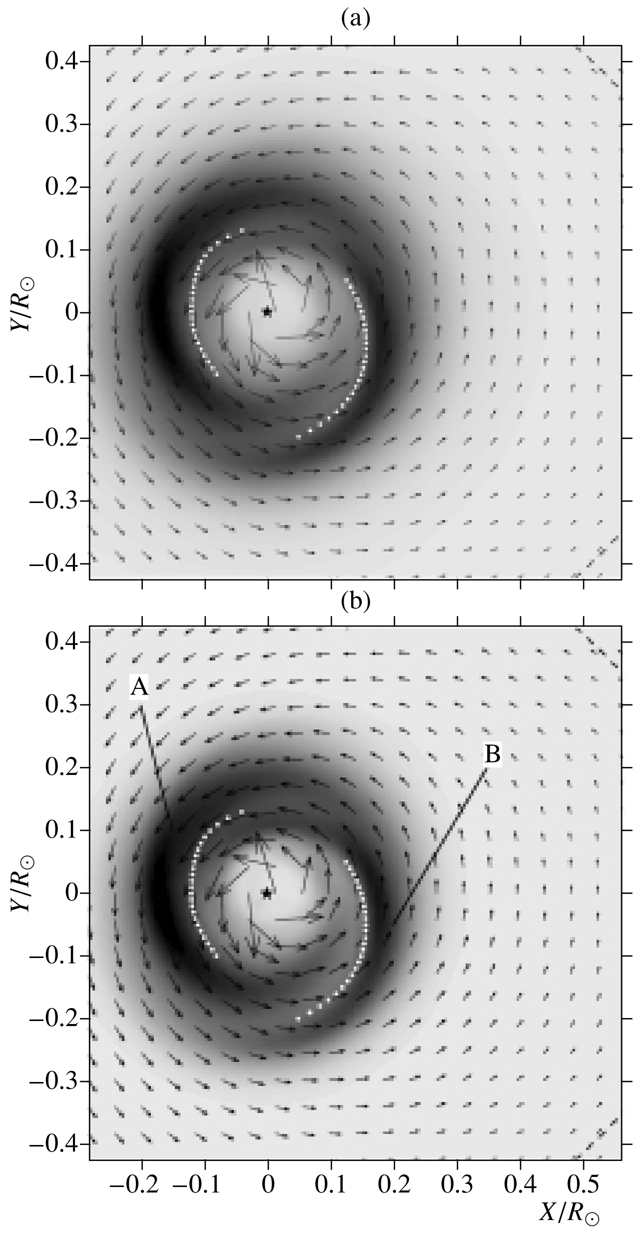

Fig. 6. (a) Density distribution in the equatorial plane of the system in its active state. The arrows denote velocity vectors in the laboratory coordinate system. The star marks the accretor. The dashed curve indicates the Roche lobe of the accretor and the white dotted curve the arm of the tidal spiral shock. (b) Distribution of $\rho^2T^{1/2}$ in the equatorial plane of the system in its active state. A and B denote zones with a maximum intensity of radiation. The remaining notation is the same.

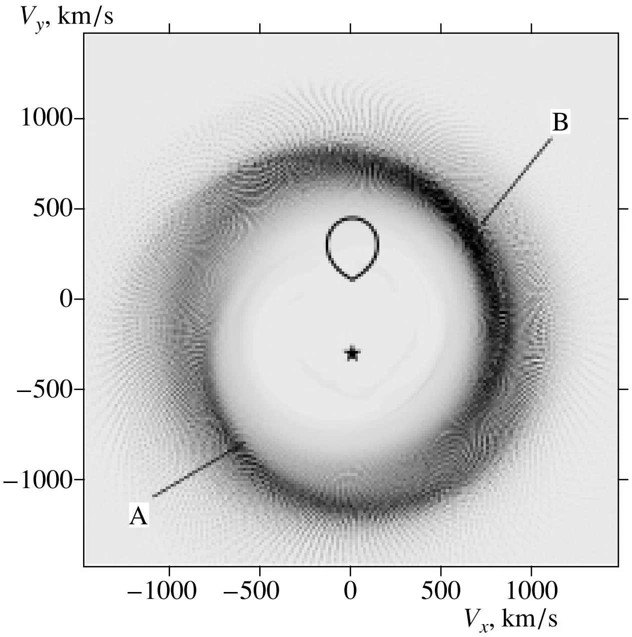

Fig. 7. Synthetic Doppler tomogram for the system in its active state. The Roche lobe of the donor star (thick black curve) and the accretor (star) are shown. A and B denote the same zones with maximum radiation as in Fig. 6b.

however, for comparison with observations, the intensity was integrated along the z coordinate in accordance with (4); i.e., the tomogram takes into account all z layers of the computation grid. Figure 4 also displays two gas-dynamical trajectories (in velocity coordinates) corresponding to the gas-dynamical trajectories in Fig. 3a.

The shock resulting from the interaction between the circumbinary envelope and the flow from L_1 is located along the edge of the stream (in spatial coordinates). The three latest, larger symbols (white circles and black squares) on the gas-dynamical trajectories in Fig. 3a denote the sections of these trajectories that pass through the shock. The location of these parts of the trajectories on the Doppler map corresponds to the area A to the left of the donor star. A spiral arm is located in the same region in the tomogram, starting approximately from the center of the donor star and situated above L_1. It can be shown that this feature is related to the circumbinary envelope flowing around the stream rather than to the shock wave. We can see from Fig. 4 that the flow from L_1 (the origin of the trajectory is denoted by the white curve with circles) is transformed into the spiral arm B in the third quadrant[2] of the Doppler map. The region of enhanced radiation near the disk apogee, area C, is seen in Fig. 4 as a darkening at the boundary between the first and fourth quadrants. The tidal spiral shock (white dotted curve in Figs. 3a, 3b) or,

---

[2] The quadrants of the coordinate plane are numerated in the usual way: the first quadrant is in the upper right (corresponding to $V_x > 0$, $V_y > 0$), and the other quadrants are numbered consecutively counterclockwise.

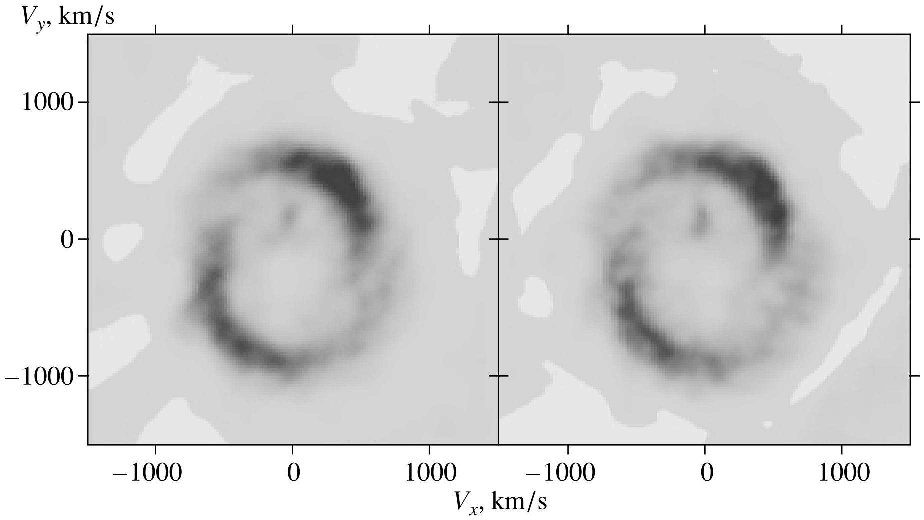

Fig. 8. Doppler map of IP Peg in the He IIλ 4686 Å line for the active state of the system for two different nights of observations [49]. Reproduced by courtesy of Morales-Rueda.

more exactly, the zone of enhanced density behind the shock (area D) forms the bright arm in the first and second quadrants of the Doppler map.

In summary, we can conclude from the synthetic Doppler tomogram for IP Peg in its quiescent state that four elements of the flow make the largest contributions to the luminosity: the hot line, the brightest part of the stream from L_1, a condensation near the disk apogee, and a condensation behind the spiral shock. Obviously, the contributions of these elements to the total luminosity can differ, depending on the specific features of the system. It is also evident that model computations cannot be used to estimate which of these elements dominates. However, by comparing synthetic tomograms with observations, we are able not only to answer this question but also to determine which of the gas-dynamical model parameters require adjustment, enabling us to refine our picture of the flow pattern in the system.

Doppler tomograms for the quiescent state of the system were constructed in [19, 39–42]. As a typical example, Fig. 5 presents Doppler maps in the H_γ and H_β lines from [40]. The distinctive features of these tomograms are a bright spot in area A and a brightening in area C. When compared with our synthetic tomogram (Fig. 4), it is apparent that the largest contributions to the luminosity in the observed tomogram of IP Peg in its quiescent state are made by the shock at the edge of the stream (the hot line) and the condensation near the disk apogee. The absence of appreciable signs of a spiral wave implies either that it is absent from the system or that it is present but has a relatively small intensity. Note also that, in the observational tomogram, the contribution from the stream (area B) is fairly small.

5. RESULTS FOR THE ACTIVE STATE OF IP Peg

Observations (see, for example, the latest reviews [6, 43]) indicate that the disk dominates in the flow structure during outburst. We can therefore assume that the relative role of the stream from the inner Lagrange point decreases in that stage. At present, we have only a qualitative idea of the nature of the outbursts and their parameters; therefore, any correct modeling of outburst-induced variations of the flow structure remains impossible. In order to obtain a gas-dynamical pattern qualitatively similar to flow structures during flares, following [44, 45], we assumed the absence of mass exchange between the components of the system after steady-state flow was attained. There is no doubt that this type of model cannot describe an outburst and the subsequent increase of the disk size; however, qualitatively, it can adequately reflect the effects of the disk and stream on the flow pattern. Studies of the structure of the residual accretion disk have indicated that the flow pattern is already altered substantially at times of about 0.3–0.4 of the orbital period after the time of mass-exchange termination. The stream from L_1 disappears, and, consequently, its effect on the flow structure is terminated. The shape of the accretion disk changes from quasi-elliptical to virtually circular. A second spiral shock appears in the disk structure, resulting from the gravitational effect of the donor star, which is suppressed by the stream from L_1 in systems with constant mass transfer. It is clear that this flow pattern qualitatively displays all features observed in the active state of IP Peg. This provides hope that synthetic tomograms based on this gas-dynamical solution will reveal characteristic features of IP Peg during outburst.

Figure 6a presents distributions of the density and velocity vectors, while Fig. 6b shows the distribution of $\rho^2 T^{1/2}$ in the equatorial plane of the system. We can see that condensations behind the arms of the spiral shock make the largest contribution to the radiation of the system. These areas are marked A and B in Fig. 6b. Figure 7 presents the synthetic Doppler tomogram for this model. Both bright arms in the first and third quadrants are due to emission in the areas of enhanced density behind the arms of the tidal spiral shock.

Observational Doppler tomograms for the active state of the system were constructed in [19, 46–49]. Figure 8 presents typical examples of such tomograms [49]. Their characteristic features are bright arms in the first and third quadrants of the Doppler map. When the synthetic and observational tomograms are compared, it is apparent that both bright arms are due to radiation from condensations behind the arms of the tidal spiral shock (areas A and B).

6. CONCLUSIONS

Gas-dynamical modeling combined with Doppler tomography can be used to identify the basic elements of flows in Doppler maps without solution of an ill-posed inverse problem. Comparison of synthetic tomograms with observations enables refinement of gas-dynamical models and interpretion of the observational data.

We have presented synthetic Doppler tomograms for gas flows in the binary system IP Peg based on three-dimensional gas-dynamical computations. Identification of the areas in the flow responsible for zones of enhanced emission in Doppler maps shows that these areas are different for the quiescent and active states of the system. In the quiescent state, the major contribution to the luminosity in tomograms is made by the shock at the edge of the stream (hot line) and a condensation near the disk apogee. The contributions from the single arm of the spiral wave and the stream are small. The situation changes dramatically when the system goes into its active state. In the corresponding gas-dynamical flow pattern, the stream from L_1 does not play an important role and the system is dominated by the disk and two arms of the spiral shock. Comparison between synthetic and observational tomograms indicates that both zones of enhanced luminosity in the active state of IP Peg are associated with the arms of the spiral shock in the disk.

This study was supported by the Russian Foundation for Basic Research (project nos. 99-02-17619, 99-02-17589, 00-02-16471, 00-01-00392, 00-02-17253) and by Presidential Grants of the Russian Federation (99-15-96022, 00-15-96722, 00-15-96553) and the INTAS (grant no. 00-491). The authors are grateful to V. V. Neustroyev for useful discussions concerning the observational tomograms.

REFERENCES

1. T. R. Marsh and K. Horne, Mon. Not. R. Astron. Soc. **235**, 269 (1988).
2. R. Narayan and R. Nityananda, Ann. Rev. Astron. Astrophys. **24**, 127 (1986).
3. E. L. Robinson, T. R. Marsh, and J. Smak, in *Accretion Disks in Compact Stellar Systems,* Ed. by J. C. Wheeler (World Scientific, Singapore, 1993), p. 75.
4. H. C. Spruit, astro-ph/9806141 (1998).
5. B. R. Frieden, in *Picture Processing and Digital Filtering,* Ed. by T. S. Huang (Springer-Verlag, Heidelberg, 1979), p. 177.
6. T. R. Marsh, in *Proceedings of Astro-Tomography Workshop, Brussels, 2000,* Ed. by H. Boffin, D. Steeghs, and J. Cuypers (in press); astro-ph/0011020 (2000).
7. K. Horne and T. R. Marsh, Mon. Not. R. Astron. Soc. **218**, 761 (1986).
8. C. J. Ferland, Publ. Astron. Soc. Pac. **92**, 596 (1980).
9. M. T. Richards and M. A. Ratliff, Astrophys. J. **493**, 326 (1998).
10. V. A. Lipovetskiĭ and Dzh. A. Stepanyan, Astrofizika **17**, 573 (1981).
11. V. P. Goranskij, S. Yu. Shugarov, E. I. Orlowsky, and V. Yu. Rahimov, Int. Bull. Var. Stars, No. 2653 (1985).
12. J. Smak, Acta Astron. **46**, 377 (1996).
13. J. Wood and C. S. Crawford, Mon. Not. R. Astron. Soc. **222**, 645 (1986).
14. A. W. Shafter, Astrophys. J. **267**, 222 (1983).
15. T. R. Marsh, Mon. Not. R. Astron. Soc. **231**, 1117 (1988).
16. J. C. Martin, D. H. P. Jones, and R. C. Smith, Mon. Not. R. Astron. Soc. **224**, 1031 (1987).
17. J. C. Martin, M. T. Friend, R. C. Smith, and D. H. P. Jones, Mon. Not. R. Astron. Soc. **240**, 519 (1989).
18. G. Beekman, M. Somers, T. Naylor, and C. Hellier, Mon. Not. R. Astron. Soc. **318**, 9 (2000).
19. T. R. Marsh and K. Horne, Astrophys. J. **349**, 593 (1990).
20. K. Horne, H. H. Lanning, and R. H. Gomer, Astrophys. J. **252**, 681 (1982).
21. V. S. Dhillon, T. Marsh, and D. H. P. Jones, Mon. Not. R. Astron. Soc. **252**, 342 (1991).
22. B. Paczyński, Ann. Rev. Astron. Astrophys. **9**, 183 (1971).
23. E. T. Harlaftis, Astron. Astrophys. **346**, L73 (1999).
24. M. S. Catalán, R. C. Smith, and D. H. P. Jones, Mon. Not. R. Astron. Soc. (2001) (in press).
25. J. H. Wood, T. R. Marsh, E. L. Robinson, *et al.*, Mon. Not. R. Astron. Soc. **239**, 809 (1989).

26. S. Wolf, K. H. Mantel, K. Horne, *et al.*, Astron. Astrophys. **273**, 160 (1993).
27. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **77**, 31 (2000) [Astron. Rep. **44**, 26 (2000)].
28. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **219**, 75 (1986).
29. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **72**, 367 (1995) [Astron. Rep. **39**, 325 (1995)].
30. P. L. Roe, Ann. Rev. Fluid Mech. **18**, 337 (1986).
31. S. Chakravarthy and S. Osher, AIAA Pap., No. 85-0363 (1985).
32. R. H. Kaitchuck, E. M. Schlegel, R. K. Honeycutt, *et al.*, Astrophys. J., Suppl. Ser. **93**, 519 (1994).
33. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, and V. M. Chechetkin, Astron. Zh. **74**, 880 (1997) [Astron. Rep. **41**, 786 (1997)].
34. D. V. Bisikalo, A. A. Boyarchuk, V. M. Chechetkin, *et al.*, Mon. Not. R. Astron. Soc. **300**, 39 (1998).
35. D. V. Bisikalo, A. A. Boyarchuk, O. A. Kuznetsov, *et al.*, Astron. Zh. **75**, 40 (1998) [Astron. Rep. **42**, 33 (1998)].
36. T. S. Khruzina, A. M. Cherepashchuk, D. V. Bisikalo, *et al.*, Astron. Zh. **78**, 625 (2001) [Astron. Rep. **45**, 538 (2001)].
37. K. Sawada, T. Matsuda, M. Inoue, and I. Hachisu, Mon. Not. R. Astron. Soc. **224**, 307 (1987).
38. K. Sawada, T. Matsuda, and I. Hachisu, Mon. Not. R. Astron. Soc. **221**, 679 (1986).
39. E. T. Harlaftis, T. R. Marsh, V. S. Dhillon, and P. A. Charles, Mon. Not. R. Astron. Soc. **267**, 473 (1994).
40. S. Wolf, H. Barwig, A. Bobinger, *et al.*, Astron. Astrophys. **332**, 984 (1998).
41. A. Bobinger, H. Barwig, H. Fiedler, *et al.*, Astron. Astrophys. **348**, 145 (1999).
42. A. Bobinger, Astron. Astrophys. **357**, 1170 (2000).
43. D. Steeghs, in *Proceedings of Astro-Tomography Workshop, Brussels, 2000,* Ed. by H. Boffin, D. Steeghs, and J. Cuypers (in press); astro-ph/0012353 (2000).
44. D. V. Bisikalo, A. A. Boyarchuk, A. A. Kilpio, *et al.*, Astron. Zh. **78**, 707 (2001) [Astron. Rep. **45**, 611 (2001)].
45. D. V. Bisikalo, A. A. Boyarchuk, A. A. Kilpio, and O. A. Kuznetsov, Astron. Zh. **78**, 780 (2001) [Astron. Rep. **45**, 676 (2001)].
46. D. Steeghs, K. Horne, T. R. Marsh, and J. F. Donati, Mon. Not. R. Astron. Soc. **281**, 626 (1996).
47. D. Steeghs, E. T. Harlaftis, and K. Horne, Mon. Not. R. Astron. Soc. **290**, L28 (1997).
48. E. T. Harlaftis, D. Steeghs, K. Horne, *et al.*, Mon. Not. R. Astron. Soc. **306**, 348 (1999).
49. L. Morales-Rueda, T. R. Marsh, and I. Billington, Mon. Not. R. Astron. Soc. **313**, 454 (2000).

Translated by K. Maslennikov

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 882–887. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 1008–1013.
Original Russian Text Copyright © 2001 by Tutukov, Fedorova.

Evolution of Stars with High Metallicities

A. V. Tutukov and A. V. Fedorova
Institute of Astronomy, Russian Academy of Sciences, ul. Pyatnitskaya 48, Moscow, 109017 Russia

Received March 20, 2001

Abstract—We have calculated evolutionary tracks for stars with high abundances of heavy elements. An abundance increase from the solar level (~ 0.02) to 0.1 (for $\Delta Y/\Delta Z = 0$–2.4), which corresponds to the central regions of the disk components of giant galaxies, shifts the main sequence towards lower effective temperatures; however, it does not appreciably affect either the luminosity of the stars or their lifetime on the main sequence. Increasing the heavy-element abundance to 0.2 for $\Delta Y/\Delta Z = 2.4$ shifts the main sequence towards higher effective temperatures, appreciably increases the luminosity of the stars, and substantially accelerates their evolution. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The evolution of single and binary stars with solar compositions and with compositions characteristic of the Galactic halo, where the relative abundances of heavy elements are nearly an order of magnitude lower than in the solar region of the Galactic disk, has been rather well studied. Variations of the basic observed parameters of stars in the course of evolution driven by nuclear synthesis, such as surface temperature and luminosity, have been calculated. The evolution of the internal structure of stars has also been modeled, from main-sequence stars to the formation of the final products of stellar evolution: degenerate dwarfs, neutron stars, and black holes (see, for example, [1]).

The structure and evolution of stars with heavy-element contents enhanced by factors of two or three compared to the Sun have also been investigated in numerous studies. We note here, for example, a pioneering work [2] (for Z up to 0.05) and a recent study [3] (for Z up to 0.04). However, there are at least two reasons to expand stellar-evolution modeling to stars with even higher heavy-element abundances. In their study of 13 spiral galaxies, van Zee *et al.* [4] concluded that the abundance of heavy elements increased from the edge of the disk toward its center. The gradient $d\log Z/dR$ was -0.04 to -0.07 kpc^{-1}. For our Galaxy, this gradient is -0.05 to -0.1 kpc^{-1} ([5–7]). As a result, the heavy-element abundance in the central regions of the disk of our Galaxy can be a factor of three to ten higher than in the solar region and can reach ~ 0.1. It is particularly notable that the color characteristics of galaxies vary with their radii together with increasing heavy-element abundance, so that, for most galaxies, the color of the stellar population reddens toward the center in accordance with the metal-abundance increase [8].

Another incentive to study the evolution of stars with high heavy-element abundances is provided by the modeling of galactic evolution on timescales of $\sim 10^{13}$ yr. Tutukov *et al.* [9] indicated that, during the evolution of disk galaxies, the abundance of metals increases with time, reaching ~ 0.1–0.2 by ages of $\sim 10^{12}$ yr.

Here, we study the evolution of single stars for $Z = 0.1$–0.2. The helium abundance in cases of high heavy-element abundances is of particular interest. Empirical estimates yield a broad interval for the ratio $\Delta Y/\Delta Z$, from two [2] to six [11]. We have chosen the value 2.4 [12] to construct two series of models with $Z = 0.1$ and $Z = 0.2$. However, calculations of the evolution of galaxies [9] indicate that, while the heavy-element content increases with the age of a galaxy, the helium abundance can remain almost constant up to ages of $\sim 10^{12}$ yr. Therefore, we also constructed a series of models with $Z = 0.1$ and a helium abundance close to the solar value ($\Delta Y/\Delta Z = 0$).

2. METHOD FOR STELLAR-EVOLUTION CALCULATIONS

We computed the evolution of stars using a code based on the trial run method. We applied the tabulated equation of state calculated in [13, 14]. The opacity was estimated using tables from [15, 16]. The nuclear-reaction rates were taken from [17, 18]. We took into account neutrino energy losses following the formulas derived by Beaudet *et al.* [19]. The ratio l/H_p from the mixing-length theory was taken to be 1.8, since this corresponds to the standard model for the contemporary Sun obtained using this same code. The boundaries of the convective zones were determined using the Schwartzschild criterion. We assumed zero mass loss in the course of the stellar evolution, since we are concerned here primarily with stars on the main sequence (MS) and in the initial stage of helium burning.

There are no opacity tables for $Z > 0.1$; we calculated the tracks for stars with $Z = 0.2$ by adopting the opacity tables for oxygen for the opacities of the higher abundances. To test the validity of this approach, we compared our results with opacity coefficients extrapolated from the tables for $Z = 0.08$ and $Z = 0.1$. This comparison indicated that the opacities obtained using these two methods differ by only a few percent.

We carried out stellar-evolution calculations for four series of models, whose initial chemical compositions are presented in Table 1. Series 1 corresponds to a solar-type initial chemical composition; series 3, to a composition at late stages of galactic evolution (an approximate age of 10^{12} yr); and series 2 and 4, to possible current chemical compositions of the central regions of the stellar disk in our Galaxy and in other giant galaxies for various values for $\Delta Y/\Delta Z$.

3. RESULTS

3.1. Position of the Main Sequence

Figure 1 presents the locations of the four calculated series of models on a zero-age Hertzsprung–Russell (HR) diagram. For comparison, the position on the main sequence of a star with mass 100 $M_\odot$ and the chemical composition of series 1 are indicated. Increasing Z with constant Y (series 4) results in the well known shift of the main-sequence towards lower effective temperatures, which explains the reddening of stellar populations with approach toward the central regions of galaxies with chemical compositions corresponding to model series 4 and 2.

For series 2 and 3, the shift of the main sequence with an increase in Z is relatively small, even when the hydrogen abundance decreases by nearly an order of magnitude (series 3). For series 3, however, the main sequence is shifted in the opposite direction, towards higher effective temperatures. The positions of stars on the main sequence are also altered for this series of models: as Z increases, a star moves upward, in the direction of increasing luminosity. This is due to a decrease in the mean opacity resulting from an increase in the helium abundance. The differences in the positions for stars with $Z = 0.02$ and $Z = 0.1$ are relatively small; however, increasing Z to 0.2 shifts the stars more appreciably, primarily due to the very small abundance of hydrogen corresponding to this Z. The largest shift occurs for masses near 1 $M_\odot$. For small masses near 0.1 $M_\odot$, the shift decreases substantially.

Table 1. Stellar chemical composition

Series of models	Z	X	Y	$\Delta Y/\Delta Z$
1	0.02	0.700	0.280	–
2	0.10	0.428	0.472	2.4
3	0.20	0.088	0.712	2.4
4	0.10	0.620	0.280	0.0

Table 2. Basic parameters of ZAMS stars

$M/M_\odot$	Series	$\log T_{\mathrm{eff}}$ (K)	$\log(L/L_\odot)$	T_c, 10^7 K	ρ_c, g/cm^3
0.1	1	3.479	−2.900	0.447	345
0.1	2	3.463	−2.859	0.589	243
0.1	3	3.652	−2.110	0.981	256
0.1	4	3.441	−2.978	0.478	269
1	1	3.734	−0.176	1.28	72.7
1	2	3.764	0.089	1.74	80.3
1	3	4.017	1.332	2.43	51.2
1	4	3.656	−0.539	1.20	51.4
3	1	4.085	2.007	2.53	37.4
3	2	4.088	2.174	2.55	22.1
3	3	4.333	3.216	3.13	14.8
3	4	3.957	1.694	2.31	27.3
10	1	4.403	3.746	3.32	10.8
10	2	4.405	4.043	3.29	6.50
10	3	4.557	4.736	3.81	5.13
10	4	4.326	3.672	3.05	7.36
30	1	4.586	5.024	3.92	3.37
30	2	4.566	5.242	3.83	2.85
30	3	4.620	5.671	4.27	2.66
30	4	4.515	5.025	3.64	2.93

Table 3. Relative mass of convective nuclei for MS stars

$M/M_\odot$	Series 1	Series 2	Series 3	Series 4
1	–	0.074	–	–
3	0.196	0.202	0.259	0.161
10	0.312	0.335	0.487	0.284
30	0.515	0.566	0.754	0.502

Table 2 presents the basic parameters of stars with masses 0.1, 1, 3, 10, and 30 $M_\odot$ on the zero-

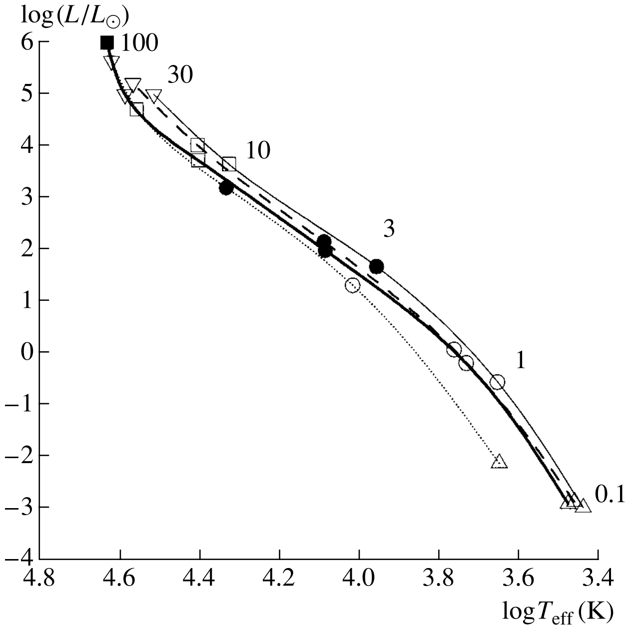

Fig. 1. Zero-age main sequences in the HR diagram. The thick solid curve corresponds to series 1; the dashed curve, to series 2; the dotted curve, to series 3; and the thin solid curve, to series 4. Various symbols indicate the positions of stars with masses 0.1, 1, 3, 10, and 30 $M_\odot$ (and also 100 $M_\odot$ for series 1).

