Astronomy Reports, Vol. 49, No. 6, 2005, pp. 431–434. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 483–487.

The Active Galaxy NGC 3862 in a Compact Group in the Cluster A1367

A. A. Lipovka[1], N. M. Lipovka[2], I. V. Gosachinskiĭ[2], and E. Chavira[3]

[1] *Center of Physical Research, Sonora University, Aptdo. Postal 5-088, Hermosillo, Sonora, Mexico*

[2] *St. Petersburg Branch of the Special Astrophysical Observatory, Russian Academy of Sciences, Pulkovskoe sh. 65, St. Petersburg, 196140 Russia*

[3] *National Institute of Astronomy, Optics, and Electronics, Tonanzintla, Aptdo. Postal 51 y 216, Pueblo, Mexico*

Received November 6, 2003; in final form, December 3, 2004

Abstract—We study a compact group of 18 galaxies in the cluster A1367 with redshifts $z = 0.0208–0.025$. The group's center of activity in the radio is the galaxy NGC 3862, whose radio flux is an order of magnitude stronger than for the other members of the group. We present coordinates derived from the Palomar plate archive together with recessional velocities, and analyze other characteristics of the group's galaxies. The results of 1400 MHz observations of NGC 3862 with the RATAN-600 radio telescope are presented. These observations indicate that the galaxy's radio emission is variable. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The evolution of galaxies and their clusters is a central topic of modern astrophysics. Interest in these questions comes from attempts to follow the evolution of galaxies and explain the formation of galaxies and clusters of galaxies, as well as from possibilities for deriving information on the physical conditions in clusters, galaxies, and galactic nuclei. To achieve progress in these areas, it is important to study isolated, compact groups of galaxies in clusters together with their morphology, radio and optical spectra, and X-ray emission.

It is also interesting in its own right to study active galactic nuclei that display activity over a wide wavelength range, from the X-ray to the radio. This study deals with a compact group containing an active galaxy. We have already presented preliminary studies of galaxies in the clusters A569 [1], A1185 [2], A2151 [3], and a cluster in Cetus [4]. In those studies, we found and identified a number of radio objects with known radio spectra, whose optical properties are of considerable interest. Some 45% of these radio sources were identified with compact galaxies, and a large fraction have their strongest radio emission in their active nuclei.

Here, we study a compact group of galaxies surrounding NGC 3862 in the cluster A1367. The second section presents a catalog of the galaxies, while the third section discusses the radio properties of the galaxies based on our RATAN-600 observations and published data together with their optical properties.

2. THE CATALOG

A compact group of galaxies with a low velocity dispersion is present in the cluster A1367, in the vicinity of NGC 3862 (Fig. 1). The group is interesting because of its morphology, because it is distinct from the rest of the cluster members, and also because NGC 3862 is an active galaxy with one of the few known galactic optical jets.

When compiling the catalog, we used coordinates for the galaxies calculated at the Institute of Optics and Electronics (INAOE, Mexico) based on glass copies of the Palomar Sky Atlas plates using a blink comparator and the technique described in [5]. The rms uncertainties in the coordinates were calculated using reference stars from the PPMN catalog, and are $\Delta\mathrm{RA} = \pm 1.5''$, $\Delta\mathrm{DEC} = \pm 2.0''$.

The table presents the resulting coordinates of the galaxies, along with their optical characteristics, measured using an Automatic Plate Measuring facility [6]. The columns of the table present (1) a running number for each galaxy; (2, 3) the equatorial coordinates for equinox 2000.0; (4) the magnitude from [6]; (5) the angular size of the galaxy's major axis (calculated by us using data from [6], in arcseconds); (6) the ellipticity, $1 - b/a$, where a and b are the major and minor axes [6]; (7) the position angle, PA [6]; (8) the radial velocity from the NASA Extragalactic Database [7]; and (9) the galaxy's name [7]. Note that galaxies 12 and 13 are considered a single galaxy in [6], and we determined the sizes of these galaxies from the Palomar images.

As we noted above, this compact group is characterized by a relatively low velocity dispersion ($V_{mean} =$

1063-7729/05/4906-0431$26.00

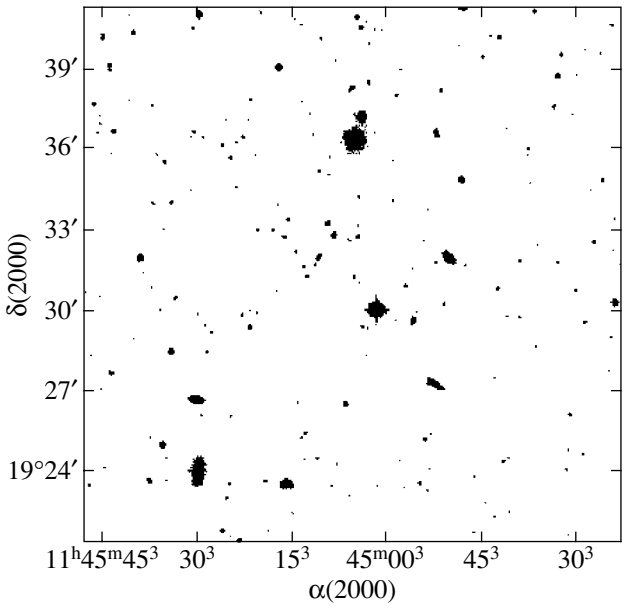


Fig. 1. The compact group of galaxies in the cluster A1367.

6450 ± 511 km/s), confirmation that the group galaxies are close together in space. Another interesting feature is displayed by the spatial positions of the galaxies: they form a chainlike curve, which can be fit with high significance. To estimate the probability for such an arrangement to occur by chance, we approximated the curved path with a second-order polynomial. The square deviation of the galaxy positions from the approximating curve is an order of magnitude lower than the density of galaxies in the vicinity of the compact group, providing evidence that the group is real and confirming that the arrangement of the galaxies in the plane of the sky is not random. The galaxies may have formed along the ridge of an expanding front, or might be the result of the ejection of matter from NGC 3862 that moved along the path now traced by the galaxy positions, gradually sweeping up matter and causing the galaxies to grow. This last possibility is supported by the fact that the galaxies' sizes increase and their color indices change systematically with increasing distance from the central galaxy in the group, NGC 3862: the closer to NGC 3862, the redder the galaxies' colors. On average, the colors of the galaxies that are closest to NGC 3862 are 0.6^m redder than the colors of galaxies at the group periphery, providing evidence for strong extinction in the neighborhood of NGC 3862. To conclude this section, we note that the statistical mean magnitude for this group of galaxies is $12.2^m \pm 2.5^m$ and the mean size is $30'' \pm 18.0''$ [6], corresponding to a mean galaxy size of 12 kpc for a mean recessional velocity of 6450 km/s. The size of NGC 3862 is 17.8 kpc for its recessional velocity of 6511 km/s and $H = 75$ km s^{-1}Mpc^{-1}.

3. THE RADIO PROPERTIES OF NGC 3862 AND OUR RATAN-600 RADIO OBSERVATIONS

Weak 21-cm radio emission was found for galaxies 2 and 4 in the group (with the fluxes being 111.9 and 81.5 mJy, respectively) [8]; weak radio emission for galaxy 5 with a flux of 10 mJy was also reported in [9]. The group's radio activity center is NGC 3862, for which an abundance of radio data are available [7–9].

It was found in 1968 [10] that the size of the radio source 3C 264 (NGC 3862) increased with decreasing frequency, which was interpreted as a spectral-index gradient across the galaxy. Further higher-resolution studies [11] at 11 cm ($\theta = 12'' \times 34''$) and 6 cm ($\theta = 6.5'' \times 19''$) revealed the presence of a large-scale asymmetry in the galaxy's radio-brightness distribution to the northeast, leading to the galaxy's classified as a head–tail radio galaxy.

Further observations [12] at 1465 MHz with even higher resolution, $\theta = 2.7'' \times 4''$, demonstrated that NGC 3862 has a nucleus that is $<3''$ in size and an extended component elongated toward the northeast. Bridle and Vallee [12] explained the large-scale structural asymmetry of NGC 3862 as the result of the diffusion of relativistic electrons that were left behind the moving galaxy. However, later observations [8, 9] demonstrated the presence of two features toward the northeast which, in our opinion, are related to the activity of NGC 3862 and represent periodic ejections of matter toward PA $= 32°–37°$.

The complete radio spectrum of NGC 3862 is presented in Fig. 2. Note the scatter of the flux densities over a wide frequency range.

Jointly analyzing the data with various resolutions, we can distinguish and construct the spectra for the corona and a nuclear region about $3''$ in size (marked "A" in Fig. 2). In Fig. 2, measurements from 1964 to 1980 are displayed using circles, whereas measurements obtained after 1980 are shown by crosses. The flux density appears to slowly decrease with time.

To confirm earlier suspicions of radio variability of NGC 3862, we obtained observations with $160''$ resolution using the RATAN-600 radio telescope in March 2003. Our flux-density measurements at 21 cm were obtained using feed cabin 2 of the northern sector of the RATAN-600 radio telescope. We carried out seven independent observations to improve the accuracy of our 21-cm flux measurements.

Parameters of the studied galaxies

No.	RA (2000.0)	DEC (2000.0)	R	θ	$1-a/b$	PA	v, km/s	Galaxy name
1	$11^h44^m47.61^s$	19°34′59.7″	12.68^m	23.5°	0.07	7°	6256	
2	11 44 49.79	19 32 4.5	9.53	49.1	0.38	46	6255	NGC 3857
3	11 44 51.76	19 36 39.4	12.10	29.0	0.36	14	6556	
4	11 44 51.91	19 31 55.2	17.31	12.3	0.38	125		
5	11 44 52.08	19 27 21.7	9.49	61.2	0.65	59	5468	NGC 3859
6	11 44 55.56	19 29 41.9	12.58	24.5	0.35	172	6457	
7	11 45 2.60	19 38 37.5	15.63	14.2	0.17	23		
8	11 45 3.69	19 37 20.0	10.94	29.3	0.13	46	6908	IC 2955
9	11 45 4.43	19 41 1.4	13.70	17.2	0.05	34		
10	11 45 4.83	19 36 32.6	9.38	41.5	0.08	96	6511	NGC 3862
11	11 45 4.89	19 37 26.4	19.32	8.5	0.20	117		
12	11 45 4.96	19 38 26.8		10				
13	11 45 5.43	19 38 20.0		8				
14	11 45 15.17	19 23 34.5	9.94	41.7	0.28	84	6997	NGC 3864
15	11 45 29.38	19 24 7.1	9.40	75.1	0.72	175	7457	NGC 3867
16	11 45 29.80	19 26 44.9	9.68	47.5	0.49	80	6386	NGC 3868
17	11 45 34.91	19 25 6.0	12.99	22.0	0.08	21		
18	11 45 36.94	19 23 41.5	14.36	18.2	0.25	35		

The system noise temperature was ≈70 K, the continuum bandwidth 10 MHz, and the time constant 6 s. We used 3C 123, which has a flux of 46.61 Jy at 21 cm, as a reference source. The observations and data reduction were done using standard software written at the RATAN-600 radio-spectrometry complex. The effective area of the telescope at the elevation of 3C 123 was 982 m^2, and the right-ascension correction was found to be -0.38^s.

We reduced the observations using the standard RATAN-600 software for the reduction of spectroscopic data. We measured the observational parameters using code designed to carry out a Gaussian analysis of the transit curves after correction for smoothing by the output device of the continuum channel. The resulting mean parameters for five observations were $\mathrm{RA}(2000.0) = 11^h45^m06.12^s \pm 0.2^s$ and the flux at 21 cm $P = 4.70 \pm 0.1$ Jy.

Previous observations at 21 cm obtained in 1981 and 1995 yielded $P = 5.94 \pm 0.17$ Jy [13] and $P = 5.45 \pm 0.12$ Jy [14]. Our 2003 RATAN-600 flux, $P = 4.70 \pm 0.1$ Jy, is 1.24 ± 0.27 Jy lower than the value observed in 1981, representing a decrease by about 20% in 22 years.

Let us now estimate the characteristic magnetic fields for the two emitting regions: the corona and plane of the galaxy.

Combined with the higher-resolution observations [8, 9, 15], our RATAN-600 data can be used to separate out the individual radio-emitting components and reveal the nature and character of the radio emission from these components. The galaxy's corona is responsible for more than 40% of the radio emission, and has a spectral index of $\alpha = 0.95$, whereas the spectral index of the radio emission from the galactic plane is $\alpha = 0.67$.

We can use the known flux density, size of the emitting region, and distance to the galaxy to estimate the magnetic-field strength (see, for example, [16]). Our estimated magnetic-field strength in the plane of NGC 3862 is $H(\mathrm{pl}) = 3.5 \times 10^{-6}$ Oe for $\alpha = 0.67$, with the field in the corona being $H(\mathrm{cor}) = 0.86 \times 10^{-6}$ Oe for $\alpha = 0.95$.

It is also of interest to compare the jet's physical parameters derived from optical data with those we have estimated from the radio emission. For this purpose, we use recent data on the optical activity of the jet of NGC 3862. Hubble Space Telescope

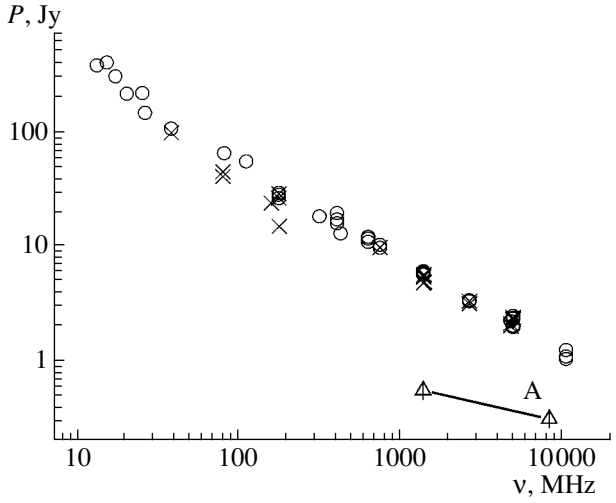


Fig. 2. Radio spectrum of NGC 3862; the spectrum of the compact nucleus is marked "A."

high-resolution optical measurements at 3400 Å and 4850 Å [17] display a compact core and a 270-pc-long jet in position angle PA = 37°, close to the direction of the radio jets [8, 9]. Crane *et al.* [17] conclude that the optical jet emission is synchrotron radiation with a spectral index of 1.4, and that the emission from the nucleus has a spectral index of $\alpha = 1.2$.

Using the standard known formulas for synchrotron radiation [16], we find that the 3400 Å and 4850 Å radiation are emitted by relativistic electrons with energies of $E = 1.8 \times 10^4 - 1.8 \times 10^6$ MeV in a magnetic field of $H = 5 \times 10^{-4}$ Oe, with the density of the relativistic electrons being $K = 10^{-9} - 10^{-11}$ erg/cm^3. The lifetime of such electrons is estimated to be about 10^3 years, in good agreement with the size of the optical jet.

4. CONCLUSIONS

We have studied a compact group of galaxies in the vicinity of NGC 3862 in the cluster A1367 ($z = 0.0208 - 0.025$). We measured the coordinates of the 18 members of the group with the blink comparator of the Institute of Optics and Electronics using the coordinates of reference stars. Four group members have been found to have radio emission, with the most active galaxy in the radio being NGC 3862, which may be ejecting matter in position angle PA = 32°–37°, observed as a jet in both the radio and optical. By analyzing this galaxy's radio emission observed with various resolutions over a wide wavelength range, we were able to separate out the radio emission from the galaxy's nucleus, plane, and corona. We find a flattening of the radio spectrum toward the nucleus. We had earlier obtained a similar result based on analysis of a large amount of data for the cluster A569 [1].

Our RATAN-600 observations of NGC 3862 demonstrate that the 21-cm emission is time-variable, with the 21-cm flux currently decreasing by approximately 70 mJy per year. Given the small uncertainties in the measured fluxes, we consider this result to be firm. Since the radio structure is associated with a jet, we suggest that the activity of NGC 3862 is periodic.

ACKNOWLEDGMENTS

We thank the anonymous referee for helpful comments. One of us (N.M.L.) thanks the Sonora University (Mexico) for hospitality and for support kindly provided for this study through grant no. P/PIFOP 2002-26-01.

REFERENCES

1. A. A. Lipovka and N. M. Lipovka, Astron. Zh. **79**, 963 (2002) [Astron. Rep. **46**, 867 (2002)].
2. A. M. Botashov *et al.*, Astron. Zh. **76**, 83 (1999) [Astron. Rep. **43**, 65 (1999)].
3. E. Chavira *et al.*, Preprint No. 140, SPbGU (St. Petersburg State Univ., St. Petersburg, 2002).
4. N. M. Lipovka, A. A. Lipovka, O. V. Verkhodanov, and E. Chavira, Astron. Zh. **77**, 3 (2000) [Astron. Rep. **44**, 1 (2000)].
5. E. Chavira-Navarerete, O. V. Kiyaeva, N. M. Lipovka, and A. A. Lipovka, Preprint No. 81, SPbGU (St. Petersburg State Univ., St. Petersburg, 1992).
6. M. Irwin, http://www.ast.cam.ac.uk/apmcat/ (1998).
7. O. V. Verkhodanov, S. A. Trushkin, H. Andernach, and V. N. Chernenkov, in *Astronomical Data Analysis Software and Systems VI*, Ed. by G. Hunt and H. E. Payne, ASP Conf. Ser. **125**, 322 (1997).
8. R. L. White, R. H. Becker, *et al.*, Astrophys. J. **475**, 479 (1997); http://third.llnl.gov/cgi-bin/firstcutout.
9. J. J. Condon, W. D. Cotton, *et al.*, The NRAO VLA Sky Survey; http://www.nrao.edu/NVSS/postage.html (1996).
10. G. H. Macdonald, S. Kenderdine, and A. C. Neville, Mon. Not. R. Astron. Soc. **138**, 259 (1968).
11. K. J. Northover, Mon. Not. R. Astron. Soc. **177**, 307 (1976).
12. A. H. Bridle and J. P. Vallee, Astron. J. **86**, 1165 (1981).
13. H. Kuhr *et al.*, Astron. Astrophys., Suppl. Ser. **45**, 367 (1981).
14. M. Ledlow *et al.*, Astron. J. **109**, 853 (1995).
15. MGV: MIT...T00L...Herold, Lori: Identifications of Radio Sources in the MG-VLA Survey (1996).
16. S. A. Kaplan and S. B. Pikel'ner, *Interstellar Medium* (Fizmatgiz, Moscow, 1963) [in Russian].
17. P. Crane *et al.*, Astrophys. J. Lett. **402**, L37 (1993).