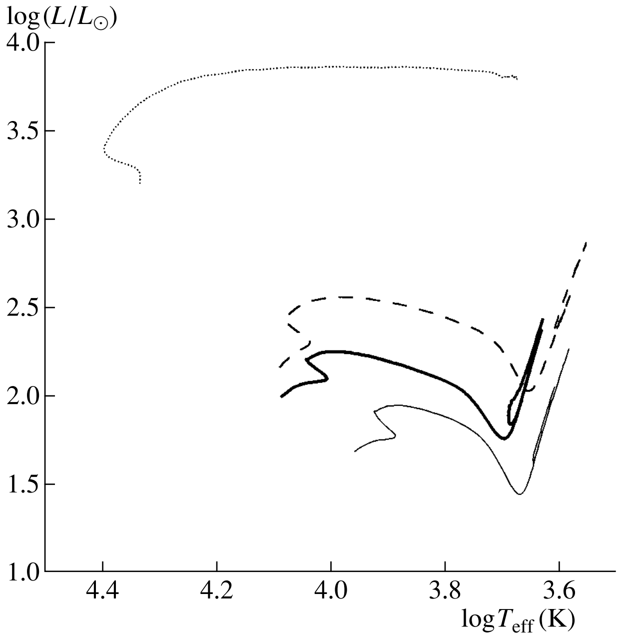

Fig. 3. Same as Fig. 2 for a star with mass 3 $M_\odot$.

age main sequence (ZAMS) for the four calculated chemical compositions. Compared to series 1, the parameters for series 2 and 3, on the one hand, and for series 4, on the other hand, vary in opposite directions.

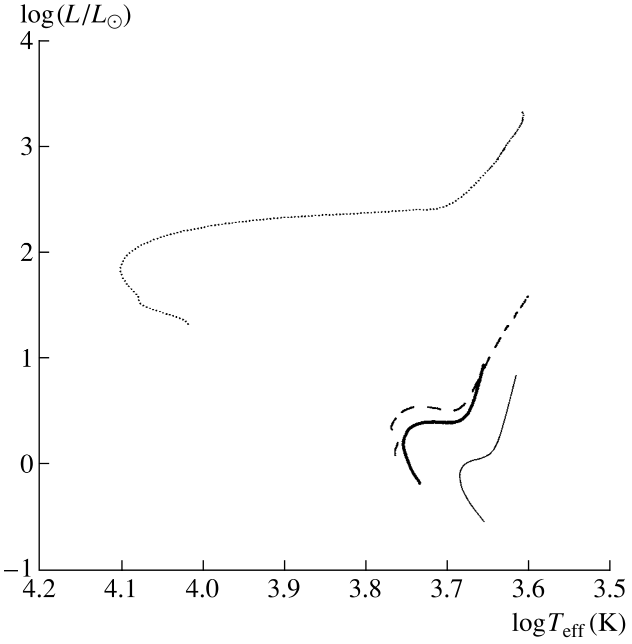

Fig. 2. Tracks of a star with mass 1 $M_\odot$ in the HR diagram. The thick solid curve corresponds to series 1; the dashed curve, to series 2; the dotted curve, to series 3; and the thin solid curve, to series 4.

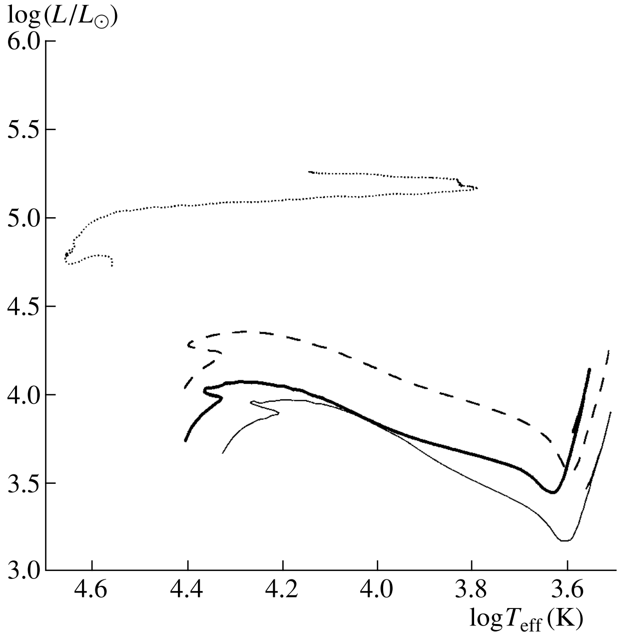

Fig. 4. Same as Fig. 2 for a star with mass 10 $M_\odot$.

The most pronounced variations of the characteristic stellar evolution occur for series 3, that is, for $Z = 0.2$.

Table 3, which presents the relative masses of the convective cores of stars for series 1–4, illustrates the structural variation of zero-age main-sequence stars due to variations in their chemical compositions.

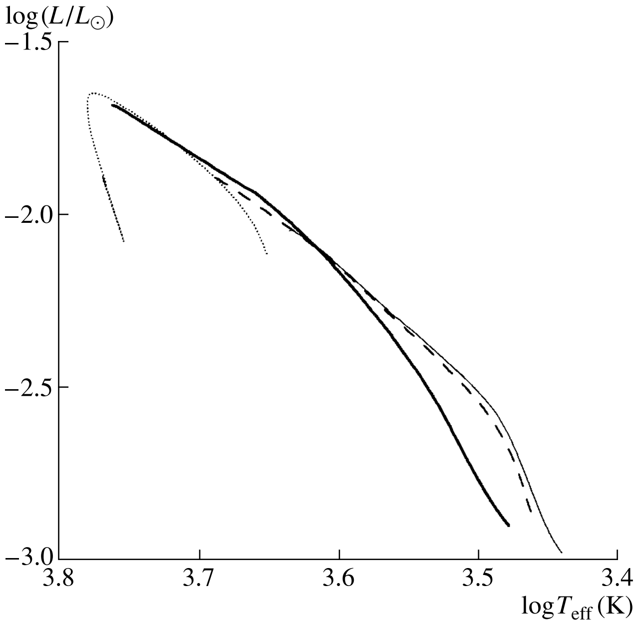

Fig. 5. Same as Fig. 2 for a star with mass 0.1 $M_\odot$.

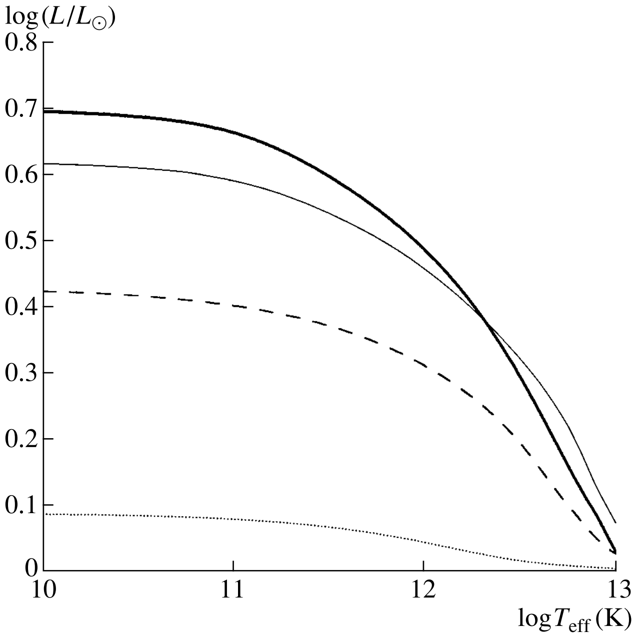

Fig. 6. Variation of hydrogen concentration as a function of time in a star with mass 0.1 $M_\odot$. The thick solid curve corresponds to series 1; the dashed curve, to series 2; the dotted curve, to series 3; and the thin solid curve, to series 4.

Note that a star with mass 1 $M_\odot$ possesses a small convective core only in series 2. For stars of other masses, increasing Z in series 2 and 3 results in appreciable increases in the convective core's mass, whereas, in series 4, on the contrary, the mass of the convective core decreases due to the decrease in stellar luminosity.

3.2. Evolution of Stars with Moderate and High Masses

Figures 2, 3, and 4 present tracks in the HR diagram for stars with masses 1 $M_\odot$, 3 $M_\odot$, and 10 $M_\odot$, respectively. These figures clearly indicate shifts of the evolutionary tracks towards lower luminosities and effective temperatures for series 4 and in the opposite direction for series 2 and 3. This is especially pronounced for $Z = 0.2$ (series 3): the tracks for a star of a given mass for $Z = 0.2$ are located in the region occupied by tracks of stars with solar composition and masses that are nearly a factor of three larger.

Table 4 presents times of hydrogen burning in the center of the star (i.e. the lifetime of the star on the main sequence) for masses 1, 3, 10, and 30 $M_\odot$ for series 1–4. In accordance with luminosity increase for an increase in Z, the stellar evolution is accelerated in series 2 and 3. For series 3, there is additional acceleration due to the low hydrogen abundance, so that the lifetime of a star on the main sequence becomes one to two orders of magnitude shorter than for series 1. In accordance with the obtained luminosity decrease, some deceleration of the stellar evolution occurs for series 4.

3.3. Evolution of Low-mass Stars

We studied the evolution of low-mass stars using a star with mass 0.1 $M_\odot$ as an example. Figure 5 presents the tracks of such a star on a HR diagram for four calculated chemical compositions (the evolution of the star is towards increasing luminosity and effective temperature). The tracks were calculated to an age of 10^{13} yr. Note that the star remains entirely convective during this time for all chemical compositions. The small initial hydrogen abundance for series 3 and also the acceleration of the evolution for this track due to the luminosity increase somewhat change the characteristic evolution of the star for $Z = 0.2$: after 5.4×10^{12} yr, it begins to cool due to hydrogen burning and its luminosity begins to decrease. However, when the central hydrogen abundance decreases to 3.56×10^{-3}, the luminosity and effective temperature of the star begin to increase again. This is associated with the nonequilibrium character of hydrogen burning, which results in an increase in the ^{3}He abundance.

Figure 6 shows the corresponding variations of hydrogen density in the star as a function of time. The times for the hydrogen density to halve are 2.44×10^{12}, 2.76×10^{12}, 9.79×10^{11}, and 3.45×10^{12} yr for the tracks for series 1, 2, 3, and 4, respectively. The substantial increase in the stellar luminosity for series

Table 4. Time in years for the central hydrogen density to decrease to 10^{-6}

$M/M_\odot$	Series 1	Series 2	Series 3	Series 4
1	1.126×10^{10}	4.508×10^{9}	1.005×10^{8}	2.293×10^{10}
3	3.644×10^{8}	1.861×10^{8}	6.000×10^{6}	6.931×10^{8}
10	2.979×10^{7}	1.325×10^{7}	1.431×10^{6}	2.941×10^{7}
30	9.382×10^{6}	7.007×10^{6}	6.986×10^{5}	1.058×10^{7}

3, which is almost an order of magnitude higher than for the track for series 1, appreciably accelerates the hydrogen burning. Thus, the evolution of low-mass stars with chemical compositions characteristic of late stages of galactic evolution (ages of $\sim 10^{12}$ yr) is changed, especially in the final stages of their evolution. However, the degree of acceleration of the evolution is not as high as for more massive stars.

4. CONCLUSIONS

Our study of the evolution of stars with enhanced metal abundances corresponding to those in the central regions of giant galaxies has led to the following conclusions.

(1) An increase in the heavy-element abundance to 0.1 shifts the main sequence and giant branch towards lower effective temperatures, independent of the value of $\Delta Y/\Delta Z$, provided this ratio lies in the interval 0–2.4. This provides an explanation of the reddening towards the center observed in most disks of giant spiral galaxies (for the same star-formation history) [8]. There exist a few galaxies whose disks become bluer toward their central regions; however, this may be due to star-formation activity in the nuclei of these galaxies [20] or may occur when the galactic nucleus accretes material from old stars from the spherical component or intergalactic gas with low abundances of heavy elements [21]. Unfortunately, this effect is difficult to estimate quantitatively due to the lack of sufficient data on the star-formation histories of such galaxies. However, it is clear qualitatively that, whatever the origin of acceleration of star-formation (for example, interaction with nearby galaxies), it results in a blueing of the corresponding disk galaxies [22].

(2) Increasing the heavy-element abundance to 0.1 does not substantially affect either the luminosity of stars with masses 0.1–30 $M_\odot$ or their main-sequence lifetimes, independent of $\Delta Y/\Delta Z$, provided this ratio lies in the interval 0–2.4. This justifies the use of stellar models with solar chemical composition for studies of the evolution of galaxies to ages of $\sim 10^{12}$ yr.

(3) Increasing the heavy-element abundance to 0.2 for $\Delta Y/\Delta Z = 2.4$, which corresponds to the late stages of evolution of galaxies with ages of the order of 10^{12} yr, substantially alters the evolution of the corresponding stars: the main sequence shifts towards higher effective temperatures, the luminosity of the stars increases, and their evolution is appreciably accelerated. Therefore, modeling of the future evolution of galaxies based on tracks for stars with solar chemical composition is applicable only to $Z = 0.1$ and ages $\sim 10^{12}$ yr. For higher ages and heavy-element abundances, the lifetimes of stars become considerably shorter; the mass and nature of their remnants also vary in accordance with the masses of their convective cores. The decrease in the stellar lifetimes accelerates the evolution of galaxies at these phases.

5. ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 99-02-17619).

REFERENCES

1. A. G. Masevich and A. V. Tutukov, *Evolution of Stars: Theory and Observations* [in Russian] (Nauka, Moscow, 1988).
2. A. G. Masevich, G. V. Ruben, S. P. Lomnev, and E. I. Popova, Nauchn. Inf. Astrosoveta Akad. Nauk SSSR **1**, 2 (1965).
3. G. Bono *et al.*, Astrophys. J. **543**, 955 (2000).
4. L. van Zee, J. J. Zalzev, M. P. Hagens, *et al.*, Astron. J. **116**, 2805 (1999).
5. J. L. Hou, N. Prantzos, and S. Boissier, Astron. Astrophys. **362**, 94 (2000).
6. G. Carraro, Y. K. Ng, and L. Portinari, Mon. Not. R. Astron. Soc. **296**, 1045 (1998).
7. L. S. Lyubimkov, *Chemical Composition of Stars* [in Russian] (Astroprint, Odessa, 1995).
8. N. Tamura and K. Ohta, Astron. J. **120**, 533 (2000).
9. A. V. Tutukov, B. M. Shustov, and D. S. Wiebe, Astron. Zh. **77**, 803 (2000) [Astron. Rep. **44**, 711 (2000)].
10. M. Peimbert, A. Peimbert, and M. T. Ruiz, Astrophys. J. **541**, 688 (2000).
11. B. Pagel *et al.*, Mon. Not. R. Astron. Soc. **255**, 325 (1992).
12. T. Thuan and Y. Izotov, astro-ph/0003234.
13. G. Fontaine, H. C. Graboske, and H. M. van Horn, Astrophys. J., Suppl. Ser. **35**, 293 (1977).

14. D. Saumon, G. Chabrier, and H. M. van Horn, Astrophys. J., Suppl. Ser. **99**, 713 (1995).
15. C. A. Iglesias and F. J. Rogers, Astrophys. J. **464**, 943 (1996).
16. D. R. Alexander and J. W. Ferguson, Astrophys. J. **437**, 879 (1994).
17. M. J. Harris, W. A. Fowler, G. R. Caughlan, and B. A. Zimmerman, Ann. Rev. Astron. Astrophys. **21**, 165 (1983).
18. G. R. Caughlan, W. A. Fowler, M. J. Harris, and B. A. Zimmerman, At. Data Nucl. Data Tables **32**, 197 (1985).
19. G. Beaudet, V. Petrosian, and E. E. Salpeter, Astrophys. J. **150**, 979 (1967).
20. E. Krugel and A. V. Tutukov, Astron. Astrophys. **275**, 416 (1993).
21. D. S. Wiebe, B. M. Shustov, and A. V. Tutukov, Astron. Astrophys. **345**, 93 (1999).
22. S. A. Pustilnik, A. Y. Kniazev, V. A. Lipovetski, and A. Y. Ugryumov, astro-ph/0104334.

Translated by K. Maslennikov

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 888–898. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 1014–1024.
Original Russian Text Copyright © 2001 by Khaliullin, Khaliullina, Krylov.

Precision $WBVR$ Photoelectric Photometry of the Eclipsing System RR Lyncis

Kh. F. Khaliullin, A. I. Khaliullina, and A. V. Krylov

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

Received March 23, 2001

Abstract—We carried out accurate ($\sigma_{\rm obs} \approx 0\overset{m}{.}003$) $WBVR$ photoelectric photometry of RR Lyn and obtained light curves of this eclipsing system. Our analysis of the light curves using an iterative differential-correction method yields a self-consistent system of geometrical and physical characteristics of the two components of the system and their evolutionary states. The system's age is estimated to be $t = (1.08 \pm 0.15) \times 10^9$ yr. Observations in all filters are fitted satisfactorily by the same geometry ($r_{1,2}$, i, e, and ω). An analysis of blanketing effects in the W, B, V, and R bands indicates that the atmospheric chemical compositions of both components of the system are peculiar: the primary shows an excess ($[{\rm Fe/H}]_{\rm I} = 0.31 \pm 0.08$) and the secondary a deficit ($[{\rm Fe/H}]_{\rm II} = -0.24 \pm 0.06$) of heavy elements. This is in qualitative and quantitative agreement with the results of an earlier spectroscopic study of RR Lyn by Lyubimkov and Rachkovskaya (1995). The derived physical characteristics of RR Lyn provide evidence that the metallicity effects are probably restricted to the stellar surface layers, while their interiors have normal chemical abundances.

1. INTRODUCTION

The eclipsing binary RR Lyn (=BD+56°1125 = HD 44 691; $P = 9\overset{d}{.}945$; A6IV + F0V; $V = 5\overset{m}{.}54$; $e = 0.08$) is interesting first and foremost because its primary component shows very obvious features of so-called "metal-line" stars, whose nature remains poorly understood [1].

The most detailed spectroscopic studies and radial velocity curves for RR Lyn are presented by Popper [2], Kondo [3], and Lyubimkov and Rachkovskaya [4, 5]. Photoelectric light curves for the system have been published by Huffer [6], Magalashvili and Kumsishvili [7], Botsula [8], Linnell [9], and Lavrov *et al.* [10]. In 1977, the star was proposed for use as a test target for studies of the relativistic advance of orbits of eclipsing systems [11]. Botsula [12] suspected an apsidal advance with a period $U \approx 2000$ yr for this system; however, this result is far from that expected theoretically ($U_{\rm th} \approx 28\,000$ yr) and could not be confirmed by later observations.

The photometric elements for RR Lyn derived by various authors from its light curves differ considerably. Discrepancies between the times of minima observed at close epochs that could not be attributed to observational errors were noted, as well as deviations from Russell's model, in particular, the need to invoke gaseous flows and a "third light" ($L_3 \approx 15$–30%) to obtain a satisfactory description of the light curves [8, 9, 12–14]. However, other evidence for the presence of a third star in the system is lacking [9].

To resolve these contradictions and derive more accurate geometrical and physical parameters of the components, we included RR Lyn in our program of photometric studies.

2. OBSERVATIONS

Our $WBVR$ photometric measurements of RR Lyn were acquired using the 50-cm reflector (AZT-14) at the Tian Shan Mountain Observatory of the Sternberg Astronomical Institute (SAI) and the 60-cm Zeiss-600 reflector at the SAI Crimean Observatory. As in previous observations by other authors, 11 Lyn ≡ BD + 56°1136 = HD 46 590 was used as the comparison star. The star 11 Lyn was one of the standards used when preparing our "Catalog of $WBVR$ Magnitudes of Bright Northern Stars" [15]. For this reason, it was studied photometrically, for variability, over many years and its brightness is very stable. Table 1 presents magnitudes and color indices for the variable and comparison star. The rms errors of the magnitudes are about $0\overset{m}{.}010$ in all filters; the errors of the color indices are about $0\overset{m}{.}005$.

We determined the errors of the differential measurements, $\sigma_{\rm obs}$, for each night from the internal agreement of the observations of standard stars, as well as of the comparison star. For example, on one of the best nights, JD 2 444 595, the $\sigma_{\rm obs}$ values were $0\overset{m}{.}0029$, $0\overset{m}{.}0024$, $0\overset{m}{.}0020$, and $0\overset{m}{.}0023$ for the W, B,

1063-7729/01/4511-0888$21.00 © 2001 MAIK "Nauka/Interperiodica"

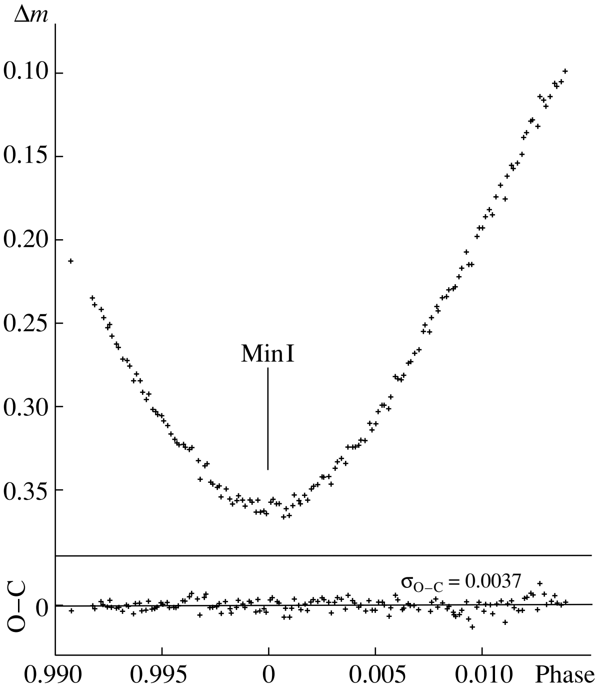

Fig. 1. V light curve of RR Lyn near Min I from observations at the SAI Crimean Observatory. The lower part of the figure shows O–C deviations of individual observed data points from the theoretical curve computed with the derived photometric elements.

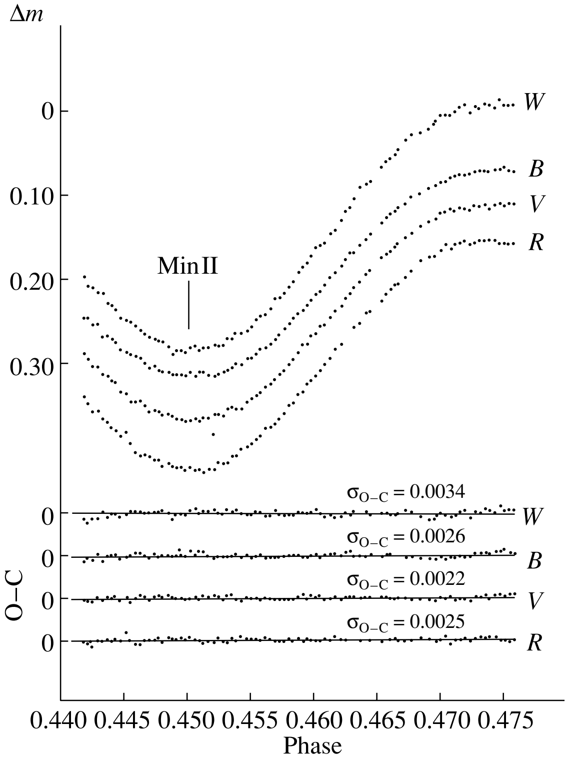

Fig. 2. W, B, V, and R light curves of RR Lyn near the secondary minimum from observations at the Tian Shan Observatory. The lower part of the figure shows O–C deviations of individual observed data points from the theoretical curves computed with the derived photometric elements.

V, and R filters, respectively (based on 95 measurements of RR Lyn in each of the four filters near the secondary minimum over more than eight hours of observations at the Tian Shan Observatory). This good accuracy of individual measurements, typical of many nights of our observations of RR Lyn, is due to the excellent astronomical conditions for photometric measurements at the Tian Shan Observatory, which is at an altitude of 3000 m above sea level. At the same time, the standard error of our measurements on one of the best nights at the SAI Crimean Observatory (altitude of about 500 m, JD 2444988, 145 measurements of RR Lyn during 6 hours of observations, Min I) was only $0\overset{m}{.}0033$ in the V filter. Altogether, we obtained 275 individual brightness measurements of RR Lyn in the V filter and 124 measurements in each of the W, B, and R filters.

The results of our differential measurements of RR Lyn are presented in Tables 2 and 3 and Figs. 1 and 2. They have been corrected for atmospheric extinction using the technique described in [16] and reduced to the standard $WBVR$ photometric system [17]. The Julian times of the observations, $\mathrm{JD}_{\odot}$, are heliocentric. The tables contain differences between the corresponding magnitudes for the variable and comparison star, $\Delta m = m(\mathrm{RR\ Lyn}) - m(11\ \mathrm{Lyn})$. Figure 3 shows the variations of the $W-B$, $B-V$, and $B-R$ color indices with the phase of the orbital cycle.

3. PHOTOMETRIC AND ABSOLUTE ELEMENTS

We determined the photometric elements using the iterative differential-correction method developed by us for the analysis of light curves of eclipsing systems with elliptical orbits [18]. We first analyzed the individual V measurements from Tables 2 and 3, i.e., those having the best accuracy and completeness. Table 4 presents the photometric elements of RR Lyn along with their error estimates obtained by us in an unrestricted (simultaneous) search. Here and below, $r_{1,2}$ are the components' relative radii, expressed as fractions of the system's major axis; $u_{1,2}$ is the limb darkening coefficients of the disks; $L_{1,2,3}$ is the component luminosities, expressed as fractions of the total luminosity; i is the orbital inclination; e is the eccentricity; ω_1 is the longitude of periastron of the primary ($\omega_2 = \omega_1 + 180°$); E_{I} is the heliocentric time of the primary minimum (expressed as $\mathrm{JD}_{\odot}$); E_{II} is the heliocentric time of the secondary minimum ($\mathrm{JD}_{\odot}$); and $\sigma_{\mathrm{O-C}}$ is the rms deviations of individual data points from the theoretical curve.

The subscript "1" and term "primary" refer to the component eclipsed in the primary (deeper) minimum

Table 1. Magnitudes of RR Lyn and the comparison star 11 Lyn

Star	W	B	V	R	$W-B$	$B-V$	$V-R$	Sp
RR Lyn	$5^m.880$	$5^m.761$	$5^m.540$	$5^m.362$	$0^m.119$	$0^m.221$	$0^m.178$	A6IV + F0V
11 Lyn	$5^m.927$	$5^m.882$	$5^m.877$	$5^m.880$	$0^m.045$	$0^m.005$	$-0^m.003$	A2V

Table 2. Individual Tian Shan $WBVR$ measurements of RR Lyn

$\mathrm{JD}_\odot -$ 2 444 000	Δm(RR Lyn − 11 Lyn)				$\mathrm{JD}_\odot -$ 2 444 000	Δm(RR Lyn − 11 Lyn)			
	W	B	V	R		W	B	V	R
595.08774	0.152	0.047	−0.168	−0.338	595.20032	0.225	0.109	−0.100	−0.258
595.09116	0.163	0.047	−0.161	−0.330	595.20380	0.224	0.106	−0.103	−0.266
595.09474	0.165	0.055	−0.153	−0.320	595.20720	0.218	0.101	−0.101	−0.266
595.09841	0.171	0.053	−0.152	−0.321	595.21129	0.218	0.102	−0.102	−0.275
595.10290	0.172	0.070	−0.143	−0.312	595.21559	0.215	0.096	−0.110	−0.277
595.10626	0.184	0.067	−0.141	−0.309	595.21902	0.211	0.095	−0.113	−0.281
595.10956	0.186	0.071	−0.135	−0.301	595.22258	0.202	0.090	−0.117	−0.286
595.11298	0.191	0.075	−0.133	−0.298	595.22636	0.198	0.084	−0.120	−0.290
595.11685	0.201	0.084	−0.122	−0.290	595.22999	0.196	0.083	−0.127	−0.298
595.12106	0.204	0.087	−0.120	−0.295	595.23348	0.190	0.078	−0.130	−0.300
595.12455	0.206	0.090	−0.110	−0.281	595.23785	0.177	0.072	−0.136	−0.303
595.12895	0.211	0.092	−0.110	−0.271	595.24134	0.176	0.067	−0.142	−0.310
595.13249	0.216	0.095	−0.111	−0.267	595.24516	0.171	0.059	−0.147	−0.316
595.13572	0.221	0.100	−0.106	−0.268	595.24886	0.163	0.053	−0.154	−0.326
595.13918	0.224	0.106	−0.098	−0.266	595.25227	0.157	0.048	−0.159	−0.329
595.14295	0.227	0.104	−0.099	−0.262	595.25609	0.147	0.042	−0.168	−0.339
595.14648	0.229	0.109	−0.098	−0.262	595.25944	0.140	0.037	−0.177	−0.341
595.14970	0.229	0.110	−0.094	−0.254	595.26268	0.135	0.033	−0.177	−0.350
595.15391	0.235	0.111	−0.095	−0.258	595.26592	0.128	0.025	−0.187	−0.354
595.15730	0.244	0.114	−0.092	−0.225	595.26929	0.117	0.018	−0.190	−0.359
595.16061	0.240	0.114	−0.089	−0.253	595.27263	0.115	0.017	−0.198	−0.366
595.16395	0.240	0.114	−0.088	−0.253	595.27640	0.112	0.007	−0.200	−0.375
595.16758	0.242	0.116	−0.086	−0.250	595.28068	0.101	−0.001	−0.290	−0.381
595.17125	0.237	0.111	−0.087	−0.253	595.28415	0.097	−0.007	−0.217	−0.392
595.17491	0.234	0.116	−0.089	−0.250	595.28971	0.085	−0.015	−0.223	−0.398
595.17828	0.240	0.111	−0.087	−0.250	595.29266	0.075	−0.022	−0.232	−0.403
595.18164	0.237	0.111	−0.091	−0.247	595.29616	0.070	−0.030	−0.238	−0.410
595.18549	0.236	0.115	−0.090	−0.251	595.29964	0.060	−0.030	−0.245	−0.417
595.18980	0.233	0.116	−0.094	−0.250	595.30475	0.046	−0.040	−0.251	−0.422
595.19330	0.234	0.113	−0.094	−0.259	595.30915	0.042	−0.045	−0.260	−0.436
595.19688	0.233	0.111	−0.092	−0.257	595.31262	0.039	−0.052	−0.265	−0.437

Table 2. (Contd.)

JD$_\odot$ − 2 444 000	Δm(RR Lyn – 11 Lyn)				JD$_\odot$ − 2 444 000	Δm(RR Lyn – 11 Lyn)			
	W	B	V	R		W	B	V	R
595.32180	0.022	−0.065	−0.280	−0.450	615.39020	−0.035	−0.124	−0.333	−0.511
595.32512	0.016	−0.071	−0.283	−0.457	615.39367	−0.044	−0.121	−0.338	−0.514
595.32866	0.012	−0.076	−0.287	−0.463	626.27671	−0.039	−0.126	−0.338	−0.518
595.33211	0.007	−0.079	−0.293	−0.469	626.28157	−0.044	−0.124	−0.339	−0.516
595.33410	0.002	−0.082	−0.296	−0.469	627.27740	−0.036	−0.125	−0.338	−0.512
595.33994	−0.010	−0.087	−0.303	−0.477	627.28157	−0.037	−0.122	−0.338	−0.511
595.34578	−0.016	−0.095	−0.309	−0.485	627.28921	−0.036	−0.125	−0.339	−0.514
595.35227	−0.019	−0.099	−0.316	−0.497	627.29337	−0.036	−0.122	−0.342	−0.518
595.35875	−0.023	−0.105	−0.322	−0.501	627.46207	−0.053	−0.142	−0.330	−0.521
595.36320	−0.029	−0.108	−0.325	−0.501	627.46707	−0.053	−0.121	−0.339	−0.520
595.36396	−0.032	−0.109	−0.325	−0.503	627.47115	−0.049	−0.112	−0.343	−0.503
595.36805	−0.039	−0.114	−0.332	−0.509	627.47532	−0.054	−0.142	−0.333	−0.523
595.37188	−0.039	−0.114	−0.333	−0.516	628.22551	−0.033	−0.118	−0.333	−0.512
595.37570	−0.039	−0.119	−0.337	−0.510	628.23355	−0.037	−0.121	−0.335	−0.514
595.37918	−0.042	−0.120	−0.335	−0.516	628.23772	−0.037	−0.122	−0.366	−0.511
595.38281	−0.048	−0.122	−0.338	−0.518	628.24466	−0.038	−0.122	−0.338	−0.517
595.38731	−0.050	−0.121	−0.337	−0.520	629.08970	−0.038	−0.120	−0.330	−0.516
595.39098	−0.041	−0.122	−0.335	−0.516	629.09456	−0.034	−0.121	−0.339	−0.511
595.39464	−0.048	−0.125	−0.341	−0.519	692.09814	−0.034	−0.124	−0.336	−0.512
595.39891	−0.046	−0.126	−0.340	−0.521	629.10498	−0.039	−0.123	−0.333	−0.510
595.40267	−0.051	−0.126	−0.342	−0.520	629.23206	−0.032	−0.120	−0.337	−0.517
595.40636	−0.050	−0.125	−0.337	−0.521	629.23554	−0.039	−0.120	−0.336	−0.511
595.41100	−0.047	−0.127	−0.342	−0.521	692.23970	−0.042	−0.123	−0.335	−0.507
595.41471	−0.055	−0.128	−0.343	−0.517	629.24595	−0.051	−0.123	−0.339	−0.514
595.41830	−0.050	−0.130	−0.341	−0.517	629.45359	−0.037	−0.128	−0.340	−0.512
595.42200	−0.050	−0.126	−0.343	−0.516	630.05429	−0.035	−0.124	−0.331	−0.519
595.42541	−0.050	−0.125	−0.343	−0.517	630.05845	−0.045	−0.121	−0.340	−0.516
615.37631	−0.035	−0.119	−0.326	−0.504	630.06262	−0.043	−0.121	−0.338	−0.509
615.37979	−0.046	−0.121	−0.331	−0.507	630.30012	−0.007	−0.090	−0.297	−0.477
615.38326	−0.042	−0.124	−0.335	−0.514	630.30359	−0.011	−0.084	−0.289	−0.472
615.38673	−0.038	−0.123	−0.334	−0.518	630.30776	−0.020	−0.082	−0.285	−0.466

(Min I). The values of $E_{\rm I}$ and $E_{\rm II}$ were determined iteratively together with the other photometric elements. Analyzing our own and other published photoelectric timings of minima [6, 8, 19–21] using the least squares solution, we derive the following ephemeris for computing times of minima for RR Lyn:

$$E_{\rm I} = JD_\odot\, 2\,444\,988.49594(30) + 9.9450738(7)E,$$
$$E_{\rm II} = JD_\odot\, 2\,444\,595.17240(30) + 9.9450745(7)E.$$

In parentheses, we present the standard errors of the corresponding values in units of the last digit. Though the recurrence periods for the primary and secondary minima differ, this difference is currently within the derived errors. For comparison, Table 4 presents the photometric elements obtained in [6, 9, 13].

Table 5 presents the solution for our W, B, V, and R light curves for a fixed geometry. The geometrical

Table 3. SAI Crimean Observatory V measurements of RR Lyn

$\mathrm{JD}_{\odot}$ − 2 444 000	ΔV	$\mathrm{JD}_{\odot}$ − 2 444 000	ΔV	$\mathrm{JD}_{\odot}$ − 2 444 000	ΔV	$\mathrm{JD}_{\odot}$ − 2 444 000	ΔV	$\mathrm{JD}_{\odot}$ − 2 444 000	ΔV
988.40404	−0.119	988.45744	−0.007	988.50254	0.035	988.54834	−0.032	988.59504	−0.138
988.41394	−0.097	988.45884	−0.006	988.50394	0.030	988.54964	−0.032	988.59634	−0.145
988.41554	−0.093	988.46054	−0.007	988.50544	0.034	988.55134	−0.030	988.59844	−0.149
988.41784	−0.090	988.46294	0.001	988.50674	0.028	988.55284	−0.037	988.59974	−0.146
988.41924	−0.085	988.46434	0.012	988.50824	0.022	988.55434	−0.049	988.60134	−0.157
988.42084	−0.079	988.46574	0.004	988.50984	0.025	988.55594	−0.048	988.60284	−0.164
988.42234	−0.081	988.46724	0.003	988.51124	0.027	988.55734	−0.047	988.60494	−0.156
988.42374	−0.074	988.46874	0.014	988.51254	0.022	988.55884	−0.050	988.60644	−0.169
988.42524	−0.069	988.47014	0.015	988.51404	0.025	988.56044	−0.057	988.60814	−0.176
988.42674	−0.067	988.47154	0.017	988351564	0.018	988.56184	−0.058	988.60954	−0.174
988.42824	−0.060	988.47304	0.016	988.51714	0.016	988.56394	−0.063	988.61094	−0.177
988.42984	−0.059	988.47444	0.023	988.51924	0.015	988.56534	−0.065	988.61284	−0.182
988.43164	−0.056	988.47604	0.018	988.52064	0.011	988.56704	−0.076	988.61424	−0.192
988.43314	−0.047	988.47774	0.024	988.52204	0.011	988.56854	−0.080	988.61564	−0.195
988.43464	−0.051	988.47904	0.027	988.52364	0.010	988.57024	−0.076	988.61704	−0.202
988.43604	−0.047	988.48064	0.025	988.52504	0.015	988.57174	−0.084	988.61854	−0.203
988.43744	−0.040	988.48214	0.022	988.52664	0.006	988.57334	−0.091	988.62014	−0.199
988.43884	−0.036	988.48354	0.025	988.52814	0.002	988.57484	−0.088	988.62164	−0.217
988.44054	−0.039	988.48514	0.028	988.52994	0.000	988.57934	−0.096	988.62324	−0.215
988.44194	−0.030	988.48654	0.025	988.53144	0.003	988.57804	−0.097	988.62454	−0.211
988.44324	−0.029	988.48814	0.026	988.53304	−0.007	988.57944	−0.101	988.62644	−0.217
988.44434	−0.027	988.48944	0.032	988.53464	−0.007	988.58104	−0.102	988.62804	−0.225
988.44594	−0.026	988.49094	0.025	988.53614	−0.007	988.58264	−0.103	988.62964	−0.223
988.44744	−0.023	988.49234	0.032	988.53764	−0.008	988.58404	−0.109	988.63114	−0.226
988.44884	−0.020	988.49374	0.031	988.63914	−0.011	988.58564	−0.114	988.63274	−0.232
988.45024	−0.015	988.49534	0.033	988.54074	−0.011	988.58714	−0.124	989.30194	−0.341
988.45174	−0.012	988.49684	0.026	988.54234	−0.021	988.58874	−0.116	989.30334	−0.335
988.45304	−0.010	988.49834	0.024	988.54384	−0.017	988.59034	−0.116	989.30474	−0.335
988.45454	−0.009	988.49964	0.027	988.54534	−0.021	988.59204	−0.133	989.30614	−0.336
988.45594	−0.009	988.50104	0.027	988.54694	−0.028	988.59354	−0.138	989.30754	−0.336

parameters $r_{1,2}$, i, e, and ω were fixed in accordance with the solution of the most complete (V) light curve: there is no reason to expect a dependence of these parameters on the spectral band considered for RR Lyn. Thus, when solving the W, B, and R light curves, we varied only the photometric characteristics of the components: L_1, L_2, u_1, and u_2.

Together with the W, B, V, and R light curves of RR Lyn near the minima, Figs. 1 and 2 present the O−C deviations of the observed data points from the corresponding theoretical curves plotted for the photometric elements in Tables 4 and 5 and the corresponding $\sigma_{\mathrm{O-C}}$ values.

The derived elements lead us to the following conclusions.

(1) There are no significant systematic variations of O−C with orbital phase. This indicates that the model used is in good agreement with our observations and that there is no need to invoke a third light to describe the observations. Most of the anomalies

Table 4. Photometric elements of RR Lyn

Parameter	Our results (V)	Huffer [6] ($\lambda \sim 4500$ Å)	Botsula [13] ($\lambda \sim 4600$ Å)	Linnell [9] (V)
r_1	0.0878 ± 0.0005	0.080	0.090	0.082
r_2	0.0541 ± 0.0011	0.050	0.057	0.066
i	$87.^\circ45 \pm 0.^\circ11$	$87.^\circ8$	$87.^\circ0$	$88.^\circ3$
e	0.0782 ± 0.0009	0.081 (fixed)	0.079	0.080 (fixed)
ω	$185^\circ \pm 5^\circ$	164°	181°	167° (fixed)
L_1	0.7835 ± 0.0039	0.794	0.780	0.6738
L_2	$1 - L_1$	$1 - L_1$	$1 - L_1$	0.3262
L_3	0 (fixed)	0 (fixed)	0 (fixed)	30.4%
u_1	0.43 ± 0.08	0.0 (fixed)	0.6–0.8	0.2908
u_2	0.59 ± 0.12	0.0 (fixed)	0.8	0.295

Fixed parameters are indicated.

Table 5. W, B, V, and R light-curve solution for RR Lyn with fixed geometrical parameters.

Parameter	W	B	V	R
L_1	0.7524 ± 0.0011	0.7938 ± 0.0010	0.7835 ± 0.0008	0.7656 ± 0.0009
L_2	$1 - L_1$	$1 - L_1$	$1 - L_1$	$1 - L_1$
u_1	–	–	0.43 ± 0.08	–
u_2	0.38 ± 0.10	0.69 ± 0.08	0.59 ± 0.07	0.16 ± 0.09
$\sigma_{\rm O-C}$ (MinI)	–	–	$0.^m0037$	–
$\sigma_{\rm O-C}$ (MinII)	$0.^m0034$	$0.^m0024$	$0.^m0022$	$0.^m0023$

and discrepancies in the light curves of RR Lyn reported earlier were apparently due to the insufficient accuracy of most early observations.

(2) The observations in all bands (W, B, V, and R) can be satisfactorily described with the same geometry ($r_{1,2}$, i, e, and ω). In both the light curves and the O–C curves, we can see physical brightness fluctuations with an amplitude of about $0.^m003$ and a time scale of $\sim 1^{\rm h}$ characteristic of δ Sct stars. Additional studies are need to determine which of the stars (RR Lyn or 11 Lyn) is responsible for these fluctuations.

(3) Tables 4 and 5 show that the limb-darkening coefficients derived for the components u_1 and u_2 are appreciably below the theoretically expected value ($u_{\rm th} \approx 0.60$). However, due to partial eclipses and the ellipticity of the orbit, the values of u_1 and u_2 and of the longitude of periastron ω are not determined very accurately. This is mainly due to mutual correlations between the estimates of u_1, u_2, e, and ω, which were found simultaneously.

(4) The variations of $W-B$ and $B-V$ with orbital phase in Fig. 3 are in qualitative agreement with similar curves for $U-B$ and $B-V$ presented by Linnell [9], but our curves are appreciably more accurate and we have observations in the R band. Figure 3 clearly shows a situation that at first seems paradoxical: in the secondary minimum, $B-V$ and $V-R$ become bluer, as expected for an eclipse of the cooler secondary (Sp II = F0V), but $W-B$ simultaneously becomes much redder. Taking into account the geometry of the eclipse and the luminosity ratio of the components, the known dip in the ($U-B$, $B-V$) diagram in the region of A stars of luminosity classes IV–V (Sp I = A6IV) can explain only about 30% of this effect. It will become clear below that the remaining 70% of the unusual behavior of $W-B$ is the result of anomalies of the heavy-element abundances in the atmospheres of both components.

Based on our data presented in Tables 1–5 and published spectroscopic data [2–5], we can calculate the physical and geometric parameters of both components of RR Lyn collected in Table 6.

To determine the components' ages t_1 and t_2 presented in Table 6, we plotted $\log g(t)$ relations for the two stars (Fig. 4), i.e., evolutionary tracks

Table 6. Main physical and geometric characteristics of the eclipsing binary RR Lyn

Parameter	Primary component	Secondary component
Mass, M	$1.89 \pm 0.07 M_{\odot}$	$1.49 \pm 0.05 M_{\odot}$
Radius, R	$2.57 \pm 0.04 R_{\odot}$	$1.58 \pm 0.04 R_{\odot}$
Luminosity, L	$21.04 \pm 1.47 L_{\odot}$	$5.31 \pm 0.32 L_{\odot}$
Effective temperature, T_e	7570 ± 120 K	6980 ± 100 K
Spectral type from $WBVR$ photometry, Sp	A6IV	F0V
Gravity, $\log g$	3.894 ± 0.019	4.214 ± 0.025
Age, t	$(1.08 \pm 0.15) \times 10^9$ yr	$(0.97 \pm 0.32) \times 10^9$ yr
Radius excess, $\Delta \log R$	0.214 ± 0.007	0.045 ± 0.012
Luminosity excess, $\Delta \log L$	0.259 ± 0.029	0.092 ± 0.025
V	$5^{m}.805 \pm 0.015$	$7^{m}.202 \pm 0.024$
$W-B$	$0^{m}.177 \pm 0.017$	$-0^{m}.079 \pm 0.020$
$B-V$	$0^{m}.207 \pm 0.011$	$0^{m}.273 \pm 0.012$
$V-R$	$0^{m}.153 \pm 0.013$	$0^{m}.265 \pm 0.015$
Absolute bolometric magnitude, $M_{\rm bol}$	$1^{m}.48 \pm 0.07$	$2^{m}.88 \pm 0.07$
Metallicity, [Fe/H]	0.31 ± 0.08	-0.24 ± 0.06
Orbital period, P	$9^{d}.9450740 \pm 0^{d}.0000007$	
Orbital semi-major axis, a	$29.23 \pm 0.34 R_{\odot}$	
Orbital inclination, i	$87^{\circ}.45 \pm 0^{\circ}.11$	
Distance to the system, r	73.5 ± 2.8 pc	
Photometric parallax, $\pi_{\rm ph}$	$0''.0136 \pm 0''.0005$	
Trigonometric parallax, π	$0''.0120 \pm 0''.0010$	
Age of the system, t	$(1.08 \pm 0.15) \times 10^9$ yr	

for stars with normal chemical abundances ($Z = 0.02$) and masses $M_1 = (1.89 \pm 0.07) M_{\odot}$ and $M_2 = (1.49 \pm 0.05) M_{\odot}$, using stellar models computed by Claret and Gimenez [22]. Our choice of this particular diagram for the age determinations was based on the fact that the gravity ($g \propto M/R^2$) is very sensitive to age and the radii needed to compute this quantity can be derived from light curves much more accurately than, for instance, luminosities or other stellar parameters. The widths of the evolutionary track strips in this diagram represent the uncertainties in the component masses. The position where the evolutionary track crosses the line $y = \log g_{\rm obs}$ yields the desired age for the star. The resulting ages t_1 and t_2, presented in Table 6, are in good agreement with each other within the experimental error. Because the accuracy of t_1 is higher, we adopt this value as the age of the RR Lyn system:

$$t = (1.08 \pm 0.15) \times 10^9 \quad \text{yr};$$

this is, in good agreement with the value $t = (1.1 \pm 0.3) \times 10^9$ yr derived by Lyubimkov and Rachkovskaya [5] using other data.

4. SPECTRAL TYPES, EFFECTIVE TEMPERATURES, AND METALLICITIES OF THE COMPONENTS

The primary of RR Lyn is a metal-line star, making it difficult to determine its effective temperature T_e from spectroscopic data. Its spectral type depends on the system of lines used; the mean estimates are A3, A7, and F0 based on lines of calcium, hydrogen, and metals, respectively [2, 3]. Spectral-type estimates for the secondary based on spectroscopic data also show a wide range: F0V–F3V [2–5]. Thus, we hoped that the results of our multicolor photometry study would lead to a better estimate for T_e, especially since we were able to reliably separate the light of the two components in each of the $WBVR$ bands.

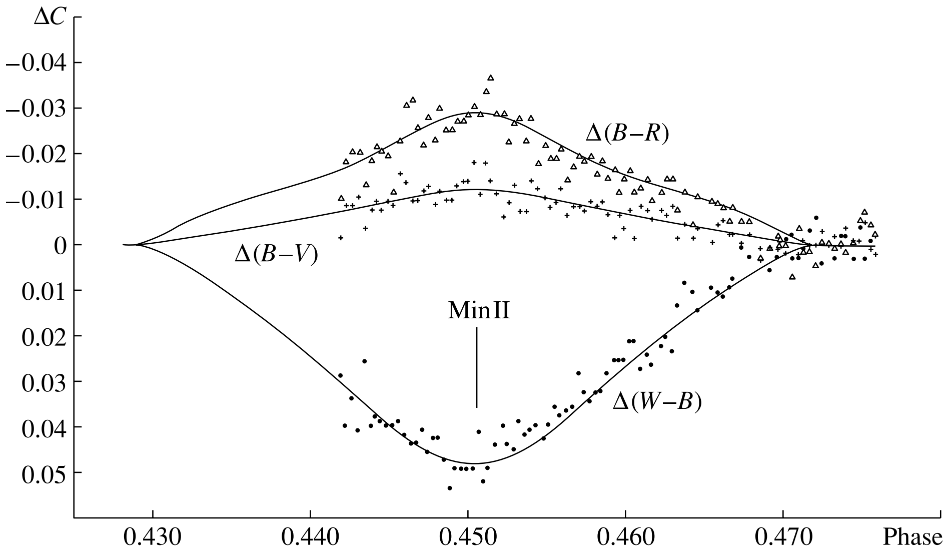

Fig. 3. Variations of $W-B$, $B-V$, and $B-R$ of RR Lyn with phases in the secondary minimum. The solid curves are theoretical curves plotted for the photometric elements in Tables 4 and 5.

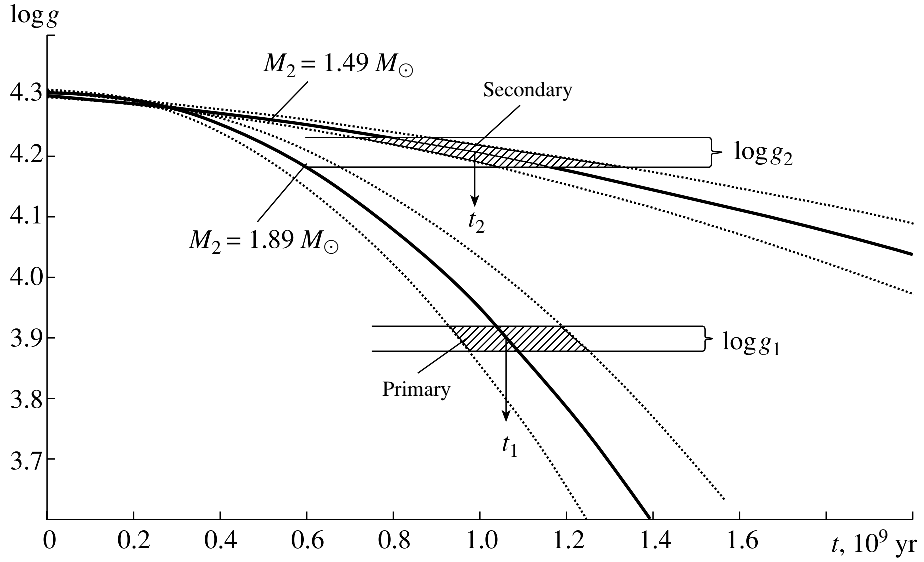

Fig. 4. Evolutionary tracks, $\log g(t)$, for stars with normal chemical composition ($Z = 0.02$) and masses $M_1 = 1.89 \pm 0.07 M_\odot$ and $M_2 = 1.49 \pm 0.05 M_\odot$ corresponding to those of the components of RR Lyn constructed to derive the components' ages t_1 and t_2.

The components' excess radii, $\Delta \log R$, and luminosities, $\Delta \log L$, presented in Table 6 were calculated relative to the zero-age main-sequence ($t = 0$) and normal chemical composition ($Z = 0.02$) stellar models from [22]. In accordance with its excess radius ($\Delta \log R_1 = 0.214$) and excess luminosity ($\Delta \log L_1 = 0.259$), we attribute the primary to luminosity class IV, though both components are usually classified spectroscopically as class V stars.

The components of the system are plotted in a ($B-V$, $V-R$) two-color diagram in Fig. 5. We have plotted the corresponding MK spectral types for luminosity classes IV–V along the axes of this diagram in addition to the color indices from the *WBVR* catalog [15]. The values of $B-V$ and $V-R$ do not depend strongly on luminosity class and metallicity in the color range of interest and are good indicators of effective temperature. We introduced small metallicity corrections in accordance with [23, 24], shown in Fig. 5 as blanketing vectors, and then estimated the spectral types of components of RR Lyn:

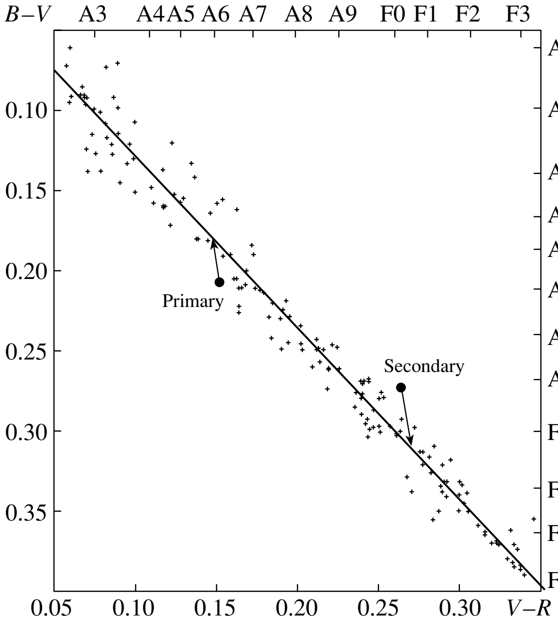

Fig. 5. Positions of the components of RR Lyn in the ($B-V$, $V-R$) diagram. The arrows are blanketing vectors. The crosses show bright, unreddened stars of luminosity classes IV–V from the catalog [18].

$$\mathrm{Sp(RR\ Lyn)} = \mathrm{A6IV} + \mathrm{F0V}.$$

Figure 4 indicates that the secondary has not yet significantly left the main sequence, so that we can use the known scales of effective temperatures T_e and bolometric corrections B.C. for stars of luminosity class V for this star [25]:

$$T_e(\mathrm{F0V}) = 6980 \pm 100\ \mathrm{K},$$
$$\mathrm{B.C.(F0V)} = -0\overset{m}{.}01.$$

At the same time, the primary has already begun to leave the main sequence and its temperature is rapidly decreasing. It is not possible to obtain unique calibrations for this comparatively rapid phase of evolution; therefore, we used the ratio of the surface brightnesses of the components to determine T_e for this star:

$$\frac{I_1}{I_2} = \frac{L_1}{L_2}\left(\frac{r_2}{r_1}\right)^2 = \frac{e^{c_2/\lambda T_2} - 1}{e^{c_2/\lambda T_1} - 1}.$$

With the above values for the relative luminosities $L_{1,2}$ and radii $r_{1,2}$ (Tables 4 and 5) and assuming the components' effective temperatures to be equal to their brightness temperatures, we obtain

$$T_e(\mathrm{A6IV}) = 7570 \pm 100\ \mathrm{K}.$$

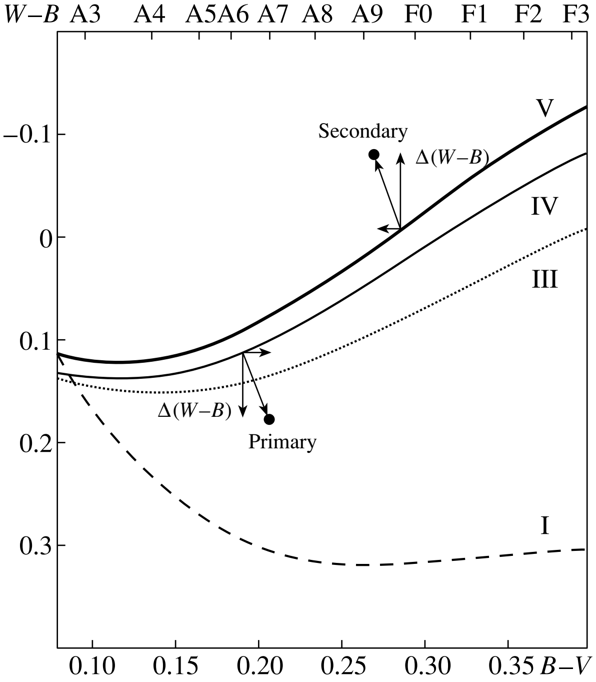

Fig. 6. The positions of the components of RR Lyn in the ($W-B$, $B-V$) diagram. The roman numerals show the normal color lines for the corresponding luminosity classes. Arrows are the blanketing vectors. The $\Delta(W-B)$ values are color excesses due to blanketing effects.

This T_e value is an average of the results derived from light-curve solutions in two bands, $T_1 = 7565$ K for B and $T_1 = 7575$ K for V. Most of the radiation of both components is restricted to these bands. The W band is strongly distorted by metal-line blanketing and cannot be used for this purpose.

The light-curve solution can be used to determine I_1/I_2 very accurately, making the T_e value for the primary obtained using this ratio the most reliable and accurate currently available. Note that the T_e value derived from our multicolor photometry is significantly lower than spectroscopic estimates, $T_1 = 8000$–8300 K [3–5]. This contradiction may reflect the unusual atmospheric structure of metal-line stars.