Translated by N. Samus'

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 435–445. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 488–499.

The Galactic Constants and Rotation Curve from Molecular-Gas Observations

V. S. Avedisova

Institute of Astronomy, Moscow, Russia

Received March 2, 2003; in final form, December 3, 2004

Abstract—We obtained the photometric distances and radial velocities for the molecular gas for 270 star-forming regions and estimated the distance to the Galactic center from ten tangent points to be $R_0 = 8.01 \pm 0.44$ kpc. Estimates of R_0 derived over the last decade are summarized and discussed; the average value is $R_0 = 7.80 \pm 0.33$ kpc. We analyze deviations from axial symmetry of the gas motion around the Galactic center in the solar neighborhood. Assuming a flat rotation curve, we obtain $\Theta_0 \sim 200$ km/s for the circular velocity of the Sun from regions beyond the Perseus arm. We used these Galactic constants to construct the Galactic rotation curve. This rotation curve is flat along virtually its total extent from the central bar to the periphery. The velocity jump in the corotation region of the central bar in the first quadrant is 20 km/s. We present analytical formulas for the rotation curves of the Northern and Southern hemispheres of the Galaxy for $R_0 = 8.0$ kpc and $\Theta_0 = 200$ km/s. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The distance to the center of the Galaxy R_0 is a key parameter for determining the physical characteristics of the Galaxy and Galactic objects. In 1985, the International Astronomical Union Assembly adopted the value $R_0 = 8.5$ kpc. However, more modern determinations with continually decreasing scatter are converging to a somewhat lower value. Especially important in this regard are results based on the motion of the star S2 around the black hole at the center of the Galaxy [1]. Estimates of another Galactic constant, Θ_0—the velocity of the Galactic rotation at the distance R_0—display a large scatter. Another fundamental characteristic of the Galaxy is its rotation curve, or the dependence of the rotational velocity on distance from the Galactic center, which is related to the mass distribution in the Galaxy. Along with the Galactic constants, the rotation curve serves as the main tool for determining the distances to objects of the disk component of the Galaxy with known velocities.

In contrast to other galaxies, whose rotation curves can essentially be determined directly from observations, the large uncertainty in the photometric distances to stars poses a major problem in the case of our own Galaxy. The location of the Sun in the Galactic plane and the presence of dust in the interstellar medium greatly reduce the possibilities of observing some objects in the optical, especially toward the inner region of the Galaxy, where only the nearest 2 to 3 kpc are accessible for observation. Only radio astronomical observations in HI and CO lines can be used to derive the rotation curve in the central regions of the Galaxy. The rotation curve in the outer Galaxy can be determined only using stars with known photometric distances. Attempts have been made to determine the rotation curve using various types of very distant objects with known velocities and independently determined distances, such as young OB stars and open clusters [2–4], classical Cepheids [5–7], and planetary nebulae and AGB-stars [8]. In the radio, rotation curves can be derived from the terminal velocities corresponding to the steep wings of the HI and CO line profiles observed in the Galactic plane in the direction of the longitudes of the first and fourth quadrants (assuming that the angular velocity of the Galaxy decreases monotonically with Galactocentric radius) [9–12]. As a rule, the rotation curves derived from different objects differ; this is due, first and foremost, to the different velocity dispersions of these objects, associated with differences in their ages. The lowest velocity dispersions are possessed by the gaseous component of the Galaxy and young stars. This makes them especially sensitive to small perturbations of the gravitational potential due to the spiral arms and other resonances. Rotation curves derived from gas have wavelike humps, some of which are due to the systematic velocities in the spiral-density waves. As we consider older objects, their velocity dispersion increases, their kinematics become closer to dynamical equilibrium, and their rotation in the Galaxy is described by smoother curves. Therefore, it is necessary to take into account these dependences for a given class of object when determining the

1063-7729/05/4906-0435$26.00

Table 1. Mean calibration of the absolute magnitudes of O stars based on data from [27, 28]

Sp	I	Ia	Iab	Ib	II	III	IV	V
O2	−6.95					−6.00		−5.60
O3	−6.95					−6.00		−5.60
O3.5	−6.95					−6.00		−5.60
O4	−6.60					−6.00		−5.50
O5	−6.60					−5.00		−5.00
O5.5	−6.60					−5.00		−5.00
O6		−6.60	−6.60	−6.60	−5.40	−5.40	−4.90	−4.90
O6.5		−6.60	−6.60	−6.60	−5.40	−5.40	−4.90	−4.90
O7		−6.50	−6.50	−6.50	−5.40	−5.40	−4.90	−4.90
O7.5		−6.50	−6.50	−6.50	−5.40	−5.40	−4.70	−4.70
O8		−6.45	−6.45	−6.45	−5.40	−5.40	−4.70	−4.70
O8.5		−6.40	−6.40	−6.40	−5.40	−5.40	−4.50	−4.50
O9		−6.00	−6.00	−6.00	−5.20	−5.20	−4.20	−4.20
O9.5		−6.00	−6.00	−6.00	−5.20	−5.20	−4.20	−4.20
O9.7		−6.05	−6.05	−6.05	−5.20	−5.15	−4.20	−4.10
B0		−6.10	−6.10	−6.10	−5.20	−5.10	−4.20	−4.00

distances for a certain class of objects by means of its rotation curve.

When determining the Galactic constants and rotation curve, a substantial role is played by the assumption of axial symmetry of the Galaxy and its circular rotation. However, surveys of atomic and molecular gas in the central region of the Galaxy revealed strong asymmetries. In the 1970s, the existence of a bar was suggested as one of the explanations for the unusual kinematics of the gas in the central part of the Galaxy [13, 14]. However, the hypothesis of radial outflows of gas associated with the 3-kpc arm was then the predominant interpretation of this kinematic asymmetry. Only new observations obtained in the 1990s [15–18] provided independent evidence for the presence of a bar in the central region of the Galaxy. A crucial role in the discovery of the bar was played by data provided by the COBE/DIRBE infrared space telescope [19–21]. Different models give slightly different parameters for the bar. Estimates of the bar corotation radius are in the range 3.5–5 kpc, and the major axis is oriented at an angle $\phi \sim 15°–45°$ relative to the direction toward the Galactic center. According to [22, 23], the bar is a triaxial structure with axial ratios $a : b : c \sim 1 : 0.6 : 0.4$ and $\phi \sim 25°$.

There is no doubt that the presence of the central bar must influence the kinematics of the outer regions. Analysis of the distribution of radial and rotational velocities relative to the Local Standard of Rest (LSR) in the nearest vicinity of the Sun revealed a well-defined stellar component that rotates more slowly and has radial velocities $u \leq -30$ km/s. This component is formed of old stars, and is related to the outer Lindblad resonance of the central bar [24]. Models of the gas kinematics for the bar and disk [21, 25] show that, beyond the bar corotation region, the motion of the gas gradually approaches the expected circular motion.