The components' absolute luminosities L and bolometric magnitudes M_{bol} computed using the above absolute radii and temperatures are also given in Table 6. These parameters enable us to find the distance to the system r and its photometric parallax π_{ph} from the relation

$$M_{\mathrm{bol}} = V + \mathrm{B.C.} + 5 + 5\log\pi_{\mathrm{ph}}.$$

Substituting the parameters of the primary from Table 6 and the bolometric correction B.C. $= -0.01$ [25], we obtain

$$\pi_{\mathrm{ph}} = 0\overset{''}{.}0137 \pm 0\overset{''}{.}0005, \quad r = 73.5 \pm 2.8\ \mathrm{pc}.$$

The Hipparcos trigonometric parallax for the star [26] is $\pi = 0''.01201 \pm 0''.00097$. We can see that the π_{ph} and π values are in agreement within the experimental error. However, our π_{ph} value is now probably the most reliable and accurate.

Figure 6 plots the components of the RR Lyn system in a $(W-B,\ B-V)$ two-color diagram. In this diagram, metal-deficient stars are located above and to the left of stars having the same luminosity and temperature but normal abundances, whereas metal-abundant stars are located below and to the right, demonstrating an ultraviolet deficiency [23, 24]. A star's metallicity, [Fe/H], can be determined from the length of the blanketing vector [27]. The blanketing vectors for both components are plotted in Fig. 6. Note that the parameters of the vectors were computed for $(U-B,\ B-V)$ diagrams and we had to introduce small ($\leqslant 10\%$) corrections for the different spectral reaction curves of the W and U bands [17, 23]. This yields

$$[\text{Fe/H}] = +0.31 \pm 0.08$$

for the primary and

$$[\text{Fe/H}] = -0.24 \pm 0.06$$

for the secondary.

These values demonstrate both qualitative and good quantitative agreement with the spectroscopic results obtained by Lyubimkov and Rachkovskaya [4].

5. CONCLUSIONS

Our accurate ($\sigma_{\text{ind}} \leqslant 0^m.003$) multicolor *WBVR* photoelectric measurements have enabled us to develop a self-consistent geometrical and physical model for the eclipsing system RR Lyn and to derive the photometric and physical elements presented in Table 6.

An important result of our study is the conclusion that both components of the system have peculiar atmospheric chemical compositions: the primary shows an excess of heavy elements; the secondary, a deficiency. This follows from our analysis of the integrated effects of blanketing by heavy elements in the broad W, B, V, and R bands. It has been known for a long time that the primary had a peculiar chemical composition; namely, that it was a metal-line star. It was only in 1995 that Lyubimkov and Rachkovskaya [4, 5] demonstrated that the chemical abundances for the other star in the system were also abnormal based on spectroscopic studies. However, in contrast to the primary, the secondary has a heavy-element deficiency. This unexpected result required confirmation, which is provided by our multicolor photometric analysis. Since the components of RRLyn were most probably formed simultaneously and had the same initial chemical composition and this system has not experienced mass transfer, this result argues in favor of the diffusion theory. This theory explains chemical-abundance anomalies as being a consequence of the separation of elements in the stellar surface layers under the influence of the force of gravity and the opposing radiation pressure [28].

The main physical characteristics of both components of RR Lyn (their masses, radii, and luminosities) show none of the anomalies often ascribed to metal-line stars and, within the errors, are consistent with current evolutionary models for stars of normal chemical composition [22]. Attempts to describe the primary and secondary using evolutionary tracks from [22] with $Z_{\text{I}} = 0.03$ and $Z_{\text{II}} = 0.01$ lead to contradictions. This indicates that the chemical-abundance anomalies of the components of RR Lyn are restricted to the stellar surface layers, while the interiors have a normal chemical abundance, $Z = 0.02$. This result is important to our understanding of the nature of metal-line stars, and the RR Lyn system provides us with a unique opportunity to continue studies concerned with the metal-line phenomenon.

REFERENCES

1. L. S. Lyubimkov, *Chemical Composition of Stars: Method and Analysis Results* [in Russian] (Astroprint, Odessa, 1995).
2. D. M. Popper, Astrophys. J. **169**, 549 (1971).
3. M. Kondo, Ann. Tokyo Astron. Obs., Second. Ser. **16**, 18 (1976).
4. L. S. Lyubimkov and T. M. Rachkovskaya, Astron. Zh. **72**, 64 (1995) [Astron. Rep. **39**, 56 (1995)].
5. L. S. Lyubimkov and T. M. Rachkovskaya, Astron. Zh. **72**, 72 (1995) [Astron. Rep. **39**, 63 (1995)].
6. C. M. Huffer, Publ. Waschburn Obs. **15**, 199 (1931).
7. N. L. Magalashvili and Ya. I. Kumsishvili, Byull. Akad. Nauk Gruz. SSR, Abastumanskaya Astrofiz. Obs., No. 24, 13 (1959).
8. R. A. Botsula, Byull. Astron. Obs. im. Éngel'garda, No. 35, 43 (1960).
9. A. P. Linnell, Astron. J. **71**, 458 (1966).
10. M. I. Lavrov, N. V. Lavrova, and Yu. F. Shabalov, Tr. Kazan. Obs., Issue 51, p. 19 (1988).
11. R. H. Koch, Astrophys. J. **183**, 275 (1973).
12. M. I. Lavrov and N. V. Lavrova, Astron. Tsirk., No. 1165 (1981).
13. R. A. Botsula, Astron. Zh. **44**, 1253 (1967) [Sov. Astron. **11**, 1000 (1967)].
14. E. Budding, Astrophys. Space Sci. **30**, 433 (1974).
15. V. G. Kornilov, I. M. Volkov, A. I. Zakharov, *et al.*, Tr. Gos. Astron. Inst., Mosk. Gos. Univ. **63**, 4 (1991).
16. V. G. Moshkalev and Kh. F. Khaliullin, Astron. Zh. **62**, 393 (1985) [Sov. Astron. **29**, 227 (1985)].
17. Kh. Khaliullin, A. V. Mironov, and V. G. Moshkalyov, Astrophys. Space Sci. **111**, 291 (1985).
18. A. I. Khaliullina and Kh. F. Khaliullin, Astron. Zh. **61**, 393 (1984) [Sov. Astron. **28**, 228 (1984)].

19. D. B. Catton, R. L. Hawkins, and W. C. Burns, Int. Bull. Var. Stars, No. 3408 (1989).
20. D. B. Catton and W. C. Burns, Inf. Bull. Var. Stars, No. 3900 (1993).
21. J. Isles, BAAVSS, Circ. No. 72, 22 (1991).
22. A. Claret and A. Gimenez, Astron. Astrophys., Suppl. Ser. **96**, 255 (1992).
23. V. Straizys, *Multicolor Stellar Photometry* (Mokslas, Vilnius, 1977; Pachart Publ. House, Tucson, 1992).
24. V. Straizys, *Metal-poor Stars* [in Russian] (Mosklas, Vilnius, 1982).
25. D. M. Popper, Ann. Rev. Astron. Astrophys. **18**, 115 (1980).
26. M. A. C. Perryman, *The HIPPARCOS and TYCHO Catalogues* (Esa, Noordwijk, 1997), ESA SP-1200.
27. A. L. T. Powell, Mon. Not. R. Astron. Soc. **155**, 483 (1972).
28. G. Michaud, in *Atmospheric Diagnostics of Stellar Evolution: Chemical Peculiarity, Mass Loss, and Explosion (IAU Colloquium 108),* Ed. by K. Nomoto (Springer-Verlag, Berlin, 1988), p. 3.

Translated by N. Samus'

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 899–907. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 1025–1033.
Original Russian Text Copyright © 2001 by Postnov, Prokhorov.

The Relation between the Observed Mass Distribution for Compact Stars and the Mechanism for Supernova Explosions

K. A. Postnov and M. E. Prokhorov

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

Received March 26, 2001

Abstract—The observed mass distribution for the compact remnants of massive stars (neutron stars and black holes) and its relationship to the possible mechanism of ejection of the envelopes of type II and Ib/c supernovae are analyzed. The observed lack of compact remnants with masses 1.5–3 $M_\odot$ suggests a magneto-rotational mechanism for the supernovae, and a soft equation of state for neutron stars with limiting masses near 1.5 $M_\odot$. The observational consequences of this hypothesis are discussed.

1. INTRODUCTION

Current astrophysical data indicate that the masses of compact objects arising as a consequence of stellar evolution (white dwarfs, neutron stars (NS), and black holes (BH)) have different distributions within theoretically permitted intervals. The observed masses of white dwarfs lie in a wide range from several tenths of solar mass to nearly the Chandresekhar limit ($\sim 1.2\ M_\odot$), with low-mass white dwarfs encountered most often. We are not concerned with these objects here and will not consider them further. In contrast to white dwarfs, the masses of neutron stars lie within a very narrow interval: the masses for 26 neutron star radio pulsars in binary systems are consistent with a normal distribution with mean mass 1.35 $M_\odot$ and dispersion 0.04 $M_\odot$ [1]; the total scatter in the masses of five binary neutron stars whose masses have been determined accurately based on observations of relativistic effects is only 7%. As noted by Thorsett and Chakrabarty (1999) [1], there is currently not a single radio pulsar in a binary system with different components whose mass exceeds 1.45 $M_\odot$. If we add the less accurately determined masses of neutron stars in X-ray binaries [3] to this sample, the observed mean mass of neutron stars is $M_{\rm NS} = (1.35 \pm 0.15)\ M_\odot$ (the same mean as above with a larger dispersion).

More than a dozen black hole candidates in close X-ray binary systems are known (see [2, 3] and references therein). The classical lower limit for the mass of an invisible component in such a binary, M_x, is determined by the mass function $f(M_o)$, which is derived from the radial-velocity curve of the optical component. Using estimates for the inclination of the binary orbit to the line of sight, we can obtain an estimate for the mass M_x. In this way, it was discovered that the masses of black-hole candidates are nearly equally distributed in the interval ~4–40 $M_\odot$, with a mean value of about 10 $M_\odot$.

In addition to the reliable dynamical determinations of the masses of neutron stars and black holes in binary pulsars and X-ray novae, there exists a number of less accurate mass estimates for compact objects in X-ray binaries. For example, (1) the mass of the neutron star in the low-mass X-ray binary Cyg X-2 has been estimated to be $1.8 \pm 0.2\ M_\odot$ [4]; (2) mass estimates for the neutron star in the Vela X-1 X-ray pulsar are $M_{\rm NS} \sim 1.9\ M_\odot$ [5] and $M_{\rm NS} \sim 1.4\ M_\odot$ [6]; and (3) in the eclipsing low-mass X-ray binary 4U 1700–37, $M_{\rm NS} = 1.8 \pm 0.4\ M_\odot$ [7] (possibly indicating that this object is a low-mass black hole [8]). Until the high masses of these neutron stars are verified, we will consider them to be uncertain. The possible existence of low-mass black holes in transient X-ray sources is discussed below.

Thus, we assume that current reliable measurements of neutron-star masses lie in the narrow interval $M_{\rm NS} = (1.35 \pm 0.15)\ M_\odot$, masses of black holes lie in the wide interval $M_{\rm BH} > 3\ M_\odot$, and no objects have currently been reliably detected with masses in the gap between those for neutron stars and black holes from 1.5 to 3 $M_\odot$. This bimodality in the mass distributions for compact objects is noted in [9] and illustrated in Fig. 1.

Can we find an overall explanation for this pattern?

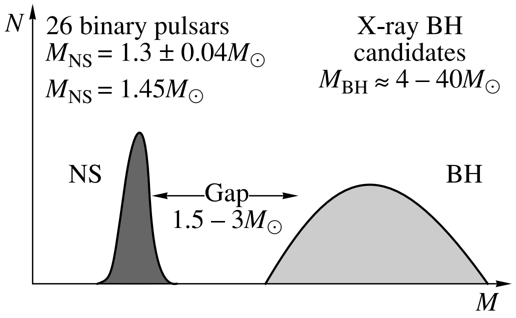

Fig. 1. Schematic of the mass distributions for neutron stars and black holes.

Neutron stars and black holes are the products of the evolution of massive stars. The matter in the cores of these stars undergoes thermonuclear evolution, forming heavy elements corresponding to those up to the iron peak. The masses of the helium cores, manifest as helium stars or Wolf–Rayet stars if the stars lose their hydrogen envelopes during their evolution, lie in a wide range from ~ 3 to $\sim 50\ M_\odot$ [9–11]. According to the computations of [12], the masses of the iron cores before collapse lie in the interval from 1.25 to 2.05 $M_\odot$. The core masses depend monotonically on the initial masses of the stars on the main sequence. The masses of the remnants formed during collapse depend on many other factors (rotation, accretion onto the collapsed core, etc.), further exacerbating the problem of the observed narrowness of the distribution of known neutron-star masses.

In this way, evolutionary considerations do not provide any preconditions for the development of the mass distribution described above. What can we try to do to obtain such a distribution?

After completion of the nuclear evolution of a fairly massive ($>$8–10 $M_\odot$) star of the sort that could form a neutron star or black hole, there is a core collapse, which should be accompanied by the ejection of a shell, leading to a type II or Ib/c supernova. After the expansion of the shell, the compact remnant of the stellar collapse becomes observable. If the shell is ejected "efficiently" (i.e., it receives an energy of the order of the binding energy associated with the remnant), a low-mass compact remnant forms whose mass is of order the mass of the collapsed core of the supernova percursor.[1] If the shell is ejected inefficiently, a large transfer of matter from the envelope to the forming compact object is inevitable. As a result, the mass of the latter can substantially grow and approach the total mass of the supernova precursor (or at least make up an appreciable fraction of this mass).

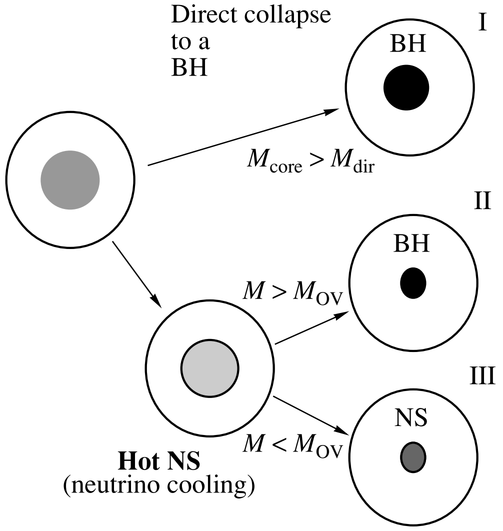

Fig. 2. Possible paths to the formation of compact remnants during core collapse in a massive star. Branch I corresponds to direct collapse into a black hole; branch II, to the formation of a black hole with a mass of the order of the mass of a proto-neutron star for the case $M > M_{OV}$; and, branch III to the formation of a neutron star.

If we suppose that the ejection of the envelope during the supernova explosion is sharply (even in a steplike manner) *weakened for supernova precursors with core masses above some threshold, the continuous sequence of precursor masses will be divided into two ranges. It remains to imagine how this might come about.*

2. THE CORE COLLAPSE

Let us consider various paths for the core collapse of a supernova precursor. The formation of a compact object during core collapse can occur in two ways. The first is via an intermediate stage with a hot proto-neutron star lasting several seconds or tens of seconds, in which there is intense radiation of thermal energy by the neutrino flux, after which the hot neutron star "cools" or, if its mass exceeds the Oppenheimer–Volkov limit for neutron matter M_{OV}, collapses into a black hole. The second is a direct collapse into a black hole, bypassing the metastable hot-neutron-star stage, if the star's mass exceeds some value $M_{dir} > M_{OV}$ (for a more detailed description of this process, see, for example, [13] and references therein). These possible paths for collapse are shown schematically in Fig. 2.

---

[1] A small transfer of mass onto the compact object is also possible, for example, from the inner regions of the shell.

As was shown by Bisnovatiĭ-Kogan [14], hot neutron stars can remain stable up to masses of the order of $\sim 7\ M_\odot$. In a dynamical collapse, the maximum mass of a hot neutron star is lower and depends substantially on the entropy per nucleon in the collapsing core material. In [14], the maximum mass somewhat exceeded $15\ M_\odot$, while, in modern computations, this value is lower, $\sim$(2–4) $M_\odot$ (see [15] and references therein).

If we make some assumptions as to the efficiency with which envelopes are ejected on each of the branches in Fig. 2, we can investigate the qualitative appearance of possible mass distributions for the resulting compact objects. Further, we will consider only the situation when the efficiency of the shell ejection does not increase with the mass of the collapsing core.

If a shell is efficiently ejected in any type of collapse (for all branches in Fig. 2), we will obtain a distribution in which neutron stars make a continuous transition to low-mass black holes. In this case, massive ($> 3\ M_\odot$) black holes do not form at all. The mass distribution for compact remnants expected in this case is shown in Fig. 3a.

If the shell ejection is weaker in the case of direct collapse to a black hole (branch I in Fig. 2) than in the formation of a proto-neutron star, then the lower two branches II and III form a continuous distribution of masses of neutron stars and low-mass black holes (as in the previous case). The maximum black-hole mass in this part of the distribution is roughly equal to the maximum possible mass of a hot neutron star formed via dynamical collapse $\sim M_{\rm dir}$ (taking into account the mass defect during its cooling and subsequent collapse). More massive black holes (with masses of the order of the mass of the supernova precursor) are separated from less massive black holes (which form after cooling of a hot proto-neutron star) by a gap, as is shown in Fig. 3b.

Note that, if for evolutionary reasons the masses of the collapsing cores do not exceed $M_{\rm dir}$, both of the cases considered cease to differ from each other and lead to the distribution depicted in Fig. 3a.

If shell ejection is efficient only on the lower branch III in Fig. 2, leading to the formation of neutron stars, then the upper branches I and II give birth to massive black holes separated from the neutron stars born via branch III by a gap (Fig. 3c). In this case, the distribution of neutron-star masses ends near $M_{\rm OV}$ (or somewhat lower, due to the mass defect)[2].

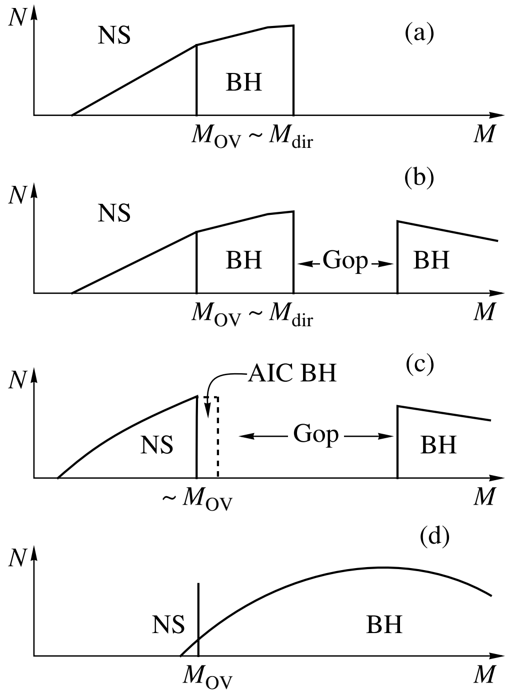

Fig. 3. Expected mass distribution for neutron stars and black holes in the following cases: (a) efficient shell ejection on all three branches I–III (shown in Fig. 2); (b) shell ejection during direct collapse to a black hole (branch I) that is weaker than that during the formation of a neutron star (branches II and III); (c) efficient shell ejection only during the formation of a neutron star (branch III), with branches I and II forming massive black holes, possibly with accretion-induced collapse (AIC) into black holes with masses $\sim M_{\rm OV}$; and (d) inefficient shell ejection for all collapse paths, so that neutron stars are either absent or are born only at the lower end of the mass distribution.

Finally, if the envelope is always "poorly ejected," this will form a broad, continuous distribution of massive black holes, which may extend to the region of low-mass black holes, and neutron stars will either not be formed at all or will be born in small numbers only at the lower end of the mass distribution (Fig. 3d).

As we can see by comparing Figs. 1 and 3, only the third possibility, corresponding to Fig. 3c, is in agreement with observations.

3. SUPERNOVA MECHANISMS

Let us now consider various mechanisms of supernova explosions in accordance with the collapse schemes discussed in the previous section.

Currently, we know of three qualitatively different mechanisms for type-II supernova explosions. For a long time, it was thought that energy was transferred to the envelope via the neutrino flux from the

---

[2] The distribution of neutron-star masses cannot be broken off appreciably below $M_{\rm OV}$, since the initial distribution of masses for the collapsing cores is continuous and broad.

hot, compact remnant. This type of model has been studied for many years, and although it has proved successful in explaining supernova explosion, subsequent studies have placed doubt on earlier conclusions (see [16] and references therein). The neutrino mechanism cannot explain the ejection of the supernova shell in either the spherically symmetric or the axially symmetric (with rotation) case [17]. There is some hope that the situation can be saved by large-scale neutrino convection [18]. Intensive investigations in this direction are underway (these are three-dimensional, much more unwieldy problems; see, for example, [19] and references therein).

The main flux of neutrinos is emitted in the hot-neutron-star stage, which lasts about 10 s [20]. In a direct collapse, the hot stage is appreciably shorter (of the order of the dynamical time scale for the collapse) and, consequently, is less efficient. Thus, if a pure neutrino mechanism were to be responsible for a supernova explosion, we would have the second scenario, (Fig. 3b) in which the envelope's ejection is weakened in the case of direct collapse[3]. Note that the neutrino mechanism acts for any type of collapse and can act jointly with other mechanisms of supernova explosion.

Another mechanism [21] is associated with the division of the rapidly rotating collapsing stellar core into two parts, two neutron stars[4]. The parts of the binary core then approach one another, due to the emission of gravitational radiation until the component with the smaller mass (and larger size) fills its Roche lobe. Further, there is an exchange of mass until the mass of the smaller component reaches the lower limit for the mass of a neutron star (about 0.1 $M_\odot$), at which point there is an explosive deneutronization of the low-mass neutron star [22]. This additional release of energy fairly far from the center of the collapsing star can efficiently eject its envelope. This mechanism can act only for the most rapidly rotating supernova precursors.

The approach of the binary core up to its merging can last from several minutes to several hours (or days), i.e., appreciably longer than the hot neutron star can exist. Analysis of this scenario is complicated due to the presence of two components. If both parts of the core collapse into black holes (either via direct collapse or via a hot-neutron-star stage), this mechanism will not operate and a single massive black hole will form at the end of the approach. The mechanism can operate when at least one of the core parts is a neutron star. The second part can be either a neutron star or a black hole, with the black hole forming either directly after cooling of a hot neutron star or as a result of accretion-induced collapse (AIC) during the approach stage (in these cases, low-mass black holes are formed). Thus, we have the situation shown in Fig. 3b, although the gap in the mass distribution can be blurred due to the large number of degrees of freedom in this model (primarily due to the unknown rate of accretion of the envelope material in the approach stage), thus leading to a situation that is closer to that shown in Fig. 3d.

The last of the supernova scenarios we will consider, magneto-rotation, was first proposed by Bisnovatyĭ-Kogan in 1970 [23]. The idea behind this mechanism is very simple: the envelope is ejected by the magnetic field of a rapidly rotating compact object. The shell is accelerated by the slowing of the neutron star's rotation. Since this simple idea involves the generation and amplification of magnetic fields, as well as complex three-dimensional hydrodynamics that are strongly affected by radiative transport, realistic computations of this scenario are extremely difficult. The results of two-dimensional computations [24] indicate that the magneto-rotational mechanism can transform several percent of the rotational energy of the compact object into kinetic energy of the shell and that the magneto-rotational outburst (the stage in which there is appreciable acceleration and ejection of the shell) lasts 0.01–0.1 s. However, this stage is preceded by a "winding" phase, when the toroidal magnetic field is linearly amplified to the critical value at which the shell is ejected ($\sim 10^{16}$–10^{17} G). The duration of this stage depends on the initial magnetic field of the neutron star and the velocity of its rotation and can vary from a fraction of a second up to minutes (or even hours). This mechanism requires that the neutron star's rotation be fairly rapid (corresponding to periods of the order of several milliseconds) but not as rapid as required by the core-division mechanism.

Depending on the relationship between the time required for amplification of the magnetic field t_B and the time for required (neutrino) cooling of the hot neutron star t_ν, we will obtain different mass distributions for the compact objects. During direct collapse into a black hole, the magnetic field does not have time to be amplified and there is no shell ejection. The magneto-rotational mechanism always operates in the lower branch of Fig. 2, and the shell is ejected most efficiently. On the middle branch, where the mass of the hot neutron star exceeds $M_{\rm OV}$, there will only be a magneto-rotational outburst when $t_B < t_\nu$. Otherwise ($t_B > t_\nu$), the neutron star cools and collapses before the magnetic field has time to become

---

[3] Currently, this scenario makes possible only the case in Fig. 3c (when there is virtually no shell ejection) and, in principle, in the presence of a very powerful neutrino flux, the case in Fig. 3a, neither of which is consistent with observations.

[4] At least the smaller part should form a neutron star.

sufficiently amplified, thus preventing ejection of the shell. When $t_B < t_\nu$, we obtain the distribution in Fig. 3b, and, when $t_B > t_\nu$, we obtain the distribution in Fig. 3c.

Thus, only a "pure" magneto-rotational mechanism (i.e., in which the magneto-rotational outburst occurs after cooling of the neutron star, since it requires $t_B > t_\nu$) can, in principle, reproduce the observed mass distribution for compact objects without invoking additional assumptions.

Here, we should comment on the anisotropy of supernova explosions that is implied by the observed high spatial velocities of neutron stars, 300–500 km/s (see, for example, [25–28]. Recent Chandra X-ray observations of a two-sided asymmetric ejection along the direction of motion of the young Vela pulsar [29] show that the spatial velocity of this neutron star is directed nearly along its rotational axis. There is also evidence that the spatial velocity and rotational angular momentum of the Crab pulsar are coincident [27]. This can be explained by the "rocket" mechanism of acceleration of motion [30], which requires a noncentral location for the magnetic dipole moment of the neutron star. Another natural explanation for the coaxiality of the rotation and spatial velocity of pulsars is provided by the magneto-rotational scenario for the ejection of the supernova shell, due to the asymmetry of the magnetic field in the presence of differential rotation of the collapsing core [31]. In the alternative supernova mechanism of Imshennik [21] for the explosion of a low-mass component, the second neutron star also acquires a high spatial velocity [32]; however this velocity should be *perpendicular to* the rotational axis of the neutron star; this is in contradiction with existing observations.

We emphasize that it is not currently known precisely which of these mechanisms (or which combination of them) gives rise to the ejection of the shells of type II supernovae in reality. It is clear that the neutrino mechanism alone is insufficient. Magneto-rotational core collapse and division of a rotating core have not been studied well enough for us to be able to use them to fully describe a supernova explosion and place constraints on the parameters of the precursor (except for the presence of rapid rotation).

4. THE OPPENHEIMER–VOLKOV LIMIT

Let us see what is required for the mass distribution for compact objects shown schematically in Fig. 3c to be consistent with observations. The broad distribution of black-hole masses does not place any constraints on the core collapse in the scheme we considered above. The wide gap from 1.5–3 $M_\odot$ between the masses of neutron stars and black holes indicates that no fewer than several solar masses from the envelope should fall onto the compact object during an inefficient shell ejection. However, this effect does not fit into our scheme and places appreciable constraints on the process for the formation of black holes during the collapse.

In contrast to the black-hole masses, the neutron-star mass distribution enables us to draw certain interesting conclusions. We will consider the narrowness of the neutron-star mass distribution in the following section. Here, we note that, according to Fig. 3c, the upper limit of the distribution of observed neutron-star masses should be close to the Oppenheimer–Volkov limit. It follows from this fact and from the observational data discussed in the introduction that

$$M_{\rm OV} \simeq (1.4 - 1.5)M_\odot.$$

This maximum neutron-star mass is possible in the case of certain soft equations of state for the neutron material, which are currently considered acceptable. Examples are the equations of state GS1, PAL6, and PCL2 from the work of Lattimer and Prakash [33].

5. NARROWNESS OF THE NEUTRON-STAR MASS DISTRIBUTION

Thus, we can explain (1) the broad distribution of black-hole masses, (2) the gap between the masses of neutron stars and black holes, and (3) the sharp upper limit for the distribution of neutron-star masses. It remains to be explained why the masses of neutron stars lie in such a narrow interval ($\Delta M/M \sim 10\%$).

We would normally expect a fairly rapid growth in the number of neutron stars as their masses decrease as a direct consequence of the monotonic dependence of the mass of the neutron star on the mass of the percursor core and also of the observed initial mass distribution for stars (a Salpeter law $dN \propto M^{-2.35}dM$), which rises sharply toward smaller masses. Precisely this type of behavior is observed for white dwarfs but not for neutron stars. *How can we explain this?*

According to theoretical calculations, the minimum mass of a proto-neutron star during the collapse of an iron core is (0.9–1.2) $M_\odot$ [34, 35]. Thus, if the collapse occurs without any additional accretion, the resulting interval of neutron-star masses will be from ~(0.9–1.2) $M_\odot$ to ~(1.4–1.5) $M_\odot$ ($M_{\rm OV}$), i.e., appreciably broader than is observed.

What will change if a *small* fraction of the envelope falls onto the collapsing core? (We will consider this question in more detail in the following section.)