The aim of the present paper is to determine the Galactic constants and construct the rotation curve using improved data on the distances and velocities of star-forming regions (SFRs). In Section 2, we present a list of HII regions or SFRs with revised photometric distances and radial velocities, and, in Section 3, estimates of the distance to the center of the Galaxy, R_0, using tangent points. A comparison of the estimates for R_0 in the first and fourth quadrants in another method suggests that the shape and kinematics of the Galaxy in the solar neighborhood are possibly asymmetric. Section 4 discusses data on the Galactic constants obtained over the last decade in the context of our current understanding of Galactic structure, and derives a most probable value of R_0. In Section 5, we estimate the Galactic rotational velocity at the solar Galactocentric distance. In

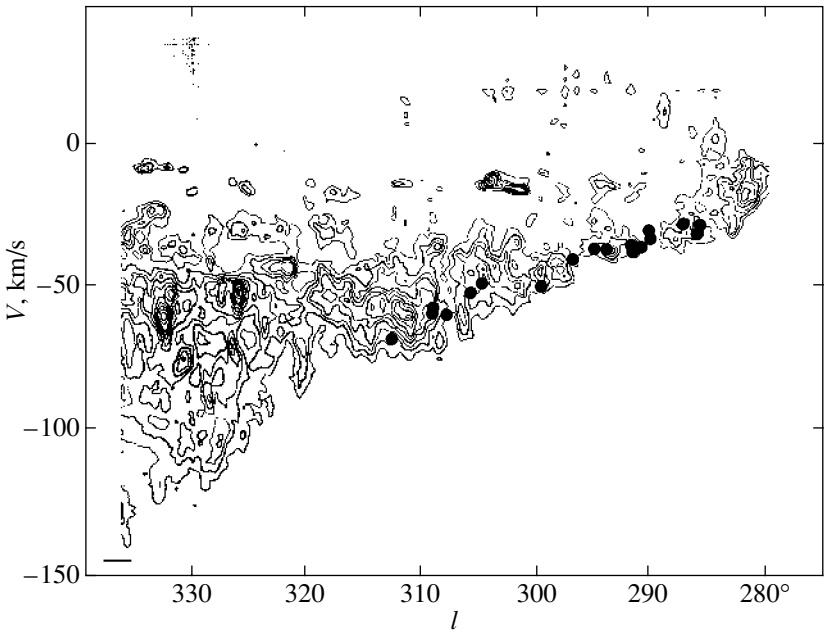


Fig. 1. $l - V$ contour map of CO emission [32] for the fourth quadrant. The filled circles show the positions of SFRs located at tangent points.

Section 6, we discuss currently available data on Θ_0 and $A - B$; in Section 7, we construct the Galactic rotation curve for the northern and southern sky.

2. THE INPUT DATA

We used SFRs for our study, mainly diffuse nebulae (HII regions) with known exciting stars. Radial velocities were taken from the catalog of star-forming regions [26], mainly for the molecular gas surrounding young stars. The distances to the SFRs were partially taken to be the distances to the clusters associated with SFRs and partially redetermined using published data on the exciting stars and applying a unified spectral class–absolute magnitude calibration. The calibration for the O stars was taken from several papers of Walborn, a long-time researcher on the spectral classification for massive stars, and his coauthors. Table 1 is based on Walborn's papers [27, 28]. For the B stars, we used the calibration of Vacca *et al*. [29]. The reddening coefficient R_v for the O stars was taken to be 3.2 after [30]. When O stars were included in the new catalog of O stars [31], we took the corresponding photometric and spectral data from this catalog. The BV photometry and spectra of the remaining stars were checked against the data in the SIMBAD database. This yielded the data in Table 2, which is available at http://cdsweb.u-strasbg.fr/cats (it is presented in electronic form only). This table contains the main characteristics of 270 Galactic SFRs, as well as the radial velocities of the molecular clouds associated with them. The columns of Table 2 give the name of the nebula or young stellar cluster, its Galactic coordinates l and b, its photometric distance and the corresponding error, its radial velocity reduced to the LSR, and the references to the data on the exciting star and the data used to determine the distance (the star's name, V and $B - V$, and the distance derived from them).

3. DETERMINATION OF R_0

To investigate the possible asymmetry of the Galaxy, all estimates were derived separately for the first and fourth quadrants. We used previously known methods to determine R_0. The first is to estimate R_0 based on data for SFRs located close to tangent points. If a region at Galactic longitude l is located at a tangent point and its distance to the Sun is known, $R_0 = D \sin l \cos b$, where D is the photometric distance to the SFR (we assume here that young stars have circular orbits).

To find SFRs located close to tangent points, all SFRs with known photometric distances were plotted on the $(l - V)$ diagram for the Galactic CO survey [32]. We selected SFRs located at the outer high-velocity ridge of the diagram, which should correspond to tangent points. Figure 1 shows the $l - V$ diagram for the fourth quadrant; the positions of the SFRs are plotted by filled circles. There are three such regions in the first quadrant and 18 in the fourth quadrant. We introduced an additional criterion to be sure that the selected regions are indeed close to tangent

Table 3. Parameters of SFRs located at tangent points

No.	l, deg	b, deg	D, kpc	V_{LSR}, km/s	R_0, kpc	Refs.	Comments
1	286.21	−0.20	2.23 ± 0.20	−21.5	7.99	[33]	1
2	290.35	1.62	2.90 ± 0.30	−19.0	8.34	[33, 34]	2
3	290.60	0.31	2.90 ± 0.40	−23.4	8.24	[33]	3
4	291.30	−0.69	2.70 ± 0.30	−24.0	7.43	[34]	4
5	296.22	−3.55	3.60 ± 0.50	−30.0	8.15	[35]	
6	298.36	2.23	3.75 ± 0.40	−29.6	7.89		5
7	305.71	1.53	4.80 ± 1.00	−57.1	8.22	[31]	6
8	307.85	0.21	4.50 ± 1.00	−50.0	7.33	[36]	7
9	309.27	−0.46	5.35 ± 0.50	−48.9	8.69	[36]	8
10	312.51	−2.69	5.64 ± 0.70	−47.7	8.35	[36]	9

Comments (the last column):
1. NGC 3324:
   LSS 1695, O8.5V, $V = 8.233$, $B - V = 0.114$, $D = 1.87$ kpc;
   LSS 1697, O6.5V, $V = 7.818$, $B - V = 0.148$, $D = 1.74$ kpc;
   LSS 1714, B1.5V, $V = 10.89$, $B - V = 0.39$, $D = 2.15$ kpc.
2. St 13, $D = 2.70$ kpc:
   LSS 2231, B0V, $V = 9.30$, $B - V = -0.01$, $D = 3.00$ kpc;
   HD 097848, O8V, $V = 8.68$, $B - V = -0.01$, $D = 3.00$ kpc.
3. NGC 3572, $D = 2.7$ kpc:
   LSS 2199, O7.5III, $V = 7.888$, $B - V = 0.06$, $D = 2.62$ kpc;
   LSS 2171, O4V, $V = 8.256$, $B - V = 0.10$, $D = 2.99$ kpc;
   LSS 2210, B0II, $V = 8.77$, $B - V = 0.20$, $D = 3.16$ kpc;
   LSS 2211, B0V, $V = 10.08$, $B - V = 0.22$, $D = 3.04$ kpc.
4. LSS 2217, O7.5 III, $V = 8.67$, $B - V = 0.19$, $D = 3.10$ kpc;
   LSS 2226, O7.5III, $V = 8.07$, $B - V = 0.16$, $D = 2.46$ kpc;
   LSS 2233, B0.5V, $V = 9.23$, $B - V = 0.05$, $D = 2.26$ kpc;
   LSS 2253, O9.5/B0V, $V = 9.61$, $B - V = 0.22$, $D = 2.71$ kpc;
   LSS 2255, B1V, $V = 10.16$, $B - V = 0.11$, $D = 2.72$ kpc;
   LSS 2151, B0.5IV, $V = 9.67$, $B - V = 0.16$, $D = 3.03$ kpc.
5. LSS 2620, O9.5V, $V = 10.23$, $B - V = 0.26$, $D = 3.35$ kpc;
   LSS 2625, B2III, $V = 11.01$, $B - V = 0.39$, $D = 3.80$ kpc;
   LSS 2626, O6V, $V = 10.42$, $B - V = 0.37$, $D = 4.13$ kpc.
6. LSS 2997, O6.5Iab, $V = 9.50$, $B - V = 0.53$, $D = 4.81$ kpc.
7. LSS 3119, O6I, $V = 9.19$, $B - V = 0.49$, $D = 4.5$ kpc.
8. LSS 3159, B1Ib, $V = 10.24$, $B - V = 0.55$, $D = 5.35$ kpc.
9. LSS 3225, B1.5III, $V = 11.43$, $B - V = 0.32$, $D = 5.6 \pm 1.3$ kpc.

points: that the error in the estimated distance to an SFR due to the difference between the position of the SFR and the tangent point be smaller than the error in the distance itself, σ_D. This can be expressed as the inequality $R\sqrt{1 - (R_0 \sin l/R)^2} < \sigma_D$, where R is the Galactocentric distance of the SFR and σ_D is the error in the distance to the region D. R_0 was varied between 6.5 and 8.5 kpc. Ten SFRs were left in the fourth quadrant after applying this criterion. Table 3 lists the Galactic coordinates, photometric distances and their errors, and radial velocities relative to the LSR of these SFRs, together with the calculated values of R_0. The resulting weighted mean value $R_0 = 8.01 \pm 0.44$ kpc was found using the data for all the SFRs, with the weight of an SFR being the inverse of the square of the distance error.

The second method we used to estimate R_0 requires the radial velocities of the SFRs. We approximated the rotational velocities of the SFRs in the first and fourth quadrants by the rotational velocities of the CO gas in them, which are determined from their terminal velocities (taken from [9] for the first quadrant and from [11] for the fourth quadrant). We assumed that the molecular gas and SFRs have similar kinematics. This procedure was applied separately to the first and fourth quadrants.

We selected from Table 2 all SFRs at longitudes $15° < l < 75°$ and $280° < l < 350°$ with positive velocities in the first quadrant and negative velocities in

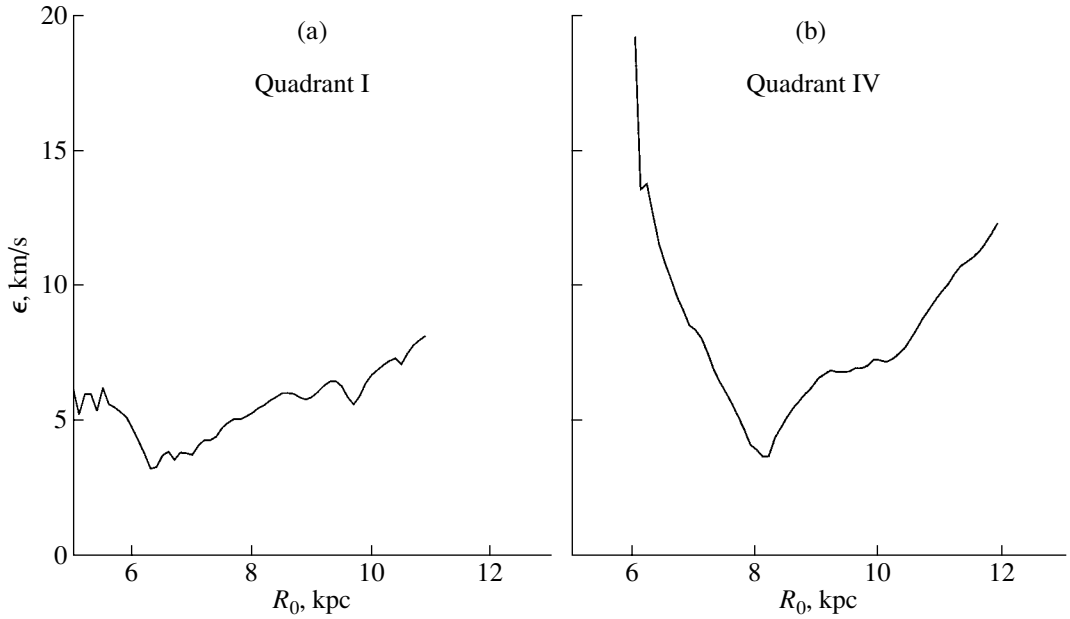


Fig. 2. The difference between the circular velocities of the SFRs and the CO gas as a function of the adopted R_0 for (a) the first quadrant and (b) the fourth quadrant.

the fourth quadrant. We determined the Galactocentric distances R and circular velocities $\Theta(\mathrm{HII})$ of these regions for the grid of values $R_0 = (6; 12)$ kpc and $\Theta_0 = (180; 250)$ km/s. We found $R(\mathrm{CO})$ and $\Theta(\mathrm{CO})$ from the terminal velocities for every longitude in the CO-survey. We then determined for every node of the grid the difference $\Theta(\mathrm{SFR}) - \Theta(\mathrm{CO})$ averaged over all HII regions, with weights that were proportional to the inverse of the error in the Galactocentric distance to the region:

$$\epsilon(R_0) = \frac{\sum_{i=1}^{i=n}[\Theta_i(\mathrm{CO}) - \Theta_i(\mathrm{HII})]/\delta R_i(\mathrm{HII})}{\sum_{i=1}^{i=n}(1/\delta R_i(\mathrm{HII}))},$$

where n is the number of HII regions that are used to determine R_0. The results are virtually independent of the Galactic constant Θ_0. The dependences of ϵ on R_0 for the first and fourth quadrants are shown in Figs. 2a and 2b. Note that we rejected regions that gave rise to sudden sharp peaks in this smoothly varying dependence from the initial list of SFRs. This is quite justified, since the distance dependence of the rotational velocity of the CO gas is far from smooth. In the first quadrant, we rejected the SFRs G55.84−3.79, G76.88+1.95, and G80.38+0.40. The plot in Fig. 2a is based on 32 SFRs. In the fourth quadrant, we rejected 284.75−3.06, 286.21−0.20, 287.28−0.88, 294.57−1.11, 299.30−0.31, 309.27−0.46, and 312.51−2.69. The remaining 23 SFRs yielded the curve in Fig. 2b. The curves in Figs. 2a and 2b are completely different. Figure 2b clearly shows a deep minimum at 8.0 kpc, while a trend toward lower R_0 and a broad minimum near 6.5 kpc can be seen in Fig. 2a (for the first quadrant). Such differences could be due to asymmetry in the orbital kinematics of objects in the first and fourth quadrants. However, this hypothesis requires additional study, which will become possible only when the distances to the HII regions are known more accurately.

The weighted mean of these R_0 estimates is close to 8.0 ± 0.36 kpc, since the estimates obtained using the second method have appreciably larger standard errors.

Thus, our analysis using data on SFRs in the solar neighborhood has revealed a possibility for asymmetry in the kinematics of the Galaxy relative to its center. If confirmed, this asymmetry could be related either to the presence of a bar in the Galactic center or to the fact that, according to the new data of [37], the Galaxy is a triaxial ellipsoid.

4. ADOPTED VALUE OF R_0

Table 4 lists various values of R_0 that have been obtained, mainly after the HIPPARCOS experiment, together with the corresponding errors, references, and the types of objects used for the estimates. These data yield the weighted mean value $R_0 = 7.80 \pm 0.32$ kpc. We have assigned the most confidence to estimates obtained using objects in the vicinity of the Galactic nucleus, since they are not affected by possible asymmetry of the kinematic data, and essentially provide a geometrical determination of the distance to the Galactic center from the orbit of the star S2 around the central black hole—Sgr A*. It is clear that R_0 lies in the range 7.5–8.2 kpc. At the

Table 4. Summary of R_0 estimates

R_0, kpc	Refs.	Objects
7.9 ± 0.8	[12]	Hydrogen HI
7.6 ± 0.4	[38]	Globular clusters
7.5 ± 1.0	[10]	Hydrogen HI
8.1 ± 0.3	[5]	278 Cepheids
7.1 ± 0.5	[7]	Cepheids
8.3 ± 1.0	[39]	RR Lyrae stars in the Galactic center
7.66 ± 0.32	[40]	Cepheids
7.1 ± 0.4	[41]	Gas distribution
8.4 ± 0.4	[42]	Red stars in the Galactic bulge
8.2 ± 0.3	[43]	Red stars in M31 and Galactic center
7.3 ± 0.3	[44]	Cepheids, open clusters, red giants
7.9 ± 0.3	[45]	RR Lyr and δ Scu stars in the Galactic center
7.9 ± 0.85	[46]	Cluster in the Galactic center
7.94 ± 0.42	[1]	The orbit of S2 around the Galactic center
8.2 ± 0.40	[47]	Open clusters
8.0 ± 0.36	This paper	Star-forming regions

same time, modeling of the Galactic bar and disk after the DIRBE experiment shows that, if $R_0 = 7.5$ kpc, the local extinction in the J band does not agree with admissible extinction in the V band. Thus, it is most likely that $R_0 > 7.5$ kpc [48]. We therefore suppose that R_0 is close to 8 kpc, and will adopt the value 8.0 kpc for convenience in our subsequent computations.

5. DETERMINATION OF Θ_0

The existence of a nearly flat rotation curve for the inner Galaxy was first suggested in [48], based on data from an HI survey. Later, an essentially flat rotation curve was obtained by Brand and Blitz [35] for the entire range of R studied, using data on optical nebulae with known photometric distances. A flat curve was obtained by Russeil [50], who adopted $R_0 = 7.1$ kpc and $\Theta_0 = 184$ km/s.

Let us suppose, as is indicated by the results of [35] and [50], that the rotation curve is flat, and derive Θ_0 using the distances and radial velocities of the SFRs only, applying the function $W(R) = (V/\sin l)\cos b$, which is valid for the outermost HII regions, located beyond the Perseus arm, and adopting $R_0 = 8$ kpc. We apply these restrictions, first, in order to exclude the Perseus arm, which clearly has abnormal kinematics and, second, because the $W(R)$ curves for different Θ_0 values diverge more with distance from the Sun. We will set the lower limit for the distances to be used to $R/R_0 = 1.35$. There are 32 HII regions with known distances in this region. The function $W(R)$ for a flat rotation curve is $W(R) = (V/\sin l)\cos b = (R_0/R)\Theta_R - \Theta_0$, and depends only on Θ_0 and R_0/R. Figure 3 shows the dependence of $W(R)$ on R_0/R for the HII regions. We approximated this dependence as $y = \Theta_0(R_0/x - 1)$ and applied the IDL 5.4 approximation routine LMFIT to find Θ_0. The result is $\Theta_0 = 202 \pm 4$ km/s for $R_0 = 8$ kpc. This result remains essentially unchanged even if we remove the point that is most distant from the center. For comparison, adopting $R_0 = 7.5$ kpc yields $\Theta_0 = 193 \pm 4$ km/s. Thus, assuming a flat rotation curve, we obtain for the Galactic constants $R_0 = 8$ kpc and $\Theta_0 = 200 \pm 4$ km/s. This means that the Oort relation is $A - B = \Theta_0/R_0 = 25$ km/s kpc.