The answer is obvious: accretion leads to an increase in the neutron-star mass and, consequently, to a narrowing of the overall mass distribution. The lower limit of the mass interval is increased, while the upper limit does not change, since neutron stars with masses exceeding $M_{\rm OV}$ become black holes and are no longer observed as neutron stars. An accretion of (0.1–0.2) $M_\odot$ is sufficient to turn the broad theoretical interval of neutron-star masses (see above) into the observed narrow interval. We must be careful here, however, since a quite small further increase in the accretion (for example, to (0.2–0.4) $M_\odot$) could result in a complete absence of neutron stars.

6. ACCRETION ONTO A NEUTRON STAR

Our analysis would not be complete without a discussion of the influence of accretion processes on the resulting mass distribution. Such processes can occur either immediately after the supernova explosion, with part of the stellar envelope falling onto the newly formed compact remnant, or at appreciably later evolutionary stages, when the rotation of the magnetized compact object becomes slow enough for the surrounding matter to reach its surface.

In the latter case, for single neutron stars and systems of two compact objects, accretion can occur only from the interstellar medium. If the maximum accretion rate in this regime is $\sim 10^{10}$–10^{11} g/s over the Hubble time ($\sim 10^{10}$ yr), the mass of the star grows by no more than several hundredths of a solar mass. Accretion in binary systems is much more efficient, even for characteristic accretion rates of $\sim 10^{-9}$–10^{-10} $M_\odot$/yr in low-mass systems, and neutron stars can increase their mass by several tenths of a solar mass.

The growth of the mass of the forming neutron star can come about due to a so-called “fall-back” of part of the supernova shell onto the compact objects (see, for example, [35, 36]). We distinguish this process from the incomplete shell ejection considered above since it occurs after the other physical processes of interest to us (cooling of the hot neutron star, the magneto-rotational outburst, etc.) and influences the mass distribution for compact objects after it has already formed (according to one of the schemes in Fig. 3).

It is probably impossible to exclude the fall-back of part of the ejected shell in any of the supernova models. In the neutrino model, the inner layers of the ejected shell have the lowest speeds[5], which could prove to be lower than the escape velocity for the innermost parts of the shell. In the magneto-rotational mechanism, the magnetic field accelerates material primarily in the plane of the rotational equator; this effect is much less powerful or even absent along the poles. Up to the magneto-rotational outburst, the amplified magnetic field hinders the infall of material onto the core but accretion becomes possible after this outburst. The question of accretion onto compact objects in supernova models with core division remains open. Current estimates of the efficiency of fall-back accretion in supernovae diverge strongly [35, 36].

Accretion acts on the mass distribution for compact object as follows: the masses of neutron stars grow, and some of them cross over $M_{\rm OV}$ and collapse into back holes (forming accretion-induced black holes). The number of remaining neutron stars decreases, while their mean mass grows. The lower limit for the observed mass distribution also rises, with this effect being more clearly expressed the weaker the variation in the mass increase ΔM from star to star. If ΔM is sufficiently high for all neutron stars (see previous section), then all neutron stars collapse into black holes.

Accretion-induced black holes have masses that slightly (by roughly ΔM) exceed $M_{\rm OV}$ and differ in no way from other low-mass black holes. Therefore, it is possible to detect such black holes only in the third scenario (Fig. 3c), which is of the most interest to us and which suggests that low-mass black holes do not form directly in the course of supernova explosions. Even if there is no fall-back of part of the supernova shell, the formation of accretion-induced black holes is expected in X-ray binary systems. The distribution of accretion-induced black holes is shown schematically in Fig. 3c by the dashed curve. It stands to reason that, if $\mathrm{M_{dir}}$ slightly exceeds the limiting mass $M_{\rm OV}$, the post-accretion distribution in Fig. 3c will differ only slightly from the distribution in Fig. 3b.

Accretion onto black holes increases their mass, but it is impossible to detect this effect, since we have essentially no information on their initial mass distribution.

7. DISCUSSION

Let us summarize the suppositions required to obtain the observed mass distribution for compact objects presented in Fig. 1 and the consequences following from them.

---

[5] The distribution of speeds in the ejected supernova shell is close to that for an isotropic expansion regime ($v \propto r$).

Suppositions	Consequences
(1) The magneto-rotational mechanism for supernova explosions with $t_B > t_\nu$.	(1) Only two groups of compact objects form directly during core collapse: NS and massive BH. Low-mass BH can form only later, after accretion.
(2) In inefficient shell ejection; no less than (1.5–2) $M_\odot$ falls onto the core.	(2) Explains the broad gap between the distributions of masses for NS and BH.
(3) $M_{\rm OV} \simeq (1.4\text{–}1.5)\ M_\odot$.	(3) Explains the observed lack of NS with masses larger than 1.45 $M_\odot$. Corresponds to the softest equations of state for NS.
(4) Must be a mechanism for increasing the lower limit for NS masses. For example, accretion of $\sim (0.1\text{–}0.2)\ M_\odot$ from the shell.	(4) Explains the narrowness of the observed distribution of NS masses.

We note the following in connection with these suppositions.

In point (2), the lower limit $M_{\rm BH}$ need not be sharp. The existence of a small number of black holes in the gap does not contradict our model.

In point (4), the mechanism for narrowing the neutron-star mass distribution need not be the one we have considered. If a different mechanism operates, this will not change our conclusions.

The question arises as to whether it is possible to reject soft equations of state for neutron matter with a lower mass limit for neutron stars of $M_{\rm OV} \simeq 1.4\ M_\odot$. First, the observed constraints on the equations of state (for example, from the $R(M)$ dependence derived from observations of X-ray bursters [38], cooling of single neutron stars [37], etc.) are not strong due to their large uncertainties. The situation is complicated by the similarity in the masses of known neutron stars, which makes it difficult to chose a single equation of state from amongst the large number of possibilities, since measurements of the mass–radius dependence have, so far, effectively been made at a single point.

Second, let us consider the minimum periods of millisecond pulsars whose rotation have been wound up by accretion in binary systems. Currently, the shortest known period is 1.56 ms, while soft equations of state allow neutron stars to achieve periods to 0.7 ms. Population syntheses for millisecond pulsars for various equations of state were recently conducted in [39]. The results indicate that we should observe a significant percent of submillisecond pulsars with masses to 1.7 $M_\odot$. Such pulsars can be detected in modern observations [40], but thus far they have not been. On the other hand, in order for the rotation to be wound up to the shortest allowed periods, the neutron star must accrete a fairly large amount of material, about (0.1–0.15) $M_\odot$ [41, 42]. At the value $M_{\rm OV} \approx 1.5\ M_\odot$, the neutron star can reach the limiting mass and collapse into a black hole before its rotation increases to the shortest allowed period. Could it be that the very absence of millisecond pulsars with the shortest allowed periods can be explained by the Oppenheimer–Volkov limit and the observed mass distribution for neutron stars?

Black holes that form as a result of the accretion-induced collapse of neutron stars cannot fill the mass gap. In accretion in massive binary systems, the growth in the mass of the accreting compact object is limited by the Eddington luminosity, $10^{38}(M_x/M_\odot)$ erg/s, which is already reached at accretion rates of $\dot{M}_{\rm cr} \simeq 10^{-8}\ M_\odot$/yr, while the accretion time is $\lesssim 10^6$ yr. In low-mass binaries, realistic accretion rates are $\dot{M} < 10^{-8}\ M_\odot$/yr, implying a total mass for the star $\lesssim 1\ M_\odot$.

However, if the collapse of neutron stars into black holes occurs during the accretion of matter in binary systems, with our supposition that $M_{\rm OV} \approx 1.5\ M_\odot$, we would expect the presence of black holes with low masses (of the order of 1.6 $M_\odot$) in low-mass X-ray sources. Currently, we know of X-ray sources in which there are no obvious signs of the presence of neutron stars (periodic pulsations, X-ray bursts), primarily X-ray transient sources. One means to distinguish the type of compact object present in an accretion binary is through analysis of quasi-periodic oscillations (QPOs) in the X-ray emission. However, observations of rapid variability by the RXTE satellite [43, 44] indicate that the properties of QPOs for sources in the low (hard) state that are neutron stars and black holes are very similar. There remains the possibility of distinguishing sources based on their high-frequency power spectra [45], if kilohertz QPOs are associated with the presence of a surface on the compact object. However, in the case of black holes with masses of the order of 1.5 $M_\odot$, high-frequency QPOs could also be generated in the inner parts of the accretion flow; thus, it is difficult to distinguish neutron stars and black holes based only on their power spectra.

Note that all the arguments we have presented could be decisively changed if the compact objects we observe are in reality so-called "strange" stars.

8. CONCLUSIONS

Using our two suppositions that

(a) the supernova shell is ejected via the magneto-rotational mechanism, with the time for amplification of the field exceeding the time for cooling of the hot proto-neutron star ($t_B > t_\nu$) and

(b) the maximum mass for neutron stars (the Oppenheimer–Volkov limit) is close to $M_{\rm OV} \simeq 1.5\ M_\odot$, we can explain important features in the observed mass distribution for compact objects: the narrowness of the distribution of measured neutron-star masses $M_{\rm NS} = 1.35 \pm 0.15\ M_\odot$ and the lack of compact remnants with masses between 1.5 and 3 $M_\odot$.

Our analysis enables us to make the following additional predictions, which are subject to verification with astrophysical observations.

(1) There should exist black holes with masses of about 1.5 $M_\odot$ (single or in binaries). They could be detected in low-mass X-ray transient sources.

(2) As a consequence of magneto-rotational supernova outbursts with the formation of neutron stars, there could be a correlation between the power of supernova explosions and the presence of neutron stars: neutron stars should be associated with the most powerful supernova remnants. Black holes are expected in remnants with lower energies. In addition, supernova remnants with neutron stars and black holes could differ in their geometrical form and dynamics of explosive expansion.

(3) Coaxiality of the rotational angular momentum of a pulsar and its spatial velocity, as observed for the Vela and Crab pulsars, should be a general property of all radio pulsars.

Confident detections of neutron stars with masses appreciably exceeding 1.4 $M_\odot$ would represent a direct refutation of our proposed ideas. Currently, there are no reliable determinations of such masses for neutron stars.

9. ACKNOWLEDGMENTS

The authors thank D. G. Yakovlev, A. M. Cherepashchuk, T. A. Lozinskaya, S. I. Blinnikov, and P. Höfling for useful discussions. This work was partially supported by the Russian Foundation for Basic Research (project nos. 99-02-16205 and 00-02-17164).

REFERENCES

1. S. Thorsett and D. Chakrabarty, Astrophys. J. **512**, 288 (1999).
2. A. M. Cherepashchuk, Usp. Fiz. Nauk **166**, 809 (1996) [Phys. Usp. **39**, 753 (1996)].
3. A. M. Cherepashchuk, Space Sci. Rev. **93**, 473 (2000).
4. J. A. Orosz and E. Kuulkers, Mon. Not. R. Astron. Soc. **305**, 132 (1999).
5. M. H. van Kerkwijk, J. van Paradijs, and E. J. Zuiderwijk, Astron. Astrophys. **303**, 497 (1995).
6. D. Stickland, C. Lloyd, and A. Radzuin-Woodham, Mon. Not. R. Astron. Soc. **296**, L21 (1997).
7. S. R. Heap and M. F. Corcoran, Astrophys. J. **387**, 340 (1992).
8. G. E. Brown, J. C. Weingartner, and R. A. M. J. Wijers, Astrophys. J. **463**, 297 (1996).
9. A. M. Cherepashchuk, Astron. Zh. **78**, 145 (2001) [Astron. Rep. **45**, 120 (2001)].
10. A. M. Cherepashchuk, in *Proceedings of the International Conference in Honour of Professor A. G. Massevitch "Modern Problems of Stellar Evolution," Zvenigorod–Moscow, 1998,* Ed. by D. S. Wiebe, p. 198.
11. A. M. Cherepashchuk, in *Proceedings of the Tartu Workshop on Thermal and Ionization Aspects of Flows from Hot Stars: Observations and Theory, Tartu, 2000,* Ed. by H. J. G. L. M. Lamers and A. Sapar, p. 249.
12. F. X. Timmes, S. E. Woosley, and T. A. Weaver, Astrophys. J. **457**, 834 (1996).
13. M. Prakash, J. M. Lattimer, J. A. Pons, *et al.*, astro-ph/0012136 (2000).
14. G. S. Bisnovatyĭ-Kogan, Astrofizika **4**, 221 (1968).
15. K. Strobel and M. K. Weigel, Astron. Astrophys. **367**, 582 (2001); astro-ph/0012321.
16. A. Mezzacappa, A. C. Calder, S. W. Bruenn, *et al.*, Astrophys. J. **493**, 848 (1998); **495**, 911 (1998).
17. H.-Th. Janka, Astron. Astrophys. **368**, 527 (2001).
18. M. Herant, W. Benz, J. Hix, *et al.*, Astrophys. J. **435**, 339 (1994).
19. K. Kifonidis, T. Plewa, H.-Th. Janka, and E. Müller, in *Proceedings of the International Workshop on Astronomy with Radioactivities, Schloss Ringberg, Kreuth, Germany, 1999,* Ed. by R. Diehl and D. Hartmann (Max-Planck-Institut für Extraterrestrische Physik, Garching, 1999), p. 141 (http://www.gamma.mpe-garching.mpg.de/mpeteam/workshop/proceedings.html).
20. D. K. Nadezhin and I. V. Otroshenko, Astron. Zh. **57**, 78 (1980) [Sov. Astron. **24**, 47 (1980)].
21. V. S. Imshennik, Pis'ma Astron. Zh. **18**, 489 (1992) [Sov. Astron. Lett. **18**, 194 (1992)].
22. S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, and A. G. Polnarev, Pis'ma Astron. Zh. **10**, 422 (1984) [Sov. Astron. Lett. **10**, 177 (1984)].
23. G. S. Bisnovatyĭ-Kogan, Astron. Zh. **47**, 813 (1970) [Sov. Astron. **14**, 652 (1971)].

24. N. V. Ardeljan, G. S. Bisnovatyi-Kogan, and S. G. Moiseenko, in *The Local Bubble and Beyond (IAU Colloquium 166),* Ed. by D. Breitschwerdt, M. J. Freyberg, and J. Trümper, Lect. Notes Phys. **506**, 145 (1998); Astron. Astrophys. **355**, 1181 (2000).
25. A. G. Lyne and D. R. Lorimer, Nature **369**, 127 (1994).
26. D. R. Lorimer, M. Bailes, and P. A. Harrison, Mon. Not. R. Astron. Soc. **289**, 592 (1997).
27. D. Lai, D. Chernoff, and J. M. Cordes, Astrophys. J. **549**, 1111 (2001); astro-ph/0007272.
28. C. M. Olbert, C. E. Clearfield, N. E. Williams, *et al.*, Astrophys. J. **554**, 205 (2001); astro-ph/0103268.
29. G. G. Pavlov *et al.*, Bull. Am. Astron. Soc. **32**, 733 (2000).
30. E. R. Harrison and E. Tademaru, Astrophys. J. **201**, 447 (1975).
31. G. S. Bisnovatyĭ-Kogan and S. G. Moiseenko, Astron. Zh. **69**, 563 (1992) [Sov. Astron. **36**, 285 (1992)].
32. V. S. Imshennik and D. V. Popov, Pis'ma Astron. Zh. **24**, 251 (1998) [Astron. Lett. **24**, 206 (1998)].
33. J. M. Lattimer and M. Prakash, Astrophys. J. **550**, 426 (2001); astro-ph/0002232.
34. J.-O. Goussard, P. Haensel, and J. L. Zdunik, Astron. Astrophys. **330**, 1005 (1998).
35. R. A. Chevalier, Astrophys. J. **346**, 847 (1989).
36. M. Herant, W. Benz, and S. Colgate, Astrophys. J. **395**, 642 (1992).
37. D. G. Yakovlev, K. P. Levenfish, and Yu. A. Shibanov, Usp. Fiz. Nauk **169**, 825 (1999).
38. L. G. Titarchuk, Astrophys. J. **429**, 340 (1994).
39. A. Possenti, M. Colpi, U. Geppert, *et al.*, Astrophys. J., Suppl. Ser. **125**, 463 (1999).
40. N. D'Amico and L. Burderi, in *Pulsar Timing, General Relativity and the Internal Structure of Neutron Stars,* Ed. by Z. Arzoumanian, F. van der Hooft, E. P. J. van den Heuvel (Publ. Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam, 1999), p. 129.
41. V. M. Lipunov and K. A. Postnov, Astrophys. Space Sci. **106**, 103 (1984).
42. L. Burderi, A. Possenti, M. Colpi, *et al.*, Astrophys. J. **519**, 285 (1999).
43. D. Psaltis, T. Belloni, and M. van der Klis, Astrophys. J. **520**, 262 (1999).
44. R. Wijnands and M. van der Klis, Astrophys. J. **514**, 939 (1999).
45. R. Sunyaev and M. Revnivtsev, Astron. Astrophys. **358**, 617 (2000).

Translated by D. Gabuzda

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 908–921. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 1034–1048.
Original Russian Text Copyright © 2001 by Zharov, Pasynok.

Theory of Nutation of a Nonrigid Earth

V. E. Zharov and S. L. Pasynok

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

Received October 5, 2000

Abstract—A general theory of terrestrial nutation is proposed assuming that the Earth is made up of four envelopes (atmosphere, mantle, fluid core, and solid core) and taking account of all important forces (viscous, electromagnetic, etc.). A theory for the effect produced on the Earth's nutation by viscous forces in the fluid core is developed based on experimental data on the viscosity of molten iron under pressure. The proposed theory predicts nutation in longitude and inclination with an rms deviation of 0.35 milliarcseconds. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The axis of the Earth's figure, instantaneous rotational axis, and angular-momentum axis move with respect to an inertial reference frame due to the gravitational forces of the Sun, Moon, and planets. This is called precessional–nutational motion. The precessional period is $\sim 26\,000$ yrs. Over the time during which telescope observations have been available, ~ 300 yr, we can assume that the precession forms a linear (or secular) motion of the axis with the nutation harmonics superimposed on it. The orbits of the Earth and Moon, as well as those of other planets revolving around the Sun, determine the periods of the nutational harmonics. The period of the main harmonic is 18.6 yr and is related to the period of motion of the nodes of the Moon's orbit. The maximum amplitude of the nutation is $\sim 9''$. The remaining nutational harmonics have smaller amplitudes.

The Earth's nutation is an elliptical motion, since it is the sum of two circular motions with equal periods but different amplitudes and opposite directions. We shall call counterclockwise and clockwise motions observed from the north celestial pole "forward" and "reverse" motions, respectively. Another representation of the nutation is a decomposition into two components: nutation in longitude and in inclination. The relationship between these two representations is discussed in the appendix. We shall take the periods (or frequencies) of the nutational harmonics for reverse motions to be negative and those of forward motions to be positive.

The traditional method for constructing a theory of the Earth's nutation is the following. An amplitude–frequency transformation function is calculated for a given model of the Earth, and this function is multiplied by the amplitudes of nutational harmonics computed for a completely rigid Earth. Then, corrections are added taking into account the difference between the model and real Earth. Using this method, Wahr [1, 2] obtained nutational series composed of 106 harmonics constructed for an elastic, ellipsoidal Earth with a fluid outer core and elastic mantle. In 1980, this series was accepted by the International Astronomy Union (IAU) as a standard for astronomical computations [3]. The model determines the normal modes (i.e, resonance properties) of the Earth. The model employed by Wahr [1, 2] defines three normal modes: the Chandler wobble (CW), nearly diurnal nutation (FCN), and a tilted mode (TOM). The difference between the model and real Earth leads to changes in the frequencies of the normal modes (and to the appearance of new modes), as well as to changes in the amplification. This is especially important when the frequencies of the lunar–solar potential are close to the frequencies of normal modes.

Beginning in the early 1980s, regular radio observations using very long baseline interferometry (VLBI) were begun. At present, VLBI observations make the most important contribution to the determination of nutational motions [4]. For most harmonics, the accuracy of VLBI measurements of nutation amplitudes is no less than 0.1 milliarcsecond (mas). The results of 18 years of VLBI observations have introduced some corrections to the IAU1980 theory [4]. The differences between the observations and theory are within ±15 and ±5 mas for nutation in longitude and inclination, respectively. These differences appreciably exceed the observational errors, thus verifying the reliability of these results.

Improvements in the accuracy of observations have shown that Wahr's [1, 2] nutation theory must be replaced by an improved theory. In 1994, the XXII IAU General Assembly (The Hague, Netherlands) created a working group on the "Nonrigid Earth

1063-7729/01/4511-0908$21.00 © 2001 MAIK "Nauka/Interperiodica"

Nutation Theory." The main goal of this working group was to develop a new nutation theory that could be used to determine the Earth's position in space with an accuracy to better than one milliarcsecond. The report of the working group [5] considers, in detail, the problems that must be solved to construct a new theory. The working group included effects produced on the nutation by the fluid core, oceans, and atmosphere among the most poorly modeled phenomena. We accordingly focus on these effects.

The basis of our theory is the analytical approach of [6], which can be outlined as follows. We first write a system of equations for the angular momenta of the whole Earth and the envelopes included in the model (in [7], a fluid core and solid core), except for the mantle. We obtain a system of algebraic equations in the frequency domain. The solution of the homogeneous system yields the frequencies and amplitudes of the normal modes. The solution of the nonhomogeneous system is represented as the product of the transformation function and the nutation amplitudes for a completely rigid Earth. We then add corrections taking into account additional effects, such as inelastic dissipation in the mantle and the effect of the oceans. The unknown internal parameters of the Earth (such as the compression of the core–mantle boundary) are determined through fitting to obtain the best agreement between theory and observation.

All Earth-nutation theories can be divided into two groups: empirical theories and theories based on solutions of equations of rotational motion. Empirical theories derive the coefficients of the transformation function through fitting to obtain the best agreement between theory and observation. The SF2000 theory [7] and the theory of Herring [8] are examples of this type of theory. Theories based on solutions of rotational equations are, in turn, divided into numerical, analytical, and semi-analytical categories. Numerical theories are based on numerical integration of equations and analytical theories on analytical solutions of the equations of rotational motion, with the Earth's internal parameters specified by some model for the Earth's internal structure [10]. However, these models cannot completely determine all of the internal parameters of the Earth with sufficient accuracy. Therefore, semi-analytical theories, which estimate unknown internal parameters based on the best agreement between theoretical nutation amplitudes and observations [6, 10], are most commonly used. Our theory is semi-analytical.

The main distinctions of our theory from previous semi-analytical theories are the following:

(1) the atmosphere is included in the momentum equations simultaneously with the solid and fluid cores, yielding a six-dimensional system of equations for the complex amplitudes;

(2) the viscosity of the fluid core is taken into account, which, together with account of the magnetic field, leads to a magneto-viscosity tensor;

(3) atmospheric tides are taken into account.

Section 2 presents five different models for the Earth for which we have computed nutational harmonics, while Section 3 describes the method used to take into account the atmosphere, viscosity, and magnetic field. Section 4 considers the results of calculations of the nutational harmonics for the given models, and Section 5 discusses our results.

2. MODELS FOR THE EARTH

Each model assumes that the Earth rotates and is ellipsoidally stratified.

Model A includes the atmosphere, mantle, fluid core, and solid core. This model takes into account torque due to viscous forces and neglects that due to magnetic forces. We adopt the corrections to the inelasticity of the mantle and the oceans, as well as the rigid-body nutation amplitudes, from [11]. The model fits the compression of the core–mantle boundary in order to match the real part of the calculated amplitude of the annual reverse nutation with the observed amplitude. The viscosity is varied to obtain the best agreement with the observations. The remaining parameters are taken from the PREM model [12] (presented in [11]), except for the dynamical compression of the whole Earth, which is taken to be $e = 0.003\,284\,915$ [11]. We adopt the atmospheric parameters from [13, 14].

Model B is the same as Model **A** except for the elimination of the atmosphere as one of the Earth's envelopes.

Model C is the same as Model **A** except for the elimination of the solid core. The fluid core occupies the entire space inside the mantle.

Model D is the most complete. The Earth consists of the atmosphere, mantle, fluid core, and solid core. The Earth's core is modeled taking into account electromagnetic and viscous forces. In the simplest case, the geophysical dynamical compression of the whole Earth is taken from [6] and is $e = \dfrac{C - A}{A} = 3.284\,520\,155\,008 \times 10^{-3}$. To fit the theoretical nutation amplitudes to the observed amplitudes, we vary the independent components of the magneto-viscosity matrix S_{ab} and the elasticity parameters γ and κ. The nutation amplitudes are calculated for various nutation series for the rigid Earth, taking into account corrections for dissipation in the mantle (calculated for various dissipation models) and for the oceans.

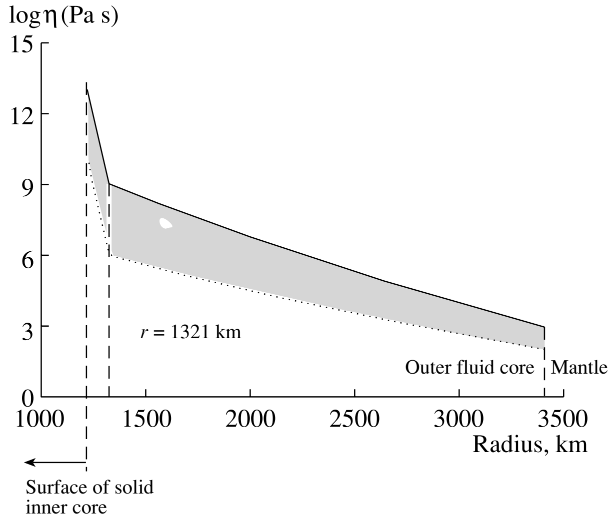

Fig. 1. The viscosity of the Earth's fluid core as a function of radius.

Model E is equivalent to Model **A** except that a rigid-body series is used, with corrections taken from Model **D** for the inelastic mantle dissipation and the oceans.

3. NUTATIONAL EFFECTS OF THE LIQUID CORE VISCOSITY, ATMOSPHERE, AND ELECTROMAGNETIC FORCES

3.1. Effects of Viscosity

Nutation theory assumes that the Earth rotates with an average angular velocity Ω_0. Each envelope of the Earth is spherically stratified; that is, the hydrostatic equilibrium produced by the gravitational and centrifugal forces determines surfaces of constant density. The mantle is assumed to be elastic.

Our computation of the Earth's nutation was similar to that performed in [15, 6]. The main dynamical equations in the theory of [15] are those of conservation of angular momentum, with the first of these applied to the whole Earth and the second to the fluid core. The more complicated theory presented in [7] takes into account the rotation of the solid inner core and uses two additional equations for the temporal evolution of the angular-velocity vector and the axis of the figure of the solid core.

There are four coordinate systems, each with its origin at the Earth's center of mass, introduced to describe the Earth's rotation and motions of the envelopes. The Cartesian coordinate system of basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is inertial, with $\mathbf{e}_3$ directed along the average rotational vector of the Earth; the nutation $\mathbf{I}$ coordinate system rotates (with respect to the inertial system) with constant angular velocity $\boldsymbol{\Omega}_0$ about the axis $\mathbf{I}_3 = \mathbf{e}_3$. The axes $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ of the terrestrial coordinate system are associated with the Tisserand axes for the mantle. In this case, the deformations produced in the mantle by exciting potentials do not contain any rotations [17]. The transformation from the $\mathbf{I}$ system to the $\mathbf{i}$ system is realized by a rotation of the mantle. The $\mathbf{i}'$ coordinate system, which is also associated with Tisserand axes, is used for the solid core.

The free rotation of the solid core implies that the $\mathbf{i}$ and $\mathbf{i}'$ axes do not (generally) coincide. The inclination of the solid core is determined by the unit vector $\mathbf{n}_s = \mathbf{i}'_3 - \mathbf{i}_3$.

We assume that the Earth's instantaneous angular velocity $\boldsymbol{\Omega}$ is related to the angular velocities of the fluid $\boldsymbol{\Omega}_f$ and solid $\boldsymbol{\Omega}_s$ cores as follows:

$$\begin{aligned} \boldsymbol{\Omega} &= \boldsymbol{\Omega}_0 + \mathbf{w} = \Omega_0(\mathbf{i}_3 + m) \\ \boldsymbol{\Omega}_f &= \boldsymbol{\Omega} + \mathbf{w}_f = \Omega_0(\mathbf{i}_3 + m + m_f) \\ \boldsymbol{\Omega}_s &= \boldsymbol{\Omega} + \mathbf{w}_s = \Omega_0(\mathbf{i}_3 + m + m_s). \end{aligned} \tag{1}$$

Here, the dimensionless vectors $\mathbf{m}, \mathbf{m}_f, \mathbf{m}_s$ are perturbations in the angular velocities of the envelopes produced by the lunar–solar, gravitational, and centrifugal potentials. The computation of the four unknown vectors $\mathbf{m}, \mathbf{m}_f, \mathbf{m}_s$, and $\mathbf{n}_s$ as functions of the exciting potentials represents a complete solution of the problem.