6. THE VALUES OF Θ_0 AND $A - B$

Summaries of previous determinations of the Oort constants $A - B = \Omega_0$ and Θ_0 are given in Tables 5 and 6. The angular and circular velocities in these tables display a large scatter, due to several combined effects. It is known that the velocity dispersions for particular types of objects increases with the age of their population. The components of the solar proper

motion also depend on the type of objects used to determine them. The azimuthal component of the solar motion is most uncertain. Detailed studies of the solar motion based on the kinematics of main-sequence stars in the HIPPARCHOS catalog [61] have shown that the radial (U_0) and vertical (W_0) components depend only weakly on the $B-V$ color index of the stellar sample, while the azimuthal component (V_0) increases with $B-V$, and is also proportional to the assumed velocity dispersion. The value of V_0 reduced to zero dispersion is 5.25 km/s; this is rather different from the values that are usually obtained. The Oort constants in the solar vicinity are also strongly influenced by the closest spiral arms and resonances due to the central bar. As was shown in [20, 24], the outer Lindblad resonance passes quite near the solar orbit, about 1 kpc closer to the Galactic center. Olling and Dehnen [59] described the kinematics of the Galaxy using new data on the solar proper motion, dividing their sample of $B-V$ values into separate groups and taking into account the influence of the spiral structure and the resonances produced by the central bar on the Oort constants. Using a sample of 10^6 stars from the ACT and TYCHO-2 catalogs, they obtained very different angular velocities for the two extreme populations of the Galaxy: $A-B=21.1$ km/s kpc for young main-sequence stars and $A-B=32.8$ km/s kpc for old red stars.

Thus, the angular velocity derived from the Oort constants depends on the velocity dispersion of the sample of stars used, the size of the velocity ellipsoid, the inhomogeneity of the sample in space and in terms of the constituent objects, and the influence of spiral arms and resonances. If we exclude data based on old red stars, extreme values [59], and the data of [52] (due to the inhomogeneity of the sample presented in [57]) from Table 5, we obtain the mean value $A-B=\Theta_0/R_0=27.2\pm2.3$ km/s kpc.

The values we have derived fit into this range well.

7. GALACTIC ROTATION CURVE FROM MOLECULAR GAS

Fich *et al.* [3] used HI velocities at tangent points for the inner Galaxy and CO velocities and photometric distances for the outer Galaxy to construct the rotation curve. They studied eight different approximating functions for the rotation curve, and identified the two that yielded the best descriptions: $y=a_1/x+a_2$ and $y=b_1x^{b_2}+b_3/x$. Here, $y=\Theta/\Theta_0$ or $y=\Omega/\Omega_0$ and $x=R/R_0$. As a result, assuming $R_0=8.5$ kpc and $\Theta_0=220$ km/s, they derived $a_1=1.00746$, $a_2=-0.017112$, and $b_1=0.49627$, $b_2=-0.00421$, and $b_3=0.49632$. Both curves are virtually identical and

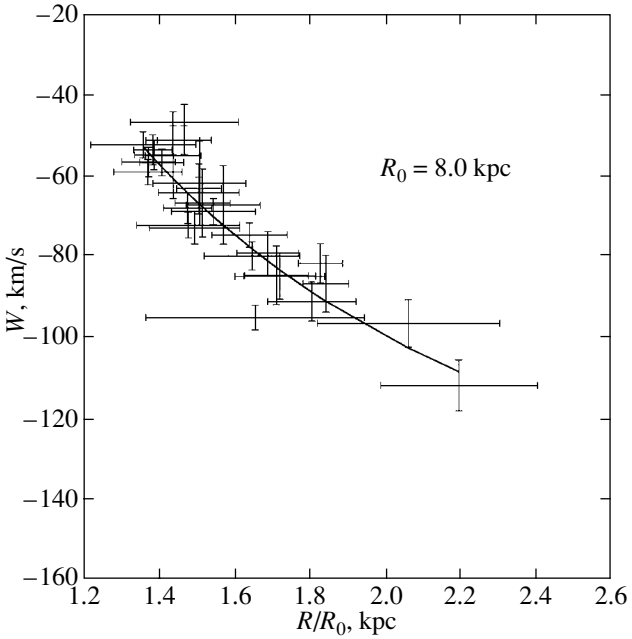


Fig. 3. Observed dependence of $W(R)$ on R/R_0 for young disk objects. The solid curve corresponds to a flat rotation curve with $\Theta_0=200$ km/s.

give almost flat rotation curves. Using 400 HII regions and reflection nebulae together with HI line data for the inner Galaxy, Brand *et al.* [35] obtained for the same R_0 and Θ_0 the values $b_1=1.00767$, $b_2=0.00394$, and $b_3=0.00712$. Assuming $R_0=7.1$ kpc and $\Theta_0=184$ km/s, Russeil [50] obtained $b_1=0.705$ and $b_2=0.35\cdot10^{-8}$ using her own data on optical HII regions with known photometric distances and Hα velocities.

Due to the large uncertainty in stellar distances, we consider the rotation curve based on the CO tangent velocities in the first and fourth quadrants for the inner Galaxy and on HII regions for the outer Galaxy to be more trustworthy. This curve is more accurate than the curve based on HII regions alone. Therefore, to construct our rotation curve, we used the CO surveys [9] in the first quadrant and [62] in the fourth quadrant for the inner Galaxy, and optical nebulae with known distances to their exciting stars and radial velocities for the associated molecular clouds or for the recombination lines of hydrogen of HII regions for the outer Galaxy (Table 2). These data yielded the radial angular-velocity distribution for $R_0=8.0$ kpc and $\Theta_0=200$ km/s.

Figure 4 shows the distribution of the observational data describing the variation of the Galactic angular velocity with radius $\Omega(R)$, derived from CO data for the first and fourth quadrants (horizontal and vertical dashes) for the inner Galaxy, and from data for HII regions alone for the outer Galaxy. We

Table 5. Summary of estimates of $A - B \equiv \Omega_0$

$A - B$, km s^{-1}kpc^{-1}	Assumed R_0, kpc	Refs.	Objects
25.3 ± 1.3	7.9	[12]	Hydrogen HI
25.9 ± 1.1	7.1	[51]	Gas distribution
30.9 ± 1.6	7.66	[40]	Cepheids
31.6 ± 1.4	8.5	[52]	1352 O–B5 stars
27.19 ± 0.87	8.5	[53]	220 Cepheids
25.2 ± 1.9	7.1	[41]	Gas distribution
28.7 ± 1.0	7.5	[54]	Cepheids and open clusters
28.0 ± 2.0	8.0	[55]	Sgr A*
28.0 ± 2.0	8.0	[56]	Sgr A*
30.1 ± 1.0	8.5	[57]	240 O–B5 stars
29.1 ± 1.0	7.1	[7]	OB associations
27.5 ± 1.4	7.5	[4]	Cepheids and open clusters
26.6 ± 1.4	8.5	[4]	Cepheids and open clusters
27.6 ± 1.7		[58]	Globular cluster M4
21.1 ± 1.2		[59]	Young main-sequence stars
32.8 ± 1.2		[59]	Old red giants

Table 6. Summary of determinations of Θ_0

Θ_0, km/s	Assumed R_0, kpc	Refs.	Objects
200 ± 10	7.9	[12]	Hydrogen HI
184 ± 8	7.1	[51]	Gas distribution
237 ± 12	7.66	[40]	Cepheids
268.7 ± 11.9	8.5	[52]	1352 O–B5 stars
243.3 ± 12.0	8.5	[52]	170 Cepheids
219 ± 20	8.0	[55]	Sgr A*
270	8.0	[60]	HIPPARCOS stars
255.5 ± 8.33	8.5	[57]	240 O–B5 stars

also show the corresponding error bars in Fig. 4. A velocity jump due to the central bar is visible at 3.8–4.0 kpc in the first quadrant. This jump is larger in the first quadrant than in the fourth quadrant, since the major axis of the bar is located in the first quadrant. We approximated this dependence using the functions indicated above with the IDL 5.4 routine LMFIT. For the total sample, after excluding points inside the bar corotation region ($R < 4.0$ kpc), we obtained the rotation-curve parameters $b_1 = 1.015$, $b_2 = -0.00397$, and $b_3 = 0.00864$. This yields for the radial dependence of the angular velocity

$$\Omega/\Omega_0 = 1.015(R_0/R)^{0.00397} + 0.00864(R_0/R).$$

This relation is plotted by the solid curve in Fig. 4a. The curve corresponding to the equation

$$\Omega = \Theta_0/R_1^{b_1}(R_0)^{a_1},$$

with $a_1 = 0.07$ and $b_1 = 1 - a_1$ is shown by the dots in the same figure.