Here, we shall not study the evolution of the axial components of the vectors $\mathbf{m}, \mathbf{m}_f, \mathbf{m}_s$ associated with changes in the length of the day. It is convenient to represent the equatorial components of an arbitrary vector $\mathbf{m}$ as a complex number $\widetilde{m} = m_1 + im_2$. Further, symbols with a tilde will represent complex numbers formed from such combinations of equatorial components.

In a reference frame fixed to the mantle, the dynamical Euler equations take the form [6]

$$\frac{\partial \mathbf{H}}{\partial t} + \boldsymbol{\Omega} \times \mathbf{H} = \mathbf{L}, \tag{2}$$

$$\frac{\partial \mathbf{H}_f}{\partial t} - \boldsymbol{\omega}_f \times \mathbf{H}_f = 0, \tag{3}$$

$$\frac{\partial \mathbf{H}_s}{\partial t} + \boldsymbol{\Omega} \times \mathbf{H}_s = \mathbf{L}_s, \tag{4}$$

where $\mathbf{H}$, $\mathbf{H}_f$, and $\mathbf{H}_s$ are the angular momenta of the whole Earth, fluid core, and solid core; and $\mathbf{L}$, $\mathbf{L}_f$, and $\mathbf{L}_s$ are the torques applied to the whole Earth and its fluid and solid cores. This system contains equations for the momenta of the whole Earth, of the Earth's outer fluid core, and of the inner solid core.

We can transform system (2)–(4) into a system of differential equations in small vectors $\mathbf{m}$, $\mathbf{m}_f$, $\mathbf{m}_s$, and $\mathbf{n}_s$. The amplitude of the perturbation (m) of the angular velocity of the mantle's rotation does not exceed 4×10^{-8} [6] in the diurnal frequency range.

The main terms in (2)–(4) are first order in $m = |\mathbf{m}|$, or $O(m)$.

Consider first the effect produced on the Earth's nutation by the viscosity of the fluid core. Equation (2) for the whole Earth does not change, since the viscous forces are internal forces and cannot change the total angular momentum $\mathbf{H}$. Equations (3) and (4) are affected by the viscosity, as we will now verify.

For the fluid core, we can express the dynamical equations in the form [6]

$$\frac{\partial \mathbf{H}_f}{\partial t} - \boldsymbol{\omega}_f \times \mathbf{H}_f = -\int\limits_V \rho \mathbf{r} \times \mathbf{G} dV, \qquad (5)$$

where the force $\mathbf{G}$ is obtained by integrating the momentum equations for an ideal fluid in the reference frame used. We can assume that the right-hand side of (5) is zero to within $O(m\varepsilon^2)$, where ε is the Earth's geometrical compression. To take account of viscosity, we must add the torque due to viscous forces:

$$\mathbf{L}_f^{(\eta)} = \oint\limits_{S_f} \mathbf{r} \times \mathbf{p}_n^{(\eta)} dS, \qquad (6)$$

where the surface density of the viscous-stress force is

$$\mathbf{p}_n^{(\eta)} = \eta(r) \sum_{i,j=1}^{3} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) n_j \mathbf{e}_i. \qquad (7)$$

Here, $\eta(r)$ is the dynamical viscosity of the fluid core; $\mathbf{v} = \mathbf{v}(v_1, v_2, v_3) = \mathbf{v}^e + \mathbf{v}^v$ is the residual velocity describing deviations of the fluid velocity from uniform rotation due to elastic deformations ($\mathbf{v}^e$) and viscous friction ($\mathbf{v}^v$); and $\mathbf{n}$ is a normal to the surface of the fluid core.

We can similarly derive expressions for the torques applied to the mantle $\mathbf{L}_m^{(\eta)}$ and inner core $\mathbf{L}_s^{(\eta)}$. In accordance with the conservation of angular momentum, we obtain

$$\mathbf{L}_s^{(\eta)} + \mathbf{L}_m^{(\eta)} + \mathbf{L}_f^{(\eta)} = 0.$$

To calculate the torques $\mathbf{L}_m^{(n)}$ and $\mathbf{L}_s^{(n)}$, we must know the viscosity of the fluid core. Various estimates yield viscosities spread over 10–14 orders of magnitude: from 10^{-2} to 10^8 Pa s according to [17] and from 10^{-3} to 10^{11} Pa s according to [18]. Either small-scale or global vortices are possible, depending on the viscosity (for low and high viscosities, respectively). Various types of fluid motions yield various torques. The use of semi-analytical theories enables us to select the viscosity appropriate for the observed nutation.

In studying molten metals exposed to pressures reaching 10 GPa, Brazhkin [19, 20] showed that there is a significant growth in the viscosity along the melting curve. Extrapolating the results obtained for molten iron to the pressures and temperatures typical of the Earth's core suggests that the viscosity of the Earth's fluid core is rather high: from 10^1–10^3 Pa s at the mantle–core boundary to 10^7–10^{11} Pa s at the boundary between the outer and inner cores [19]. Figure 1 presents the model for the viscosity distribution we used for the Earth's fluid core. The solid and dashed curves in Fig. 1 indicate the region of permitted viscosity values proposed in [19].

It is important to our model that the viscosity significantly increases only in a rather thin layer (with a thickness of about 100 km) near the boundary between the solid and fluid cores. This is due to the fact that the viscous torque is proportional to the product of the surface area and the velocity gradient. Since the surface area of the core–mantle boundary exceeds the surface area of the solid core by about a factor of ten and the viscosity decreases exponentially with radius, we can neglect the integrals over the core–mantle boundary. Therefore, we assume that $\mathbf{L}_m^{(\eta)} \ll \mathbf{L}_s^{(\eta)}$, such that

$$\mathbf{L}_s^{(\eta)} = -\mathbf{L}_f^{(\eta)}. \qquad (8)$$

Further, we must derive the viscous stress tensor from (7) and substitute it into (6). Computation of the surface integral results in an expression for the viscous torque $\mathbf{L}_f^{(\eta)}$.

Thus, if we know the velocities $\mathbf{v}$, we can calculate via (6) the viscous torque $\mathbf{L}_f^{(\eta)}$ applied to the fluid core and then immediately use (8) to obtain the viscous torque $\mathbf{L}_s^{(\eta)}$ applied to the solid core. Using the right-hand sides of (3) and (4), including the viscous torques, we can solve the corrected system of equations (2)–(4) and obtain the nutation parameters.

Let us compute the velocities $\mathbf{v}$. To avoid complicated calculations, we should select a special reference frame whose angular velocity is taken to be the angular velocity of the fluid core. To eliminate the effect of $\mathbf{v}$, this reference frame was determined in [6] assuming the corresponding torque to be zero:

$$\int_V \rho \mathbf{r} \times \mathbf{v}^e dV = 0.$$

This defines the angular velocity of the fluid core as the angular velocity of the Tisserand axes associated with the core.

To determine the nonrotating portion of the velocities $\mathbf{v}$, we use the following heuristic reasoning. First, we assume that the effects on the nonrotating portion of the velocities produced by the viscosity and elastic deformations are independent of each other. As we will show below, the viscous torque is of order $O(m)$. In this case, the effect of the interaction of the elastic and viscous forces is of order $O(m^2)$. We neglect such small quantities. Therefore, we can define a reference frame rotating with the angular velocity of the fluid core by assuming that the torques attributed to the elastic and viscous portions of the nonrotating velocities are separately equal to zero. This reference frame becomes the Tisserand frame for the fluid core with an accuracy to $O(m^2)$.

Second, let us assume that the velocities $\mathbf{v}^v$ depend on the radial distance r from the center of the Earth as follows:

$$\mathbf{v}^v(r) = \boldsymbol{\Omega}(r) \times \mathbf{r}, \qquad (9)$$

$$\boldsymbol{\Omega}(r) = \boldsymbol{\Omega}_s + \frac{\boldsymbol{\Omega}' - \boldsymbol{\Omega}_s}{\delta r}(r - r_s) + O(\delta r^2),$$
$$r_s + \delta r > r > r_s,$$

where $r_s = 1221$ km is the radius of the solid core and $\delta r = 100$ km is the thickness of the layer of high viscosity. The smallness of this thickness compared to the radius of the solid core forms the basis for this approximation and enables us to use a series expansion in a small parameter. Here, $\boldsymbol{\Omega}(r)$ is the angular velocity of the fluid layer at a distance r from the center of the Earth. Thus, we assume that the fluid core consists of two components: a nonviscous core and a viscous layer with thickness δr at the boundary between the solid and fluid cores. Let $\boldsymbol{\Omega}_s$ be the angular velocity of the Tisserand reference frame for the solid core rotating with respect to the inertial frame and $\boldsymbol{\Omega}'$ be the angular velocity of rotation at the boundary of the viscous portion of the fluid core (the outer boundary of the viscous layer with thickness δr).

If the torque attributed to the nonrotating portion of the velocities is zero, we obtain

$$\boldsymbol{\Omega}' = \boldsymbol{\Omega}_f + O(m\varepsilon^2), \qquad (10)$$

where $\boldsymbol{\Omega}_f$ is the angular velocity of rotation of the fluid core (or the angular velocity of the Tisserand reference frame defined for the mantle). The velocities (9) combined with (10) yield

$$\frac{\partial v_i}{\partial x_j} = \partial_j \left(\sum_{p,k=1}^{3} \varepsilon_{ipk} \left(\Omega_{sp} + \frac{\Omega_{fp} - \Omega_{sp}}{\delta r}(r - r_s) \right) x_k \right)$$
$$= \sum_{p,k=1}^{3} \varepsilon_{ipk} \left(\Omega_{sp} + \frac{\Omega_{fp} - \Omega_{sp}}{\delta r}(r - r_s) \right)$$
$$+ \sum_{p,k=1}^{3} \varepsilon_{ipk} \left(\frac{\Omega_{fp} - \Omega_{sp}}{\delta r} \right) n_j x_k.$$

Substituting this into (7), we obtain

$$\mathbf{p}_n^{(\eta)} = -\eta(r_s) \sum_{i,j=1}^{3} \left(\sum_{p,k=1}^{3} \varepsilon_{ipk} \left(\frac{\Omega_{fp} - \Omega_{sp}}{\delta r} \right) n_j x_k \right.$$
$$\left. + \sum_{p,k=1}^{3} \varepsilon_{jpk} \left(\frac{\Omega_{fp} - \Omega_{sp}}{\delta r} \right) n_i x_k) \right) n_j \mathbf{e}_i$$

or

$$\mathbf{p}_n^{(\eta)} = -\eta(r_s) \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \times \mathbf{r}. \qquad (11)$$

For the inner core, we obtain a similar expression but of opposite sign (since an outward normal for the inner core is an inward normal for the fluid core and vice versa).

Substituting (11) into (6) yields

$$\mathbf{L}_f^{(\eta)} = -\eta(r_s) r_s^4 \int_0^{\pi} \int_0^{2\pi} \mathbf{n} \times \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \times \mathbf{n} \, d\phi \sin\theta d\theta.$$

We transform the integrand using the formula for the double vector product

$$\mathbf{n} \times \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \times \mathbf{n} = \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r}$$
$$- \mathbf{n} \left(\frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \mathbf{n} \right), \int_0^{2\pi} \mathbf{n} \times \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \times \mathbf{n} \, d\phi$$
$$= 2\pi \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} - \frac{\pi}{\delta r} \begin{pmatrix} (\Omega_{f1} - \Omega_{s1}) \sin^2\theta \\ (\Omega_{f2} - \Omega_{s2}) \sin^2\theta \\ 2\,(\Omega_{f3} - \Omega_{s3}) \cos^2\theta \end{pmatrix}.$$

Integration yields

$$\int_0^{\pi} \int_0^{2\pi} \left[\frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} - \mathbf{n} \left(\frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} \mathbf{n} \right) \right] d\phi \sin\theta d\theta$$
$$= \frac{8\pi}{3} \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r}.$$

Therefore, the viscous torque applied to the fluid core is

$$\mathbf{L}_f^{(\eta)} = -\eta(r_s) \frac{8\pi}{3} \frac{\boldsymbol{\Omega}_f - \boldsymbol{\Omega}_s}{\delta r} r_s^4. \qquad (12)$$

Using (8), we can obtain the viscous torque applied to the solid core:

$$\mathbf{L}_s^{(\eta)} = W \Omega_0^2 (\mathbf{m}_f - \mathbf{m}_s), \qquad (13)$$
$$W = \eta(r_s) \frac{8\pi}{3} \frac{r_s^4}{\delta r \Omega_0};$$

i.e., the viscous torque is of order $O(m)$.

With the same notation used in [7], assuming that the variations of the equatorial components of the angular velocities of the envelopes $\widetilde{m}, \widetilde{m}_f, \widetilde{m}_s$ are determined by harmonics of the tidal potential with frequency σ (we express frequency in units of the average angular velocity of the Earth's rotation Ω_0; i.e., the frequency σ in usual units corresponds to the frequency $\sigma\Omega_0$), we obtain the torque equations taking into account the effects of viscosity and the solid core (i.e. for Model B for the Earth):

$$Mx = y, \tag{14}$$

where

$$M = \begin{pmatrix} \sigma(1+\kappa)+\kappa-e & (1+\sigma)\left(\frac{A_f}{A}+\xi\right) & (1+\sigma)\left(\frac{A_s}{A}+\zeta\right) & (1+\sigma)\alpha_3 e_s\frac{A_s}{A} \\ \sigma(1+\gamma) & 1+e_f+\sigma(1+\beta)-i\frac{W}{A_f} & \sigma\delta+i\frac{W}{A_f} & -\sigma\alpha_1 e_s\frac{A_s}{A} \\ \sigma(1+\theta)-e_s\alpha_3 & \sigma\chi+\alpha_1 e_s+i\frac{W}{A_s} & \sigma\nu+1+\sigma-i\frac{W}{A_s} & (1+\sigma-\alpha_2)e_s \\ 0 & 0 & 1 & \sigma \end{pmatrix},$$

$$x = \begin{pmatrix} \widetilde{m} \\ \widetilde{m}_f \\ \widetilde{m}_s \\ \widetilde{n}_s \end{pmatrix}, \qquad y = \begin{pmatrix} (\kappa - e + \sigma\kappa)\widetilde{\phi} \\ \sigma\gamma\widetilde{\phi} \\ (\sigma\theta - \alpha_3 e_s)\widetilde{\phi} \\ 0 \end{pmatrix}.$$

Here, A, A_f, A_s and e, e_f, e_s are the equatorial moments of inertia and dynamical compression coefficients for the whole Earth, the fluid core, and the solid core, respectively, while σ is the frequency of the harmonic of the tidal potential $\widetilde{\phi}$. The remaining Greek symbols denote the elastic constants and coefficients of the gravitational interactions for the Earth's envelopes.

We find the frequencies of the normal modes by setting the determinant of (14) equal to zero. Two of these frequencies are the Chandler frequencies for the Earth (CW) and for its fluid core (ICW). They are approximately

$$\sigma_{CW} = \frac{A}{A_m}(e-\kappa), \quad \sigma_{ICW} = (1-\alpha_2)e_s.$$

The viscosity does not change these frequencies but affects two almost diurnal frequencies, the frequencies of the reverse (RFCN) and forward (PFCN) free nutation of the core, which are approximately

$$\sigma_{RFCN} = -1-\left(1+\frac{A_f}{A_m}\right)(e_f-\beta)+i\left(1+\frac{A_f}{A_m}\right)\frac{W}{A_f},$$

$$\begin{aligned}\sigma_{PFCN} = -1 &+ \left(1+\frac{A_f}{A_m}\right)(e_f-\beta)(\alpha_2 e_s+\nu) \\ &+ \frac{W^2}{A_f A_s} + i\Big[\left(1+\frac{A_f}{A_m}\right)(e_f-\beta)\frac{W}{A_s} \\ &- \left(1+\frac{A_f}{A_m}\right)\frac{W}{A_f}(\alpha_2 e_s+\nu)\Big].\end{aligned}$$

We selected an initial W that depended on the viscosity η in accordance with the Earth's Q factor (see Appendix) for the frequency σ_{RFCN}. We then varied it to obtain the best agreement with the observations.

We have used the same dependence of the viscosity on the radius (Fig. 1), though we decreased somewhat the viscosity at the boundary between the solid and fluid cores to provide the best agreement with the observations. In addition, we expect the viscosity to be positive (the variations could in principle lead to an impossible negative viscosity).

3.2. Effects of the Atmosphere

To estimate the effects of the atmosphere, we must add a fourth equation to the torque equations (2)–(4) (Model **A**):

$$\frac{\partial \mathbf{H}_a}{\partial t} + \mathbf{\Omega} \times \mathbf{H}_a = \mathbf{L}_a. \tag{15}$$

In (15), $\mathbf{H}_a$ is the angular momentum of the rotation of the atmosphere and $\mathbf{L}_a$ is the torque applied to the atmosphere.

The components of $\widetilde{L}_a$ were obtained in [21] and can be expressed in the form $\widetilde{L}_a = -iU(\widetilde{c}_3^a + \widetilde{h}/\Omega_0)$, where the parameter U depends on the shape of the Earth's figure. The variations of the atmosphere's moment of inertia $\widetilde{c}_3^a$ are related to the exciting atmospheric functions χ [22] via the expression $\chi = (\widetilde{c}_3^a + \widetilde{h}/\Omega_0)/(C-A)$.

The system of torque equations taking into account the atmosphere, viscosity, and solid core (for Model **A**) becomes six-dimensional, and takes the form

$$Mx = y, \tag{16}$$

where

$$M =$$

$$\begin{pmatrix}
= \sigma(1+\kappa)+\kappa-e & (1+\sigma)\left(\frac{A_f}{A}+\xi\right) & (1+\sigma)\left(\frac{A_s}{A}+\zeta\right) & (1+\sigma)\alpha_3 e_s \frac{A_s}{A} & (1+\sigma)\frac{A_a}{A} & (1+\sigma)\frac{A_a}{A}e_a(1+\kappa') \\
\sigma(1+\gamma) & 1+S_{22}+\sigma(1+\beta) & \sigma\delta+S_{23} & -\sigma\alpha_1 e_s\frac{A_s}{A} & 0 & \sigma(\frac{\xi}{\tau}+h_f)\frac{A_a}{A_f}e_a \\
\sigma(1+\theta)-e_s\alpha_3 & \sigma\chi+\alpha_1 e_s+S_{32} & \sigma\nu+1+\sigma S_{33} & (1+\sigma-\alpha_2)e_s & 0 & \sigma(\frac{\zeta}{\tau}+h_s)\frac{A_a}{A_s}e_a \\
0 & 0 & 1 & \sigma & 0 & 0 \\
\sigma-e_a & 0 & 0 & 0 & 1+\sigma & \left(1+\sigma+\frac{U}{\Omega_0^2}\right)e_a \\
0 & 0 & 0 & 0 & 1 & \sigma,
\end{pmatrix}$$

$$x = \begin{pmatrix} \widetilde{m} \\ \widetilde{m}_f \\ \widetilde{m}_s \\ \widetilde{n}_s \\ \widetilde{m}_a \\ \widetilde{n}_a \end{pmatrix}, \quad y = \begin{pmatrix} (\kappa-e+\sigma\kappa)\widetilde{\phi}-(1+\sigma)(1+\kappa')\frac{\widetilde{c}_3^a}{A} \\ \sigma\gamma\widetilde{\phi}-\sigma(\frac{\xi}{\tau}+h_f)\frac{\widetilde{c}_3^a}{A_f} \\ (\sigma\theta-\alpha_3 e_s)\widetilde{\phi}-\sigma(\zeta/\tau+h_s)\frac{\widetilde{c}_3^a}{A_s} \\ 0 \\ -\left(1+\sigma+\frac{U}{\Omega_0^2}\right)\frac{\widetilde{c}_3^a}{A_a} \\ 0 \end{pmatrix}.$$

The components of the viscous tensor S_{ab} are $S_{22} = e_f - i\frac{W}{A_f}$, $S_{23} = i\frac{W}{A_f}$, $S_{32} = i\frac{W}{A_s}$, and $S_{33} = -i\frac{W}{A_s}$. The remaining parameters have the same meaning as previously. When solving system (16), we assume that variations in the atmosphere's moment of inertia $\widetilde{c}_3^a$ are due to atmospheric tides [23].

The formulas for deriving approximate values for the already known normal modes have the same form as those in Model **B**. However, this does not mean that there are no atmospheric effects. The effect of the atmosphere is indirect, and taking it into account changes the conditions for determining the unknown internal parameters of the Earth's structure in formulas for the normal modes.

Model **A** proposes two additional normal modes that were absent from Model **B**: the atmospheric Chandler frequency σ_{ACW} and the almost diurnal reverse frequency of the free nutation of the atmosphere σ_{RFAN}. Approximate formulas for these were obtained in [21]:

$$\sigma_{ACW} = \left(1+\frac{U}{\Omega_0^2}\right)e_a,$$

$$\sigma_{RFAN} = -\left(1+\frac{U}{\Omega_0^2}e_a+(\kappa+\beta-2\xi)\frac{A}{A_m}\right),$$

where the parameter U is determined by the shape of the Earth's surface and e_a is the dynamical compression of the atmosphere.

3.3. Effects of the Magnetic Field

The magnetic field was neglected in Models **A** and **B**. It is not necessary to compute the magnetic force since this was done in [24]. In accordance with [24], the magnetic forces are taken into account by adding the terms $K^{CMB}+K^{ICB}A_s/A_f$ to S_{22}, $-K^{ICB}A_s/A_f$ to S_{23}, $-K^{ICB}$ to S_{32}, and $K^{ICB}A_s/A_f$ to S_{33}, where K^{CMB} and K^{ICB} depend on the radial component of the magnetic field, conductivity, depth to which the magnetic field penetrates, and the radii of the boundaries between the core and mantle and between the solid and fluid cores. Then, the viscous matrix S_{ab} becomes magnetoviscous and has the following components:

$$S_{22} = \left(e_f + \mathrm{Re}\left(K^{CMB}+\frac{A_s}{A_f}K^{ICB}\right),\right. \qquad (17)$$

$$\left.\mathrm{Im}(K^{CMB})+\frac{A_s\mathrm{Im}(K^{ICB})-W}{A_f}\right),$$

$$S_{23} = \left(-\frac{A_s}{A_f}\mathrm{Re}(K^{ICB}), \frac{W-A_s\mathrm{Im}(K^{ICB})}{A_f}\right),$$

$$S_{32} = \left(-\mathrm{Re}(K^{ICB}), \frac{W}{A_s}-\mathrm{Im}(K^{ICB})\right),$$

$$S_{33} = \left(\mathrm{Re}(K^{ICB}), -\frac{W}{A_s}+\mathrm{Im}(K^{ICB})\right).$$

The remaining equations and the matrix M (see (16)) retain their previous forms. Note that this is the most general formulation of the problem:

for zero magnetic numbers K^{CMB} and K^{ICB}, it results in six-dimensional viscous models similar to Model **A**; neglecting the atmosphere gives rise to four-dimensional viscous models similar to Model **B**; without the atmosphere and viscosity, we obtain a four-dimensional magnetic nonviscous model of the type in [24]; without the atmosphere, viscosity, and magnetic forces, we obtain a four-dimensional nonviscous and nonmagnetic model of the type in [7]; neglecting the solid core, magnetic forces, and viscosity, we obtain a four-dimensional nonviscous and nonmagnetic model similar to Model **C** [14].

In general, the magneto-viscosity matrix S_{ab} enables us to determine only the following combinations of quantities: $e_f + \mathrm{Re}\,(K^{CMB}), \mathrm{Im}(K^{CMB})$, $\mathrm{Re}(K^{ICB})$, and $A_s\mathrm{Im}(K^{ICB}) - W$. However, for the Earth's magnetic field, the real part of K^{CMB} is equal to the imaginary part of K^{CMB} [24] and we can also determine e_f and $\mathrm{Re}(K^{CMB}) = \mathrm{Im}(K^{CMB})$ separately.

4. CALCULATION RESULTS

Using a system of equations of type (14) or (16), we can find the unknown quantities $\widetilde{m}$, $\widetilde{m}_f$, etc. as functions of the exciting potentials. The complex function $\widetilde{m}$ describes the position of the Earth's instantaneous rotational axis with respect to its figure axis, which is very close to the $\mathbf{i}_3$ axis. This function is none other than the polar motion. In the frequency domain, the polar motion $\widetilde{m}$ is related to the nutation $\widetilde{\zeta}$ (see Appendix) via the simple kinematic formula [15]:

$$\widetilde{\zeta}(\sigma) = -\frac{\widetilde{m}(\sigma)}{1+\sigma}, \qquad (18)$$

with the ratio $\widetilde{\zeta}(\sigma)/\widetilde{m}(\sigma)$ being independent of the Earth's internal structure.

To determine the nutation $\widetilde{\zeta}(\sigma)$ of the elastic Earth, we first find the nutation $\widetilde{\zeta}_R(\sigma)$ of the rigid Earth, then multiply it by the transformation function $\widetilde{q}(\sigma)$ [17], since

$$\widetilde{q}(\sigma) = \frac{\widetilde{\zeta}(\sigma)}{\widetilde{\zeta}_R(\sigma)} = \frac{\widetilde{m}(\sigma)}{\widetilde{m}_R(\sigma)}, \qquad (19)$$

where $\widetilde{m}_R$ is the polar motion of the rigid Earth. We can find the function $\widetilde{m}_R$ from (14) or (16), assuming that all elastic parameters and moments of inertia of the core and atmosphere are equal to zero. This yields

$$(\sigma - e)\widetilde{m}_R = -e\widetilde{\phi},$$

and we can find $\widetilde{\zeta}_R$ using (18) and (19).

The nutation of the rigid Earth has been derived using various methods (see [5] for details) and with varying accuracies. The accuracy of the theory of the elastic Earth will depend on the accuracy of this nutation theory, since the errors will also be multiplied by the function $\widetilde{q}(\sigma)$. However, the accuracy of nutation theory for the rigid Earth currently exceeds the accuracy of VLBI observations. The amplitudes obtained for the nutational harmonics of the elastic Earth are corrected to take into account the effects of the oceans and mantle dissipation (see Section 2).

The transformation function used to calculate the amplitudes was represented in the resonance form

$$q(\sigma) = R + R'(1+\sigma) + \sum_{i=1}^{N} \frac{R_i}{\sigma - \sigma_i},$$

where R, R', and R_i are the resonance coefficients, σ_i are the frequencies of the normal modes, and N is the dimension of the system of equations (14) or (16).

For the observed amplitudes of the nutational harmonics, we fit various parameters of the Earth's structure in order to reach the best agreement between the model and observations. In Models **A**, **B**, and **C**, we fit the dynamical compression of the core–mantle boundary e_f, with the criterion for goodness of fit being the agreement between the reverse annual nutation amplitude and the observed amplitudes, for which we used the nutation amplitudes presented in [24].

Table 1 compares the theoretical amplitudes of these three models with the observations. Those for which the deviations exceed the observational errors are marked in bold. If we compare the data for Models **B** and **C**, we can see that their deviations from the observations are approximately equal in magnitude but opposite in sign. Therefore, we supposed that Model **A** (atmosphere + mantle + fluid core + solid core + viscous forces) should be in almost complete agreement with observations.

However, contrary to our expectations, the discrepancy actually becomes somewhat larger for Model **A** (Table 1). This is due to the fact that the dynamical compression of the core–mantle boundary was varied in these theories in order to obtain agreement between the amplitude of the reverse annual nutation and the observations. The compression e_f in Models **B** and **C** is different; thus, the transformation functions also differ. Due to this difference in the resonance amplification, the unification of these models does not provide a linear addition of errors.