Both curves are virtually the same outside the solar circle, while, inside the solar circle, the dashed curve lies slightly below the solid curve. In the region

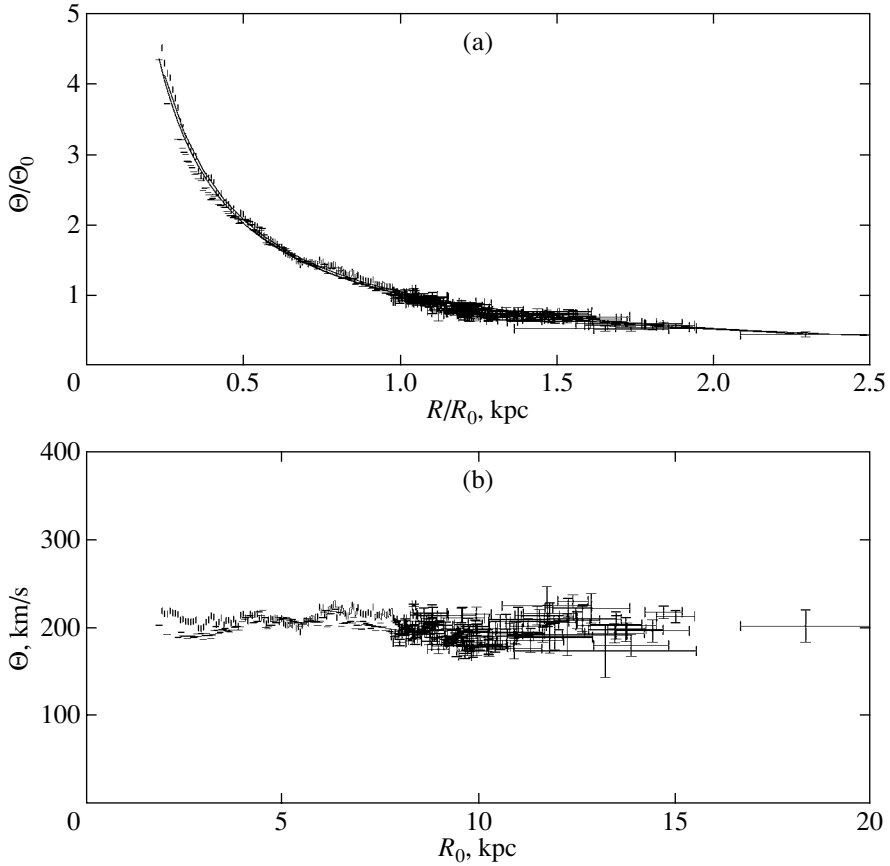


Fig. 4. (a) Radial distribution of the reduced angular velocities based on CO data for the first quadrant (horizontal dashes) and fourth quadrant (vertical dashes). The points for the outer Galaxy are based on the data from Table 2 (the corresponding errors are also shown). The smooth curve that best approximates these data has the parameters $b_1 = 1.015$, $b_2 = -0.00397$, and $b_3 = 0.00864$ (see the text). (b) Rotation curve based on the data indicated above.

of the bar, the solid curve passes exactly between the points representing near and distant parts of the Galactic bar.

Figure 4b shows the distribution of the observational data for the Galactic rotation curve. If we exclude the Perseus-arm region, the distribution is flat, to first approximation. Nevertheless, as we noted above, the rotation curves for the Northern sky (first and second quadrants) and the Southern sky (third and fourth quadrants) differ.

When deriving kinematic distances from observed velocities, it is important to use a rotation curve that includes the direction toward the objects under study. Accordingly, Figs. 5a and 5b show rotation curves for the Northern and Southern sky separately. For more convenient application of these to the distance determinations, we approximated both distributions using nonlinear equations with the IDL routines LMFIT and POLYFITW. We obtained for the Northern rotation curve the analytical expression

$$\Theta = a_0 + a_1 R + a_2 R^2 + a_3 R^3 + a_4 R^4 + a_5 R^5,$$

with $a_0 = -503.424$, $a_1 = 433.336$, $a_2 = -97.028$, $a_3 = 9.97301$, $a_4 = -0.477901$, and $a_5 = 0.00865043$.

The analytic expression for the Southern rotation curve has the form

$$\Theta = b_0 + b_1 R + b_2 R^2 + b_3 R^3 + b_4 R^4 + b_5 R^5,$$

where $b_0 = -364.553$, $b_1 = 329.293$, $b_2 = -70.3912$, $b_3 = 7.05383$, $b_4 = -0.3364$, and $b_5 = 0.00617765$.

When determining kinematic distances in the direction of the Perseus arm, one must bear in mind that, as is shown in [63], the observed radial velocities do not correspond to the distances in this case, since the gravitational well in this arm is much deeper than the variations of the gravitational field in the arms within the solar circle.

8. CONCLUSION

We have used data on 270 SFRs with photometric distances and radial velocities derived from

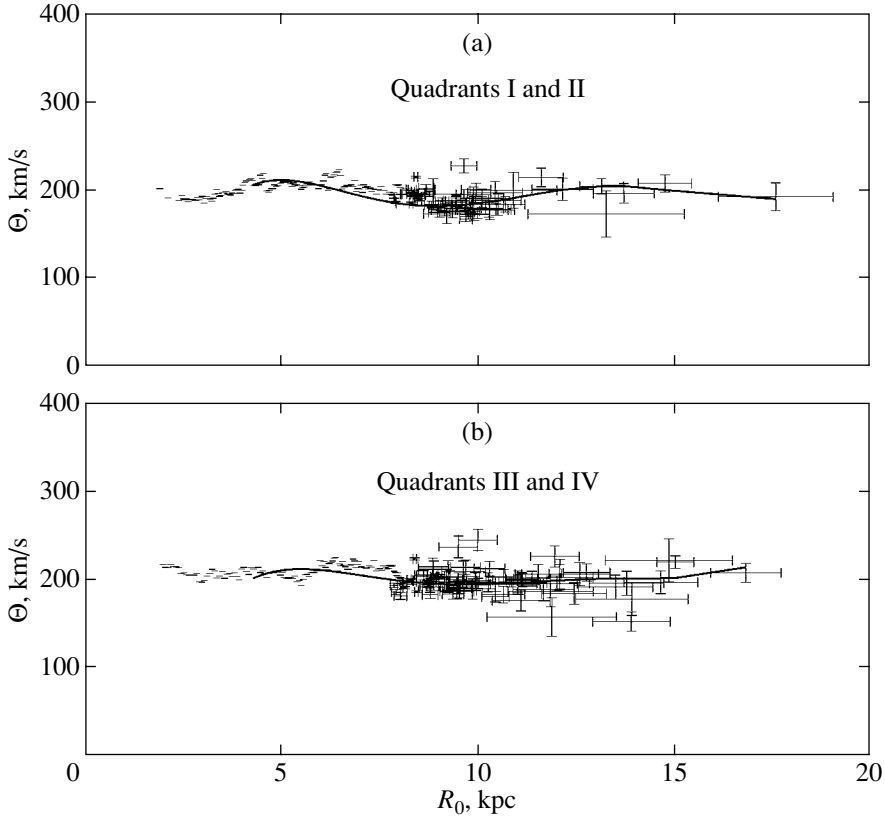


Fig. 5. Galactic rotation curve for the (a) Northern and (b) Southern sky.

their molecular gas to estimate the Galactic constants R_0 and Θ_0 and construct the Galactic rotation curve. Despite the presence of a barlike structure in the Galactic center, it is possible, to first approximation, to assume axially symmetric motion for the gas component. Estimating the distance to the Galactic center from ten tangent points under this assumption yielded the mean value $R_0 = 8.01 \pm 0.44$ kpc. Together with other determinations over the last decade, we obtain the weighted mean value $R_0 = 7.80 \pm 0.32$ kpc. Fich *et al.* [3], Brand *et al.* [35], and Russeil [50] derived almost flat rotation curves. Assuming a flat rotation curve, we obtained $\Theta_0 = 200$ km/s for the solar circular velocity using the SFRs beyond the Perseus arm. We then used the values $R_0 = 8.0$ and $\Theta_0 = 200$ km/s to derive the Galactic rotation curve. To first approximation, we can assume that, starting from the location of the central bar (i.e., from ∼4 kpc) and out to 15–16 kpc, the curve is virtually flat (∼200 km/s) and can be described by the expression $\Theta = 218.13 - 1.827R$. We obtained different analytical expressions for the rotation curves for the Northern and Southern hemispheres of the Galaxy. The velocity jump in the first quadrant in the corotation region of the central bar is 20 km/s.

ACKNOWLEDGMENTS

I am grateful to N.N. Chugai and to A.E. Piskunov for reading the manuscript and for remarks.