The unification of Models **B** and **C** into Model **A** has not provided improved agreement with the observations. However, these computations have verified the following important conclusion: it is not possible to take into account atmospheric effects in semi-analytical models via the addition of constant corrections. This is associated with the fact that such

Table 1. Deviations of theoretical amplitudes of the nutation harmonics for Models A, B, and C from observations

Period, days	Observed amplitudes [24]	Model A	Model B	Model C
−13.66	(−3.63, −0.03)	(−0.01, −0.02)	(−0.01, −0.02)	(0.01, 0.02)
−182.62	(−24.57, −0.06)	(**0.04**, 0.00)	(0.00, 0.00)	(0.01, 0.01)
−365.26	(−33.05, 0.35)	(0.00, **0.42**)	(0.00, **0.42**)	(0.00, −**0.42**)
−6798.38	(−8024.88, 1.43)	(−0.31, 0.57)	(0.01, 0.06)	(−0.47, −0.57)
6798.38	(−1180.50, −0.07)	(0.15, 0.04)	(0.10, 0.04)	(−0.06, 0.18)
365.26	(25.64, 0.12)	(**0.03**, **0.10**)	(**0.08**, **0.10**)	(−**0.04**, −**0.12**)
182.62	(−548.47, −0.51)	(**0.31**, 0.00)	(−**0.21**, 0.00)	(−**0.08**, 0.00)
13.66	(−94.20, 0.13)	(**0.14**, **0.13**)	(0.02, **0.13**)	(−**0.04**, −**0.13**)

Table 2. Deviations of theoretical amplitudes of the nutation harmonics for Models D and E from observations

Period, days	Observed amplitudes [24]	Model D	Model E
−13.66	(−3.64, −0.03)	(−0.00, −0.02)	(−0.00, −0.02)
−182.62	(−24.57, −0.06)	(**0.04**, 0.02)	(**0.04**, −0.01)
−365.26	(−33.06, 0.35)	(−0.00, 0.00)	(−0.00, −**0.42**)
−6798.38	(−8024.88, 1.43)	(−0.32, 0.34)	(−0.02, 0.64)
6798.38	(−1180.46, −0.07)	(0.13, 0.06)	(0.08, 0.02)
365.26	(25.65, 0.12)	(−0.01, −0.01)	(−0.001, **0.08**)
182.62	(−548.48, −0.51)	(−0.01, 0.02)	(**0.14**, −**0.06**)
13.66	(−94.18, 0.13)	(0.00, −0.02)	(−0.01, −0.02)

Table 3. Resonance characteristics for Model D for the Earth

Modes	Period, days	Resonance coefficients R_i
CW	394.52	$(-5.817\,951\,9 \times 10^{-4}, -2.774\,4 \times 10^{-9})$
RFCN	426.71	$(-1.201\,517 \times 10^{-4}, -8.693\,3 \times 10^{-7})$
ICW	2381.98	$(-4.078\,41 \times 10^{-8}, 1.676\,1 \times 10^{-11})$
PFCN	367.52	$(1.216\,028\,9 \times 10^{-6}, -4.871\,844 \times 10^{-7})$
ACW	23.64	$(4.6 \times 10^{-15}, 1.8 \times 10^{-16})$
RFAN	36.37	$(4.702 \times 10^{-13}, 1.1 \times 10^{-15})$
		$R = (1.051\,273\,517, -6.444\,156\,4 \times 10^{-7})$
		$R' = (-0.281\,414\,161\,4, 4.08 \times 10^{-14})$

corrections change when the parameters are varied (in particular, the compression e_f). Therefore, we must either appropriately vary the corrections or take the atmosphere into account in the torque equations. We believe that the latter approach is more correct from both the methodological and physical point of view.

In Models **E** and **D**, we varied the independent components of the tensor S_{ab} and other parameters, with the criterion for goodness of fit being the minimization of the weighted sum of the squared deviations of the theoretical from the observed amplitudes (we eliminated the 18-yr and 6-yr harmonics, since the accuracy of their detection is the lowest). We did not fit the geophysical dynamical compression of the whole Earth and took it to be $e = 3.284\,520\,155\,008 \times 10^{-3}$ [5]. Thus, we estimate the accuracy of the theory from the deviations of the amplitudes of the main

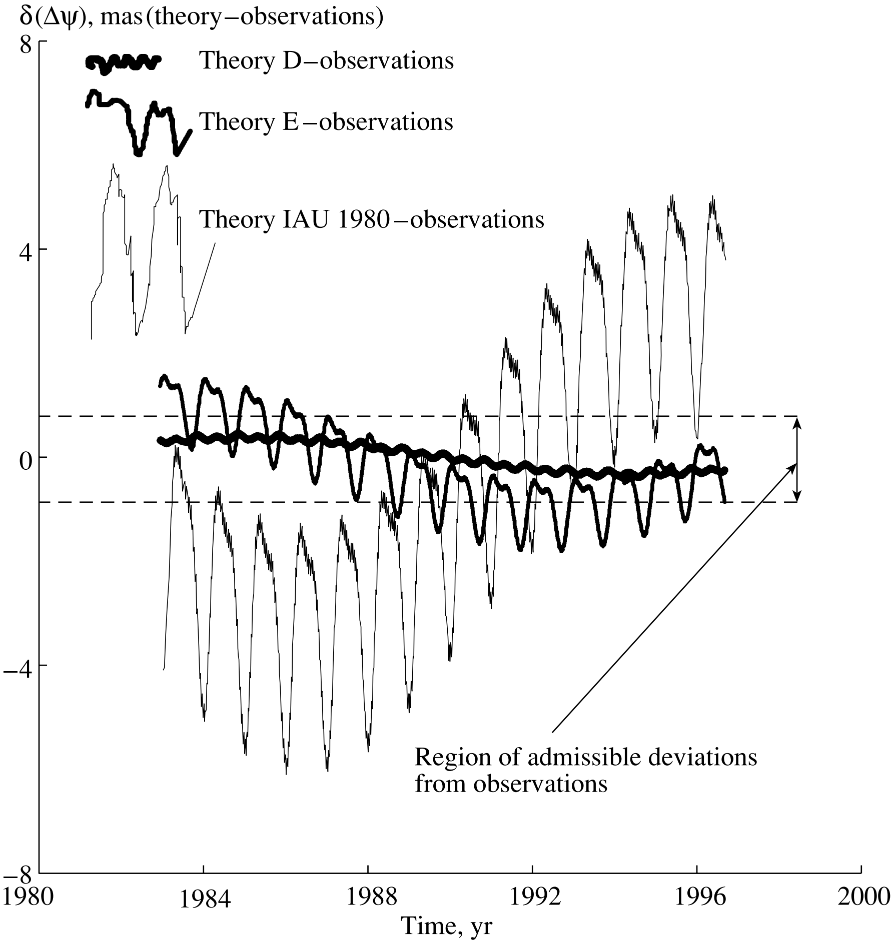

Fig. 2. Deviations of the theoretical nutation (in inclination) from the observations. Model **D** provides full agreement with observations (within the observational errors). Model **E** displays worse agreement, though it is more physical.

nutational harmonics (their real and imaginary parts) from the observed values.

We calculated the nutation amplitudes for all combinations of the nutation series for a completely rigid Earth (using three series [25, 26, 27]) and corrections taking into account the effects of inelastic dissipation in the mantle (five models) and of the oceans (six models) [25, 29]. Upon comparison of our calculations with the observations, we decided to use the rigid-body nutation series from [26], corrections for the inelastic mantle dissipation for a β model with $\alpha = 0.15$ from [25], and corrections for the effects of oceans and oceanic flows from [28]. Note that the differences between various contemporary rigid-body series are almost negligible and any of them could, in principle, be adopted; we prefer the series from [26].

We selected corrections taking into account the effects of the oceans and inelastic mantle dissipation based on achieving the best agreement with both the nutation observations and other geophysical observations. Corrections for the oceans [28] were calculated by solving the Laplace tidal equations for the world ocean level determined from satellite observations. Corrections for mantle dissipation were calculated by solving the continuous-medium equations for the mantle, with the parameter $\alpha = 0.15$ being verified by satellite observations [29]. These corrections also turn out to be the most consistent with nutation observations, confirming that our model describes well the effects of the Earth's mantle, oceans, and atmosphere.

Our fitting indicates the following Earth parameters:

$$e_f = 2.673\,074\,529\,217\,239 \times 10^{-3}, \mathrm{Re}(K^{CMB})$$
$$= \mathrm{Im}(K^{CMB}) = 2.670\,849\,687\,083\,623 \times 10^{-3},$$
$$\mathrm{Re}(K^{ICB}) = 5.373\,476\,463\,538\,518 \times 10^{-2},$$
$$-\frac{W}{A_s} + \mathrm{Im}(K^{ICB}) = 3.443\,181v818v0 \times 10^{-4},$$
$$k = 1.042\,013\,671\,875 \times 10^{-3},$$
$$\gamma = 1.965\,103\,515\,625\,008 \times 10^{-3}.$$

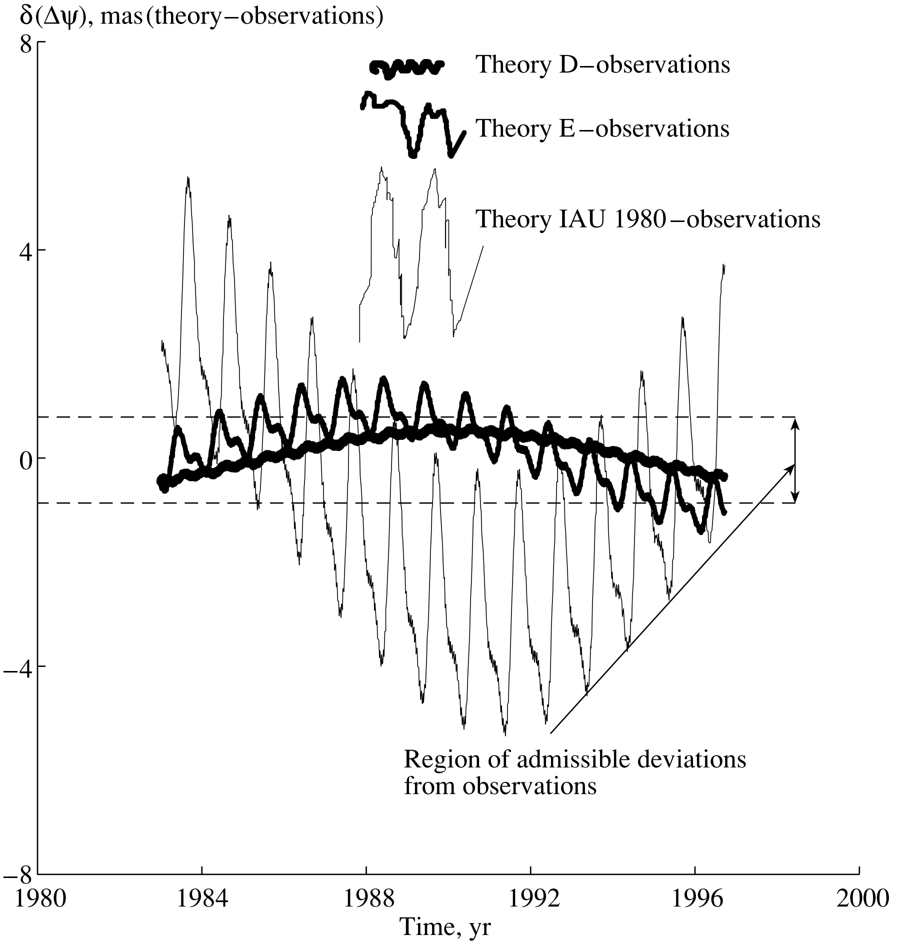

Fig. 3. Deviations of the theoretical nutation (in longitude) from observations.

The deviations of the nutation amplitudes for Model **D** and from the observations are presented in Table 2.

The agreement with observations is virtually complete for Model **D** (the nutation deviations in longitude and inclination are smaller than 1 mas). However, the imaginary part of every eigenfrequency of the matrix M for Model **D** is negative:

$$\sigma_{CW} = (2.528\,05 \times 10^{-3},\ -5.15 \times 10^{-9}),$$
$$\sigma_{ACW} = (0.042\,176\,5,\ 0.00),$$
$$\sigma_{ICW} = (4.142\,579 \times 10^{-4},\ -1.429\,6 \times 10^{-7}),$$
$$\sigma_{RFAN} = (-1.027\,417\,7,\ 0.00),$$
$$\sigma_{RFCN} = (-1.002\,337\,1,\ -1.868\,659 \times 10^{-5}),$$
$$\sigma_{PFCN} = (-0.997\,270\,0,\ -3.457\,123 \times 10^{-4}).$$

The imaginary parts of σ_{ACW} and σ_{RFAN} are smaller than 10^{-14} and negative. We assume that they are approximately equal to zero.

We should emphasize that the eigenfrequencies of M are those calculated for the Earth without oceans and mantle dissipation. Consequently, the originally elastic system becomes self-exciting when the magnetic and viscous forces are added, whereas it should display the opposite behavior (see Appendix).

We believe that Model **D** does not take into account all relevant forces in the Earth's core, such as that due to the significant inclination ($\sim 11°$) [31] of the rotational axis of the solid core with respect to that of the mantle.

Table 3 presents the periods of the normal modes (in days), dynamical compression of the core–mantle boundary, and resonance coefficients for Model **D** ($e_f = 2.673\,074\,529 \times 10^{-3}$).

With the additional requirement that the imaginary parts of the eigenfrequencies be positive, Model **E**, which includes viscosity but neglects the magnetic field, proves to be preferable. Table 2 presents the deviations of the nutation amplitudes from the observed values. This nutation model is much better than Models **B** and **C**, but its accuracy is not sufficiently high (the deviations in longitude and inclination are about 2 mas, while current requirements are for accuracies better than 1 mas). Figures 2 and 3 compare the theoretical nutations for the main harmonics for Models **D** and **E** in longitude and inclination with the IAU1980 theory and observations.

5. DISCUSSION

One can use nutation theory for two main purposes: to predict the nutation in longitude and inclination for an assumed model of the Earth's structure and to determine the parameters of the Earth's internal structure from observations of nutation angles. Our theory has been successful in connection with the first problem; i.e., it predicts the nutation angles with the required accuracy (the rms deviation of the theoretical from the observed nutation is 0.35 mas). We have constructed a nutation theory, since we have completed the nutation series using the adopted model for the Earth. However, our theory is less successful in solving the second problem, since we have not taken into account all features of the structure of the Earth's core in Model **D** (for example, the possible inclination of the inner core).

Our calculations show that six-dimensional nutation models are more accurate than four-dimensional models. This seems obvious, since the number of parameters characterizing the Earth's structure is increased. However, we note that the transition from the four-dimensional to the six-dimensional theory does not expand any possibilities for adjusting the parameters in Eqs. (14) or (16), simply because none of the atmospheric parameters were adjusted. It might seem that adjusting the parameters of the magneto-viscosity matrix and the two elastic constants γ and k would provide any result desired, but this is not the case. The reason for this is that the calculations are very insensitive to variations in k and γ. In addition, there are only two independent components in the magneto-viscosity matrix, for example, S_{22} and S_{33}, and only S_{22} varies appreciably. This means that we can vary only one complex parameter, S_{22}, determined by the structure of the Earth's core, namely by the compression of the core–mantle boundary, the magnetic field, and viscosity. Therefore, our fitting can determine the properties of only part of the Earth: the fluid core + solid core + their interactions. If there were another dependence of the nutation amplitudes on the parameters γ, k, and S_{33}, we would be able to obtain complete agreement with the observations for any combination of corrections taking account of the effects of the mantle and oceans, especially for the corrections selected by us. However, in practice, this did not occur.

The next very important question is the practical application of the nutation series for Model **D**. To give a clear and well-based answer, let us transform Eqs. (16) by expressing $\widetilde{m}_f, \widetilde{m}_s$, and $\widetilde{n}_s$ in terms of $\widetilde{m}, \widetilde{m}_a$, and $\widetilde{n}_a$ using the second, third, and fourth equations. System (16) then takes the form

$$Sg = h_1 + h_2,$$

where S is a 3×3 matrix, $g = (\widetilde{m}, \widetilde{m}_a, \widetilde{n}_a)^x$ is a column vector, and the function h_1 describes the effects of the external tidal potential on the part of the Earth made up of the mantle, atmosphere, and oceans, with their interactions (i.e., h_1) being independent of the core structure and interactions inside the core. On the contrary, the function h_2 depends on the core structure and interactions in the core. To the first approximation, h_2 takes into account only the elastic fluid core.

The fact that the corrections providing the best agreement with the nutation observations are also in agreement with other observations testifies to the fact that our theory provides a good description the part of the Earth made up of the mantle, atmosphere, and oceans, as well as of the frequency dependence of h_2. However, the negative imaginary parts of the eigenfrequencies indicate that our theory inaccurately relates the coefficients of h_2 to the Earth's structural parameters. Thus, our theory for Model **D** can be used to predict the nutation angles (the first problem of nutation theory) but not to separately estimate the viscosity, magnetic field, and other parameters of the Earth's core structure based on nutation observations (the second problem of nutation theory). Additional work is necessary to solve the second problem, which we are in the process of carrying out [1].

As a first step, we can propose to improve the initial torque equations (3) and (4) in order to take into account the large inclination of the solid core and the triaxial structure of the core–mantle boundary. Improvement of the corrections for the oceans could be the next step. Corrections for the oceans in the form of the series [28] are probably sufficient to achieve an accuracy ~ 1 mas. To increase the accuracy of the theory, we must find another dependence of the corrections on frequency. This problem has been partially solved in [24], where the corrections for the oceans are represented as approximations in the form of a quadratic trinomial that must be added to the transformation function for short periods (shorter than 50 days). This expression does not describe the effects of the oceans over the whole frequency range and is unacceptable even as an approximation. Nevertheless, the use of frequency-dependent corrections improves the accuracy of the nutation theory.

---

[1] Similar difficulties appear in other nutation theories, for example, in the MHB2000 theory (in contrast to our theory, this model neglects the viscous forces and atmosphere). This is quite natural, since our analytical theory is a generalization of all previous models.

In our opinion, in order to correctly take into account the oceans, it is necessary to solve the Laplace equations for the real oceans and seas with boundary conditions specified at coastlines. This will yield a set of ocean eigenmodes that will lead to changes in the Earth's transformation function and nonlinear variations in the corrections.

In the MHB2000 nutation theory recommended by the IAU, the imaginary parts of the frequencies in the resonance formula for the transformation function are quite physically reasonable. This was attained by replacing the formula *nutation amplitude = transformation function × nutation amplitude of a completely rigid Earth + corrections for dissipation in the oceans and mantle*, with the formula *nutation amplitude = new transformation function × nutation amplitude of a completely rigid Earth*. In this transition, the strong dissipation in the mantle and oceans hides the nonphysical components of the magnetic tensor (viscosity and atmosphere were neglected).

6. CONCLUSIONS

We have constructed a nutation theory for various models of the Earth's structure. This theory is presented as a set of nutational harmonics and amplitudes determined by the model used for the Earth. The theory includes 1498 nutation harmonics whose frequencies are determined by lunar—solar and planetary tidal potentials. It satisfies all the IAU requirements, providing an rms deviation of about 0.35 mas. Our theory is realized using software with a convenient user interface, which can be found at http: //lnfm1.sai.msu.ru/~kazaryan/TEST/homepage/russian/program/prog.htm. In addition, this software can be used to determine the nutation angles calculated by empirical series obtained from observations [4].

Currently, we are searching for additional factors that should be taken into account and considering various models for the internal structure of the Earth's core that would enable one not only to construct an approximation for h_2, but also to derive a more accurate relation between the h_2 coefficients and the Earth's core parameters consistent with available observational data (both geophysical and astrometrical).

7. ACKNOWLEDGMENTS

The authors are grateful to V. Dehant for presenting us with the complete rigid-body series of the Earth's nutation RDAN98. We thank the referee for helpful remarks. This work was supported by the Russian Foundation for Basic Research (project no. 98-05-64797) and the "Universities of Russia" (project 2-5547) and "Integration" (K0641) programs.

REFERENCES

1. J. M. Wahr, Geophys. J. R. Astron. Soc. **64**, 651 (1981).
2. J. M. Wahr, Geophys. J. R. Astron. Soc. **64**, 705 (1981).
3. P. K. Seidelmann, Celest. Mech. **27**, 79 (1982).
4. *1998 IERS Annual Report* (Observatoire de Paris, Paris, 1999).
5. V. Dehant, F. Arias, C. Bizouard, *et al.*, Celest. Mech. Dyn. Astron. **72**, 245 (1999).
6. P. M. Mathews, B. A. Buffet, T. A. Herring, and I. I. Shapiro, J. Geophys. Res. **96**, 8219 (1991).
7. T. Shirai and T. Fukushima, Astron. J. **119**, 2475 (2000).
8. *IERS Conventions,* IERS Tech. Note 21 (Observatoire de Paris, Paris, 1996).
9. J. Getino and J. M. Ferrándiz, Mon. Not. R. Astron. Soc. **306**, L45 (1999).
10. P. M. Mathews, T. A. Herring, and B. A. Buffet, J. Geophys. Res. (2001) (in press).
11. P. M. Mathews, B. A. Buffet, T. A. Herring, and I. I. Shapiro, J. Geophys. Res. **96**, 8243 (1991).
12. A. M. Dziewonski and D. L. Anderson, Phys. Earth Planet. Inter. **25**, 297 (1981).
13. T. Sasao and J. M. War, Geophys. J. R. Astron. Soc. **64**, 729 (1981).
14. V. E. Zharov, Doctoral Dissertation in Mathematical Physics [in Russian] (St. Petersburg, 1998).
15. T. Sasao, S. Okubo, and M. Saito, in *Proceedings of IAU Symposium No. 78, 1980,* p. 165.
16. H. Moritz and I. I. Mueller, *Earth Rotation: Theory and Observation* (Ungar, New York, 1987; Naukova Dumka, Kiev, 1992).
17. V. N. Zharkov, *Internal Structure of the Earth and Planets* [in Russian] (Nauka, Moscow, 1983).
18. D. L. Anderson, *Theory of the Earth* (Blackwell Sci. Publ., Boston, 1989).
19. V. V. Brazhkin, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 469 (1999) [JETP Lett. **68**, 502 (1998)].
20. V. V. Brazhkin, Usp. Fiz. Nauk **170**, 535 (2000).
21. V. E. Zharov, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron., No. 6, 65 (1997).
22. R. T. H. Barnes, R. Hide, A. A. White, and C. A. Wilson, Proc. R. Soc. London, Ser. A **387**, 31 (1983).
23. Ch. Bizouard, A. Brzezinski, and S. Petrov, J. Geodesy **72**, 561 (1998).
24. P. M. Mathews, B. A. Buffet, T. A. Herring, and M. Fessel, in *Journess: Systemes de reference spatio-temporales* (Observatoire de Paris, Paris, 1988), p. 86.
25. T. A. Herring, P. M. Mathews, B. A. Buffet, and I. I. Shapiro, J. Geophys. Res. **96**, 8259 (1991).
26. F. Roosbeek and V. Dehant, Celest. Mech. Dyn. Astron. **70**, 215 (1998).
27. J. Souchay and H. Kinoshita, Astron. Astrophys. **312**, 1017 (1996).

28. B. F. Chao, R. D. Ray, J. M. Gipson, *et al.*, J. Geophys. Res. **101** (B9), 20151 (1996).
29. V. Dehant and P. Defraigne, J. Geophys. Res. **102**, 27659 (1997).
30. J. Wahr and Z. Bergen, Geophys. J. R. Astron. Soc. **87**, 633 (1986).
31. X. Song and P. G. Richards, Nature **382**, 221 (1996).

APPENDIX

VARIOUS REPRESENTATIONS OF NUTATION

Mathematically, it is convenient to represent the nutation of the Earth's axis as a complex number $\widetilde{\zeta}(t) = \Delta\epsilon(t) + i\Delta\psi(t)\sin\epsilon_0$, where ϵ_0 is the inclination of the ecliptic toward the equator. The angles $\Delta\epsilon$ (nutation in inclination) and $\Delta\psi$ (nutation in longitude) are expressed via a series:

$$\Delta\epsilon(t) = \sum_j^N \left[\Delta\epsilon_{rj}\cos\Theta_j(t) + \Delta\epsilon_{ij}\sin\Theta_j(t)\right], \quad (20)$$

$$\Delta\psi(t) = \sum_j^N \left[\Delta\psi_{rj}\sin\Theta_j(t) + \Delta\psi_{ij}\cos\Theta_j(t)\right],$$

where $\Theta_j(t)$ is the tidal argument [4], $\Delta\psi_{rj}$ and $\Delta\epsilon_{rj}$ are the real (cophasal) coefficients of the jth term (of the N total) of the expansion, and $\Delta\psi_{ij}$ and $\Delta\epsilon_{ij}$ are the imaginary (quadrature) coefficients, with their phase being shifted by 90° with respect to $\Delta\psi_{rj}$ and $\Delta\epsilon_{rj}$. In the IAU1980 nutation theory, $\Delta\epsilon_{ij} = \Delta\psi_{ij} = 0$. For the real Earth, the quadrature terms $\Delta\epsilon_{ij}, \Delta\psi_{ij}$ differ from zero. This indicates dissipation inside the Earth at the given frequency.

The right-hand sides of (20) are frequently expressed as the sum of the forward and reverse nutation [9, 25]. The tidal argument is a linear combination of five basic arguments [9], with each being a function of time t and angular frequency $\sigma_i : \Theta_i = \sigma_i t + \phi_i$, where ϕ_i is the phase of the argument. The frequency of the nutation $\sigma_i = d\Theta_i/dt$ is positive for most harmonics. The conventional form [9] of the nutation $\widetilde{\zeta}(t)$ is

$$\widetilde{\zeta}(t) = -a(t) = -\sum_j^N \left(a_j^p e^{-iq_j\Theta_j(t)} + a_j^r e^{iq_j\Theta_j(t)}\right),$$

where a_j^p and a_j^r are the amplitudes of the forward and reverse nutation and q_j is the sign of the frequency σ_j; i.e., $q_j = \mathrm{sign}(\sigma_j)$. We represent the amplitudes a_j^p and a_j^r in the form

$$a_j^{p,r} = a_{rj}^{p,r} - ia_{ij}^{p,r},$$

where the subscripts r, i denote the real and imaginary parts of the forward (reverse) nutation, to obtain

$$\begin{aligned} a_{rj}^p &= -(\Delta\epsilon_{rj} - q_j\Delta\psi_{rj} sin\epsilon_0)/2, \qquad (21)\\ a_{rj}^r &= -(\Delta\epsilon_{rj} + q_j\Delta\psi_{rj} sin\epsilon_0)/2,\\ a_{ij}^p &= (q_j\Delta\epsilon_{ij} + \Delta\psi_{ij} sin\epsilon_0)/2,\\ a_{ij}^r &= -(q_j\Delta\epsilon_{ij} - \Delta\psi_{ij} sin\epsilon_0)/2. \end{aligned}$$

In the general case, the frequency is complex: $\widetilde{\sigma} = \sigma_1 + i\sigma_2$, with $\sigma_2 \ll \sigma_1$. The complex part is associated with energy dissipation at the tidal frequency σ_1 and, consequently, with the Earth's Q factor at this frequency:

$$\sigma = \sigma_1\left(1 + \frac{i}{2Q}\right).$$

When there is dissipation, the σ_2 must be positive. Indeed,

$$\cos\widetilde{\Theta} = \mathrm{Re}\, e^{i\widetilde{\Theta}} = \mathrm{Re}\, e^{i[(\sigma_1 + i\sigma_2)t + \phi]} = \cos\Theta e^{-\sigma_2 t},$$

and in the absence of excitement, the nutation amplitudes (20) will decrease with time only for positive σ_2.

Translated by V. Badin

Astronomy Reports, Vol. 45, No. 11, 2001, pp. 922–928. Translated from Astronomicheskiĭ Zhurnal, Vol. 78, No. 11, 2001, pp. 1049–1056.
Original Russian Text Copyright © 2001 by Kurbasova, Rykhlova.

Free Oscillations of Parameters of the Earth's Motion in the Earth–Moon System

G. S. Kurbasova[1] **and L. V. Rykhlova**[2]

[1] *Crimean Astrophysical Observatory, National Academy of Sciences of Ukraine, Nauchnyi, Crimea, 334413 Ukraine*
[2] *Institute of Astronomy, Russian Academy of Sciences, Pyatnitskaya ul. 48, Moscow, 109017 Russia*

Received August 1, 2000

Abstract—When the coupled Earth–Moon system is considered rather than a closed terrestrial system (the Earth's atmosphere, ocean, inner and outer cores, and mantle), an eigenfrequency of free oscillations of the Earth is transformed into an eigenfrequency of the Earth–Moon system, making it possible to understand the increase in the Chandler period in a natural way. The eigenmodes of the coupled Earth–Moon system are determined by solving a linear, homogeneous set of equations with two degrees of freedom. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The basic problem of astrometry is the construction, maintenance, and improvement of spatial and temporal coordinate systems, such that these systems remain as close as possible to an inertial coordinate system. Currently, such inertial systems are based on objects located beyond the Galaxy and possessing negligible proper motions. To solve applied problems within the solar system, we use a dynamical coordinate system based on precise positions of the Sun and planets, which are continuously moving with respect to the observer. Determining the relationship between an inertial and dynamical coordinate system requires an adequate theory for the translational and rotational motion of the Earth.

Such a theory should provide exact equations relating Earth-rotation parameters (the polar motion and nonuniformity of the rotation) to data on the shape, deformations, and mass distribution of the Earth's interior, taking into account all mutual forces between the Earth, planets, and Sun. This problem is extremely difficult, since neither numerical values for the parameters appearing in the equations nor the structure of these equations are known to a sufficient accuracy. As a result, there is some discrepancy between the observational data and predictions of existing theoretical models. For example, it is not possible to obtain an analytical description of the Earth's rotation parameters. This necessitates searches for new approaches to the analysis of polar motions, in particular of the excitation of free-nutational oscillation in the motion of the Earth's poles (the Chandler motion). We will consider a coupled Earth–Moon system rather than a closed system composed of the Earth's atmosphere, ocean, core, and mantle.

The Earth and Moon are separated by only 0.0026 AU. The center of gravity (barycenter) of the Earth–Moon system is 4672 km from the Earth's center, and the Earth's mass is 81.30 times greater than the Moon's. As a result, the Earth substantially perturbs the motion of the Moon in its heliocentric orbit.

On the other hand, the mass of the Moon is sufficiently large to produce appreciable perturbations in the heliocentric motion of the Earth. For comparison, the mass of the largest satellite of Neptune (Nereid) is a factor of 770 less than the mass of Neptune, the mass of the largest satellite of Saturn (Titan) is a factor of 4030 less than the mass of Saturn, and the mass of the largest satellite of Jupiter (Ganymede) is a factor of 1220 less than the mass of Jupiter. Moreover, the orbital momentum of the Moon is a factor of five greater than the angular momentum of the Earth. (In all the other systems mentioned above, the angular momentum of the planet turns out to be much greater than the total orbital momentum of all its satellites.)

The specific features of the coupling between the Earth and Moon enable us to treat the Earth–Moon system as a binary planet. At the same time, the entire lunar orbit lies beyond the region where the attraction of the Earth is greater than the attraction of the Sun: the Sun's attraction exceeds the Earth's by a factor of 2.18. The Earth and Moon represent a binary dynamical system of coupled bodies whose parameters are affected by external forces. The most substantial perturbation is produced by the Sun.

1063-7729/01/4511-0922$21.00 © 2001 MAIK "Nauka/Interperiodica"

The first studies to be undertaken on the motion of systems of coupled bodies began relatively recently. This is partially due to the difficulty of solving this problem. For example, solutions for the relative motion of two rigid bodies whose masses are attracted by Newtonian universal gravitation are known only in very special cases [1]. The need to analyze the general properties of the motions in a system of coupled bodies arose in connection with the dynamics of satellite systems, which began to be vigorously developed at the beginning of the 1960s. There is an extensive bibliography of published works in this area concerning the motions of rigid-body and gyrostatic satellites; in some cases, they take into account the elastic properties of the satellites or some their elements. The mathematical methods developed in these papers may be useful for studies of the specific features of motion of natural celestial bodies that cannot be explained in the framework of classical celestial mechanics.

One such effect is the increase in the period and variation in the amplitude of the Earth's eigenmodes (i.e., Chandler oscillations). The eigenmodes of individual bodies change when they are in a coupled system. As early as 1829, in his short paper "On the New General First Principle of Mechanics," Gauss concluded that, "It is very significant that, when free motions cannot occur under the present conditions, they are modified by Nature in precisely the same way as is done by a mathematician when he obtains a least squares fit to the results of observations corresponding to quantities related to each other by the required dependences."

We show here that the increase in the period of free nutation of the Earth (the Chandler period) can be explained very simply when the motion of the Earth–Moon–Sun coupled system is considered: an eigenfrequency of oscillations of the free Earth is transformed into an eigenfrequency of the Earth–Moon system as a result of the mutual influence of oscillations inside the system. We assume there to be a contactless coupling mechanism associated with the gravitational forces. This interpretation of the increase in the period of free nutation of the Earth was first considered in our previous work [2].

2. EQUATIONS OF THE COUPLED MOTION

Structure of Coupling in the Sun–Earth–Moon System

We will perform a preliminary analysis of the structure of the coupling in the three-body Sun–Earth–Moon system. These bodies interact with each other via gravitational forces and moments. The interaction forces for each pair of directly interacting bodies are governed by the law of universal gravitation.

The gravitational forces between the Earth and Moon vary as functions of the distance between the bodies. These can be considered internal forces in the Earth–Moon system and produce an appreciable external effect only because there is an external force field interacting with them.

We introduce the notation $e_1 = Gm_S$, $e_2 = Gm_E$, and $e_3 = Gm_M$. The quantities e_1, e_2, and e_3 characterize the intensity of the corresponding source, and G is the gravitational constant. Then, the Newtonian field whose sources are at the points $\mathbf{r}_i$ $(i = 1, 2, 3)$ will be characterized by the quantity

$$\mathbf{E} = \sum \frac{e_i}{|\mathbf{r} - \mathbf{r}_i|^3} (\mathbf{r} - \mathbf{r}_i). \tag{1}$$

The Sun plays a special role in the three-body system under consideration, since $e_1 \gg e_2$ and $e_1 \gg e_3$. Therefore, when deriving equations for the coupled motion of the Earth and Moon, we assume the motion of a point mass with the mass of the Sun (m_S) to be given.

If we take the Sun as a basis body, the positions of the Earth and Moon will be described by the radius vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of their centers of masses. The radius vector $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ describes the positions of the Earth (m_E) and Moon (m_M) at any time during their motion relative to the Sun (Fig. 1).

The motions of the Earth and Moon in their heliocentric orbits are affected by the forces of mutual attraction in the Earth–Moon system. The effect of these forces is similar to the imposition of constraints on the motion of the Earth and Moon around the Sun. In this case, the constraints represent an idealized model of the internal interaction forces. This model enables us to take into account the basic effect of the interaction without specifying the physical nature of the corresponding forces.

Lagrange Equations of the First Kind

We shall consider the special case when real rigid bodies are approximately described by three point masses m_S, m_E, and m_M. The interaction between the point masses m_E and m_M constrains their possible displacements, i.e., the domains of variability of their coordinates and velocities. This can be expressed by a constraint equation describing a surface of the form

$$f(t, \mathbf{r}_1, \mathbf{r}_2) = 0. \tag{2}$$

We will show that motions of the points $m_1 = m_E$ and $m_2 = m_M$ in the presence of the indicated constraints differ from free motion only under the action of the active force F_i $(i = 1, 2)$ from the Sun.

Let $\mathbf{r}_i$ and $\mathbf{v}_i$ be the radius vector and velocity of the ith point mass at time t. Then, in the absence of constraints, its position at time $t + dt$ will be described by the radius vector:

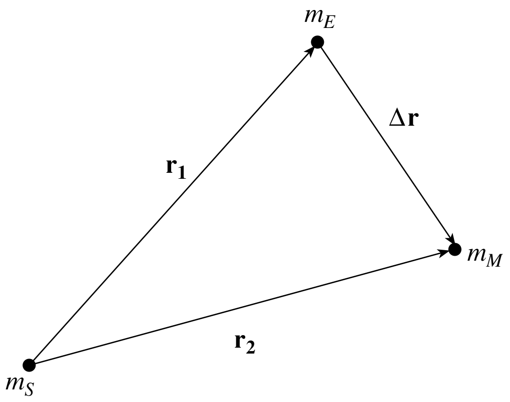

Fig. 1. Positions of the Earth (m_E) and Moon (m_M) in the course of their motion relative to the Sun (m_S).

$$\mathbf{r}_i(t+dt) = \mathbf{r}_i + \mathbf{v}_i dt + 1/2dt^2\frac{\mathbf{F}_i}{m} + \dots. \quad (3)$$

When the constraints are imposed for the same values of $\mathbf{r}_i$, $\mathbf{v}_i$, and t, the point mass m_i will have the position

$$\mathbf{r}'_i(t+dt) = \mathbf{r}_i + \mathbf{v}_i dt + 1/2dt^2\mathbf{w}_i + \dots \quad (4)$$

at time $t+dt$, where $\mathbf{w}_i$ is the acceleration of the ith point mass. Expressions (3) and (4) were derived using Taylor expansions.

The difference between (3) and (4) represents the deviation of the point mass from its free motion due to the constraints:

$$\mathbf{r}_i(t+dt) - \mathbf{r}'_i(t+dt) \quad (5)$$
$$= 1/2dt^2\left(\frac{\mathbf{F}_i}{m_i} - \mathbf{w}_i\right)\dots.$$

According to Gauss' principle of least constraint, the real motion of the system of point masses m_1 and m_2 is such that the constraint function

$$Z \equiv \sum_{i=1}^{2}\frac{1}{m_i}\left(\mathbf{F}_i - m_i\mathbf{w}\right)^2 \quad (6)$$

possesses a minimum; i.e., $\delta Z = 0$. The quantity Z can be considered the sum of the squares of the "lost forces" divided by the masses. By varying only the acceleration $\delta\overline{w}_i$, we obtain for the system of two point masses the expression

$$\sum_{i=1}^{2}\left(\mathbf{F}_i - m_i\mathbf{w}_i + \lambda\nabla f, \delta\mathbf{w}_i\right) = 0, \quad (7)$$

where ∇f is the gradient of the scalar constraint function and λ is the Lagrange factor (constraint coefficient).

Two vector Lagrange equations of the first kind follow directly from (7):

$$m_i\mathbf{w}_i = \mathbf{F}_i + \lambda\nabla f \quad \text{or} \quad \mathbf{w}_i = \frac{\mathbf{F}_i}{m_i} + \lambda_i\nabla f, \quad (8)$$

where $\lambda_i = \lambda/m_i$ $i = 1, 2$.

The same derivation can be generalized to the nonholonomic case (nonintegrable constraints). Since Gauss' principle treats the behavior of the system only at the present time, the nonstationary character of the constraint (i.e., its explicit time dependence) does not affect the derivation of (7) and (8).

Only the most general assumptions concerning the constraint equations (2) and the forces responsible for these constraints were used when deriving (7) and (8). As can be shown [3], the following are true for a system of two coupled point masses attracted by the coordinate origin:

(a) if the relative velocity v_0 of one of these points relative to the other is constant, then λ is also constant;

(b) the center of inertia of the points moves along an eclipse whose center is at the coordinate origin;

(c) when considering motion relative to the center of inertia, the straight line joining the two points will rotate uniformly with respect to the center of inertia, thus describing a plane.

Constraint Coefficient

We can write the vector sums of the Newtonian attraction forces acting on m_E and m_M as

$$m_E\ddot{\mathbf{r}}_E + \frac{k^2 m_E}{r_E^2}\frac{\mathbf{r}_E}{|\mathbf{r}_E|} = G\frac{m_E m_M}{|\mathbf{r}_M - \mathbf{r}_E|^2}\frac{(\mathbf{r}_M - \mathbf{r}_E)}{|\mathbf{r}_M - \mathbf{r}_E|}, \quad (9)$$

$$m_M\ddot{\mathbf{r}}_M + \frac{k^2 m_M}{r_M^2}\frac{\mathbf{r}_M}{|\mathbf{r}_M|} = G\frac{m_E m_M}{|\mathbf{r}_E - \mathbf{r}_M|^2}\frac{(\mathbf{r}_E - \mathbf{r}_M)}{|\mathbf{r}_E - \mathbf{r}_M|},$$

where $k^2 = Gm_s$. We divide the first equation of (9) by m_E and the second by m_M and introduce the dimensionless time $\tau = n^* t$, where n^* is the average sidereal motion of the Moon. In addition, we assume that the Moon rotates uniformly about the Earth in a circular orbit. We denote $\mathbf{r}'_E = n^{*^2}\mathbf{r}_E$, $\ddot{\mathbf{r}}'_E = n^{*^2}\ddot{\mathbf{r}}_E$, $\mathbf{r}'_M = n^{*^2}\mathbf{r}_M$, and $\ddot{\mathbf{r}}'_M = n^{*^2}\ddot{\mathbf{r}}_M$. Then, using the above notation, the system of equations (9) takes the form

$$\ddot{\mathbf{r}}'_E + \frac{k^2 n^{*4}\mathbf{r}'_E}{r'^3_E} = -\frac{\mu}{1+\mu}\Delta\mathbf{r}', \quad (10)$$

$$\ddot{\mathbf{r}}'_M + \frac{k^2 n^{*4}\mathbf{r}'_M}{r'^3_M} = \frac{1}{1+\mu}\Delta\mathbf{r}',$$

where $\mu = m_M/m_E$. We now combine the left-hand and right-hand sides of (10) to obtain

$$\left(\ddot{\mathbf{r}}'_E + \ddot{\mathbf{r}}'_M\right) + k^2 n^{*4}\left(\frac{\mathbf{r}'_E}{r'^3_E} + \frac{\mathbf{r}'_M}{r'^3_M}\right) = \frac{1-\mu}{1+\mu}\Delta\mathbf{r}'. \quad (11)$$

Using (7) and (8), we can write the right-hand side of (11) in the form

$$\frac{1-\mu}{1+\mu}\Delta\mathbf{r}' = \left(\frac{1}{m_E} + \frac{1}{m_M}\right)\lambda\nabla f. \qquad (12)$$

We denote $\tilde{\lambda} = \dfrac{m_M + m_E}{m_E m_M}\lambda$. Then,

$$\frac{1-\mu}{1+\mu}\Delta\mathbf{r}' = \tilde{\lambda}\nabla f. \qquad (13)$$

To find the coefficient $\tilde{\lambda}$, we must determine the scalar function f and calculate ∇f.

The kinematics of the Sun–Earth–Moon system correspond to the motion of a spatial triangle with revolving pairs. A schematic of the kinematic constraints is drawn in Fig. 2. The radius vectors $\mathbf{r}_1$, $\mathbf{r}_2$, and $\Delta\mathbf{r}$ correspond to those in Fig. 1. Here, X_1, X_2, and X_3 are the horizontal axes of the orthonormal bases for the corresponding bodies; ε is the inclination of the Earth's rotational axis Z_2 to the axis of the ecliptic Z_1; β is the angle between the axis of the lunar orbit Z_3 and Z_1; and q_1, q_2, and q_3 are positive angles (generalized parameters). In this case, the constraints are given by the equalities

$$\mathbf{r}_2 - \mathbf{r}_1 - \Delta\mathbf{r} = 0, \quad G_1^T G_2^T G_3 = E_3, \qquad (14)$$

where G_1, G_2, and G_3 are the transformation matrices and E_3 is a unit matrix.

The system in Fig. 2 possesses a cyclic structure. Precisely the presence of cycles is responsible for the constraints imposed on the mutual motions of the bodies forming the system of fundamental cycles.

The above constraint equations contain neither quantities associated with the specified absolute motion of the basis body (Sun) nor those characterizing the absolute motion of the Earth and Moon. This is quite natural, since the cycles are responsible for the internal coupling between the bodies, which do not depend on the absolute motion of the system. To get around difficulties associated with differentiating the constraint equations (14), we take into account the fact that, for any body of a given cycle, each geometric quantity determined in its intrinsic coordinate system can be represented by geometric quantities related to other bodies of the cycle and parameters specifying the relative motions at the cycle joints.

The parameters defining the motion of the system of point masses m_E and m_M relative to their center of mass (barycenter), which moves in a specified Keplerian orbit, are the radius vectors of the points ρ_E and ρ_M specified by the relations

$$\rho_E = -\frac{m_M}{m_M + m_E}\Delta\mathbf{r}, \qquad (15)$$

$$\rho_M = \frac{m_E}{m_M + m_E}\Delta\mathbf{r},$$

$$m_E\rho_E + m_M\rho_M = 0. \qquad (16)$$

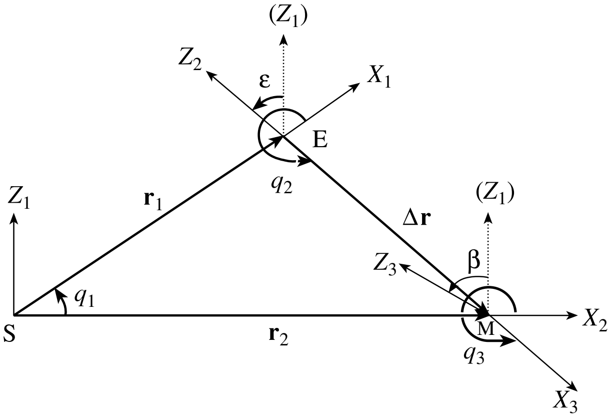

Fig. 2. Schematic of kinematic constraints in the Sun–Earth–Moon system.

We now combine the first and second equations of (15) to obtain

$$\rho_M + \rho_E = \frac{m_E - m_M}{m_M + m_E}\Delta\mathbf{r} \qquad (17)$$

or

$$\Delta\rho = \frac{m_E - m_M}{m_M + m_E}\Delta\mathbf{r}. \qquad (18)$$

We determine the scalar function f relating the coordinates of the points m_E and m_M from (18) by multiplying the right-hand and left-hand sides of this equation by n^{*2} and squaring them. We then have

$$f = |\Delta\rho'|^2 - \left(\frac{m_E - m_M}{m_E + m_M}\right)^2 |\Delta\mathbf{r}'|^2. \qquad (19)$$

Consequently,

$$|\nabla f| = -2\left(\frac{m_E - m_M}{m_E + m_M}\right)^2 |\Delta\mathbf{r}'| \qquad (20)$$

or

$$|\nabla f| = -2\left(\frac{1-\mu}{1+\mu}\right)^2 |\Delta\mathbf{r}'|. \qquad (21)$$

After substituting (21) into (13), we obtain

$$2\tilde{\lambda} = -\frac{1+\mu}{1-\mu}. \qquad (22)$$

The constraint coefficient $\tilde{\lambda}$ does not depend on the coordinate systems and is specified by the coupling of the physical characteristics of the bodies m_E and m_M.

If we do not use the adopted idealized scheme (representing the Earth and Moon as point masses), the complex system of two coupled bodies can be represented as two separate (partial) systems that are mutually coupled. The degree of physical coupling between such partial systems is specified by the parameter σ, which is called the coupling strength [4].

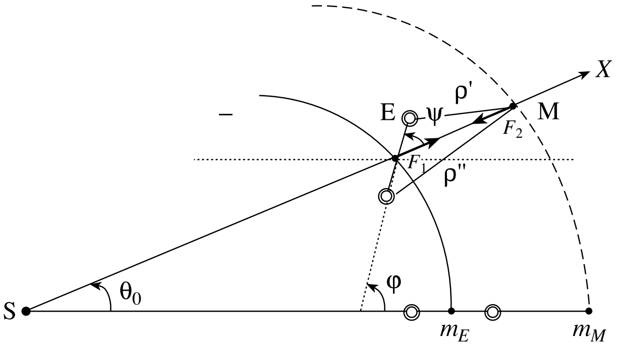

Fig. 3. Uniform rotation of the centers of masses of the Earth and Moon about the ecliptic axis.

Taking into account (22), the coupling strength of the Earth—Moon system is

$$\sigma = \left(\frac{1+\mu}{1-\mu}\right) / \left|\nu_1^2 - \nu_2^2\right|, \tag{23}$$

where ν_1 and ν_2 are the partial frequencies of the Earth and Moon. We can see that the coupling strength depends not only on the constraint coefficient but also on the closeness of the partial frequencies.

3. EIGENFREQUENCIES OF THE SYSTEM OF COUPLED BODIES

In a system of coupled bodies (such as the Earth—Moon system), the eigenfrequencies (partial frequencies) of the individual bodies differ from those of the system. This difference is determined by the amplitude of the constraint coefficient and the closeness of the partial frequencies, as can be seen in (23).

The appearance of a tendency for the oscillations of the Earth and Moon to become coherent is illustrated by the following idealized model of the system. Let us suppose that this model retains the structure of a real system determined by the number of bodies, their sequence, and the mechanism of their interaction described by the number of degrees of freedom in it.

We describe the motion of the centers of masses of the Earth and Moon in a polar coordinate system using some simplifications (Fig. 3). We take into account the extended nature of the Earth by representing it as two point masses $m = m_E/2$ separated by a distance of two Earth radii and rotating about the center of mass of the Earth with the period of its proper motion.

The classical linear theory is based on the empirical fact that the angle between the axis of rotation of the Moon and the axis of the ecliptic, and also the inclination of the lunar orbit to the ecliptic, are small. We shall assume that the centers of mass of the Moon and Earth rotate uniformly about the axis of the ecliptic in circular orbits under the action of a Newtonian force determined by the corresponding force functions with a period $T = 365.25$ days. If we neglect the interaction between the Earth and Moon, the motion of the point masses m_E and m_M will be stable with respect to the radius vectors of each point.

The interaction between the Earth and Moon results in oscillations about the stable motion determined by the external force. We identify the effect of the external force (from the Sun) with the imposition of stationary constraints in the coordinate system rotating with the radius vector SM (Fig. 3).

We define the forces of Newtonian interaction between the Earth and Moon $\mathbf{F}_1 = -\mathbf{F}_2$ to be internal forces in the Earth—Moon system. These forces are directed along the line joining the centers of mass of the Earth—Moon system and do not affect the equilibrium motion of the point masses m_E and m_M. This is true if we do not take into account the physical properties of the Earth—Moon system that constrain the motion of the individual bodies. For example, if the rotation and extent of the Earth are approximately taken into account, as is shown in Fig. 3, the effect of the lunar attraction force on the Earth will differ from the action of the same force on the point mass m_E. To illustrate this, we expand the force $\mathbf{F}_1$ into the two components $\mathbf{F}'_1$ and $\mathbf{F}''_2$, each directed along the lines ρ' and ρ'' joining the masses m and the center of mass of the Moon. We represent the resulting forces as the sum of tangential and normal components:

$$\mathbf{F}_1 = \mathbf{F}'_{1\Psi} + \mathbf{F}'_{1r}, \quad \mathbf{F}''_1 = \mathbf{F}''_{1\Psi} + \mathbf{F}''_{1r}. \tag{24}$$

Let us consider what relation should exist between $\mathbf{F}'_1$ and $\mathbf{F}_2$ in order for the system of point masses m and m_M to be in equilibrium with respect to the axis EX. The motion of the point m can be conveniently described by the polar coordinates r and ψ; the motion of m_M, by the Cartesian coordinates X and Y. In this case, the equilibrium state is equivalent to imposing ideal constraints of the form

$$r = R, \quad Y = 0. \tag{25}$$

Moreover, ψ and X' are related through the formula

$$X' = R\cos\psi + \rho'\sin q_1 \tag{26}$$

$$= R\cos\psi + \rho'\sqrt{1 - \left(\frac{R}{\rho'}\right)^2 \sin^2\psi},$$

where $\psi = \varphi - \theta_0$, R is the Earth's radius, and q_1 is the angle of inclination of ρ' to the axis EX. The statistical principle of virtual work in this case will take the form

$$RF_{1\psi}\delta\psi - F_2\delta X' = 0. \tag{27}$$

We now find the variation

$$\delta X' = -R\sin\psi \tag{28}$$

$$\times \left[1 + \frac{R\cos\psi}{\rho'\sqrt{1 - \left(\frac{R}{\rho'}\right)^2 \sin^2\psi}}\right]\delta\psi.$$

Substituting (28) into (27), we obtain

$$F'_{1\psi} = -F_2 \sin\psi \times \left[1 + \frac{R\cos\psi}{\rho'\sqrt{1 - \left(\frac{R}{\rho'}\right)^2 \sin^2\psi}}\right]. \qquad (29)$$

Thus, for equilibrium motion, it is necessary that the tangential component $F'_{1\psi}$ applied to the point m satisfy condition (29).

The same condition for the diametrically opposite point m has the form

$$F''_{1\psi} = -F_2 \frac{R\cos\psi\sin\psi}{\rho''\sqrt{1 - \left(\frac{R}{\rho''}\right)^2 \sin^2\psi}}, \qquad (30)$$

where ρ'' is the distance from the corresponding point to the center of mass of the Moon. If we take $\rho' \approx \rho'' \approx \rho$, then the torque will be

$$\tau = F'_{1\psi}R - F''_{2\psi}R \qquad (31)$$

or

$$\tau = -F_2 R \sin\psi. \qquad (32)$$

In this case, the constraints imposed on the normal components $\mathbf{F}'_{1r}$ and $\mathbf{F}''_{1r}$ follow from the definition of the interaction forces $\mathbf{F}_1$ and $\mathbf{F}_2$.

Variations of the physical conditions in the system result in small oscillations about the equilibrium motions. From a mathematical point of view, these small oscillations are associated with the quadratic dependence of the potential energy on the coordinates. In this case, if there is sufficient physical coupling between the bodies in the system and the free (partial) frequencies are quite close to each other, the oscillations will be "equalized;" i.e., eigenfrequencies of the system (different from the partial frequencies) will arise. The appearance of eigenfrequencies in the system of constrained bodies is a consequence of Gauss' basic principle of constrained systems. These eigenfrequencies are determined by the properties of the mechanical system.

Taking into account the kinematics of the coupled Earth—Moon system and the form of the constraint equation, the following statements are true [3]:

(1) the motion of the barycenter of the constrained system is subject to the same laws as is the free system;

(2) the law of areas is in effect;

(3) the principle of conservation of energy holds in the same form as for the free system.

We now find the interrelation between the partial frequencies and the eigenfrequencies of the coupled Earth—Moon system, where the basic frequencies are the eigenfrequency of the free motion of the Earth ν_1 and the eigenfrequency of the Moon ν_2. The basic requirement for ν_1 and ν_2 is that they be close to each other, which ensures the required coherence of the oscillations. Then, taking into account the constraint coefficient $\tilde{\lambda}$ and the corresponding choice of coordinates and initial conditions, we obtain a system of homogeneous equations with two degrees of freedom:

$$\begin{aligned} &\ddot{\xi} + \nu_1^2\xi - 2\tilde{\lambda}\left(\nu_2^2 - \nu_1^2\right)\eta = 0, \\ &\ddot{\eta} + \nu_2^2\eta - 2\tilde{\lambda}\left(\nu_2^2 - \nu_1^2\right)\xi = 0, \end{aligned} \qquad (33)$$

where ξ, η are coordinates characterizing small deviations from the equilibrium motions.

Solving the above equations, we obtain the possible proper oscillations of the system ω_1 and ω_2 for various ratios of the partial frequencies ν_1 and ν_2:

$$\begin{gathered} \nu_2 > \nu_1, \\ \omega_1^2 = k_1\nu_1^2 - k_2\nu_2^2, \quad \omega_2^2 = k_1\nu_2^2 - k_1\nu_1^2, \end{gathered} \qquad (34)$$

$$\nu_2 = \nu_1 = \omega_1 = \omega_2, \qquad (35)$$

$$\begin{gathered} \nu_2 < \nu_1, \\ \omega_1^2 = k_1\nu_2^2 - k_2\nu_1^2, \quad \omega_2^2 = k_1\nu_1^2 - k_2\nu_2^2. \end{gathered} \qquad (36)$$

The coefficients k_1 and k_2 ($k_1 > k_2$) are the roots of the equation

$$k^2 - k - 1.05043 = 0 \qquad (37)$$

and depend on the constraint coefficient ($\tilde{\lambda} = 0.512453$). For $\tilde{\lambda} = 0.5$, Eq. (5) takes the form

$$k^2 - k - 1 = 0; \qquad (38)$$

i.e., its roots are Fibonacci numbers ($k_1 = \Phi$, $k_2 = \Phi^{-1}$). It is interesting that these numbers arise in various ways in descriptions of parameters and proportionalities of the solar system. For example, the periods of revolutions and beat periods of the planets form a geometric progression with denominator Φ, and so on.

Let us take $\nu_1 = 0.90588$ yr^{-1} ($T_{\nu_1} = 403.2$ days) and $\nu_2 = 1$ yr^{-1} ($T_{\nu_2} = 365.25$ days). Substituting these values into (34), we obtain two eigenfrequencies of the Earth—Moon system: $\omega_1 = 0.84004$ ($T_{\omega_1} = 434.8$ days) and $\omega_2 = 1.05594$ ($T_{\omega_2} = 345.9$ days).

4. CONCLUSIONS

Oscillations in the Earth–Moon system obey the basic dynamical principle of constrained systems. The mutual influence of the partial frequencies gives rise to proper oscillations in the constrained system. This process is based on a complex redistribution of energy, involving all the movable elements of the system.

Using a model taking into account all the basic features of the real Earth–Moon system, we have derived a general law for the appearance of eigenfrequencies in the constrained system. In this model, taking into account the derived constraint coefficient $\tilde{\lambda}$ and the partial oscillations with periods $T_{\nu_1} = 403.2$ days and $T_{\nu_2} = 365.25$ days, the periods of proper oscillations in the coupled Earth–Moon system are $T_{\omega_1} = 434.8$ days and $T_{\omega_2} = 345.9$ days.

Therefore, the increase in the period of the proper oscillations of the Earth (the Chandler period) turns out to be a natural consequence of the appearance of eigenfrequencies in the constrained system.

Changing the physical conditions in the real Earth–Moon system (e.g., making the parameters dependent on energy) leads to deviations of the frequencies for the proper motion from the mean values specified by (33).

5. ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (project no. 98-02-16803).

REFERENCES

1. V. P. Demin, *The Motion of Satellites in a Noncentral Gravitational Field* [in Russian] (Nauka, Moscow, 1968).
2. G. S. Kurbasova and L. V. Rykhlova, Astron. Zh. **72**, 945 (1995) [Astron. Rep. **39**, 845 (1995)].
3. A. M. Lyapunov, *Lectures on Theoretical Mechanics* [in Russian] (Naukova Dumka, Kiev, 1982).
4. L. I. Mandel'shtam, *Complete Collected Works* [in Russian] (Akad. Nauk SSSR, Moscow, 1955), Vol. IV; *Lectures on Oscillation Theory* [in Russian] (Nauka, Moscow, 1972).

Translated by Yu. Dumin