REFERENCES

1. F. Eisenhauer, R. Schodel, R. Genzel, *et al.*, Astrophys. J. Lett. **597**, L121 (2003).
2. J. Hron, Astron. Astrophys. **176**, 34 (1986).
3. M. Fich, L. Blitz, and A. A. Stark, Astrophys. J. **342**, 272 (1989).
4. M. V. Zabolotskii, A. S. Rastorguev, and A. K. Dambis, Pis'ma Astron. Zh. **28**, 516 (2002) [Astron. Lett. **28**, 454 (2002)].
5. F. Pont, M. Mayor, and G. Burki, Astron. Astrophys. **285**, 415 (1994).
6. F. Pont, D. Queloz, P. Bratschi, and M. Mayor, Astron. Astrophys. **318**, 416 (1997).
7. A. K. Dambis, A. M. Mel'nik, and A. S. Rastorguev, Pis'ma Astron. Zh. **21**, 331 (1995) [Astron. Lett. **21**, 291 (1995)].
8. L. H. Amaral, R. Ortiz, J. R. D. Lepine, and W. J. Maciel, Mon. Not. R. Astron. Soc. **281**, 339 (1996).
9. D. P. Clemens, Astrophys. J. **295**, 422 (1985).
10. I. I. Nikiforov and I. V. Petrovskaya, Astron. Zh. **71**, 725 (1994) [Astron. Rep. **38**, 642 (1994)].
11. H. Alvarez, J. May, and L. Bronfman, Astrophys. J. **348**, 495 (1990).

12. M. R. Merrifield, Astron. J. **103**, 1552 (1992).
13. W. W. Shane, Astron. Astrophys. **16**, 118 (1971).
14. W. L. Peters, Astrophys. J. **195**, 617 (1975).
15. L. Blitz and D. N. Spergel, Astrophys. J. **379**, 631 (1991).
16. S. Nikolaev and M. D. Weinberg, Astrophys. J. **487**, 885 (1997).
17. K. Z. Stanek, A. Udalski, *et al.*, Astrophys. J. **477**, 163 (1997).
18. B. Paczynski, K. Z. Stanek, *et al.*, Astrophys. J. Lett. **435**, L113 (1994).
19. J. Binney, O. Gerhard, and D. Spergel, Mon. Not. R. Astron. Soc. **288**, 365 (1997).
20. D. Raboud, M. Grenon, L. Martinet, *et al.*, Astron. Astrophys. **335**, L61 (1998).
21. B. J. Weiner and J. A. Sellwood, Astrophys. J. **524**, 112 (1999).
22. R. Fux, Astron. Astrophys. **345**, 787 (1999).
23. M. R. Merrifield, Astron. Soc. Pac. Conf. Ser. **317**, 289 (2004).
24. W. Dehnen, Astrophys. J. Lett. **524**, L35 (1999).
25. P. Englmaier and O. Gerhard, Mon. Not. R. Astron. Soc. **304**, 512 (1999).
26. V. S. Avedisova, Astron. Zh. **79**, 216 (2002) [Astron. Rep. **46**, 193 (2002)].
27. N. R. Walborn, I. D. Howarth, D. J. Lennon, *et al.*, Astron. J. **123**, 2754 (2002).
28. N. R. Walborn, Astron. J. **124**, 507 (2002).
29. W. D. Vacca, C. D. Garmany, and J. M. Shull, Astrophys. J. **460**, 914 (1996).
30. P. Patriarchi, L. Morbidelli, and M. Perinotto, Astron. Astrophys. **410**, 905 (2003).
31. J. Maiz-Apellaniz and N. R. Walborn, Astrophys. J., Suppl. Ser. **151**, 103 (2004).
32. T. M. Dame, H. Ungerechts, R. S. Cohen, *et al.*, Astrophys. J. **322**, 706 (1987).
33. J. C. Mermilliod, Astron. Astrophys., Suppl. Ser. **24**, 159 (1986); CDS, II/124A.
34. Y. M. Georgelin, D. Russeil, *et al.*, Astron. Astrophys. **357**, 308 (2000).
35. J. Brand and L. Blitz, Astron. Astrophys. **275**, 67 (1993).
36. Y. M. Georgelin, PhD Thesis (Univ. of Marseille, Marseille, 1975).
37. K. Kuijken and S. Tremaine, Astrophys. J. **421**, 178 (1994).
38. W. J. Maciel, Astrophys. Space Sci. **206**, 285 (1993).
39. B. W. Carney, J. P. Fulbright, *et al.*, Astron. J. **110**, 1674 (1995).
40. M. R. Metzger, A. R. Caldwell, and P. L. Schechter, Astron. J. **115**, 635 (1998).
41. R. P. Olling and M. R. Merrifield, Mon. Not. R. Astron. Soc. **297**, 943 (1998).
42. B. Paczynski and K. Z. Stanek, Astrophys. J. Lett. **494**, L219 (1998).
43. K. Z. Stanek and P. M. Garnavich, Astrophys. J. Lett. **503**, L131 (1998).
44. E. V. Glushkova, A. K. Dambis, A. M. Mel'nik, and A. S. Rastorguev, Astron. Astrophys. **329**, 514 (1998).
45. D. H. McNamara, J. B. Madsen, J. Barnes, and B. F. Ericksen, Publ. Astron. Soc. Pac. **112**, 202 (2000).
46. R. Genzel, C. Pichon, A. Eckart, *et al.*, Mon. Not. R. Astron. Soc. **317**, 348 (2000).
47. T. P. Gerasimenko, Astron. Zh. **81**, 102 (2004) [Astron. Rep. **48**, 103 (2004)].
48. H. T. Freudenreich, Astrophys. J. **492**, 495 (1998).
49. J. E. Gunn, G. R. Knapp, and S. D. Tremaine, Astron. J. **84**, 1181 (1979).
50. D. Russeil, Astron. Astrophys. **397**, 133 (2003).
51. R. P. Olling and M. R. Merrifield, Mon. Not. R. Astron. Soc. **297**, 943 (1993).
52. M. Miyamoto and Zi Zhu, Astron. J. **115**, 1483 (1998).
53. M. Feast and P. Whiteoak, Mon. Not. R. Astron. Soc. **291**, 683 (1997).
54. A. S. Rastorguev, E. V. Glushkova, A. K. Dambis, and M. V. Zabolotskikh, Pis'ma Astron. Zh. **25**, 689 (1999) [Astron. Lett. **25**, 595 (1999)].
55. M. J. Reid, A. C. S. Readhead, R. C. Vermeulen, and R. N. Treuhaft, Astrophys. J. **524**, 816 (1999).
56. D. C. Backer and R. A. Sramek, Astrophys. J. **524**, 805 (1999).
57. M. Uemura, H. Ohashi, T. Hayakawa, *et al.*, Publ. Astron. Soc. Jpn. **52**, 143 (2000).
58. L. R. Bedin, G. Piotto, I. R. King, and J. Anderson, Astron. J. **126**, 247 (2003).
59. R. P. Olling and W. Dehnen, Astrophys. J. **599**, 275 (2003).
60. R. A. Mendez, I. Platais, and T. M. Girard, Astrophys. J. Lett. **524**, L39 (1999).
61. W. Dehnen and J. J. Binney, Mon. Not. R. Astron. Soc. **298**, 387 (1998).
62. B. J. Robinson, R. N. Manchester, and J. B. Whiteoak, Astron. Astrophys. **193**, 60 (1988).
63. T. G. Sitnik, Pis'ma Astron. Zh. **29**, 356 (2003) [Astron. Lett. **29**, 311 (2003)].

Translated by L. Yungel'son

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 446–449. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 500–504.

Gamma-Ray Bursts from a Close Pulsar Binary System?

Ya. N. Istomin

Lebedev Physical Institute, Leninskiĭ pr. 53, Moscow, Russia

Received July 22, 2004; in final form, December 3, 2004

Abstract—The close neutron-star binary system comprised of the radio pulsars PSR J0737−3039 A,B is discussed. An analysis of the observational data indicates that the wind from pulsar A, which is more powerful than the wind from pulsar B, strongly distorts the magnetosphere of pulsar B. A shock separating the relativistic wind from pulsar A and the corotating magnetosphere of pulsar B should form inside the light cylinder of pulsar B. A weakly diverging "tail" of magnetic field is also formed, which stores a magnetic energy on the order of 10^{30} erg. This energy could be liberated over a short time on the order of 0.1 s as a result of reconnection of the magnetic-force lines in this "tail," leading to an outburst of electromagnetic radiation with energies near 100 keV, with an observed flux at the Earth of 4×10^{-11} erg cm^{-2} s^{-2}. Such outbursts would occur sporadically, as in the case of magnetic substorms in the Earth's magnetosphere.

1. INTRODUCTION

Lyne *et al.* [1] report the discovery of a unique binary system: two neutron stars, both observed as radio pulsars (PSR J0737−3039 A and B) and with the system having a small orbital period of 0.1 d. Component A is a millisecond pulsar with period $P_A = 22.7$ ms, while component B is an ordinary pulsar with period $P_B = 2.773$ s. The uniqueness of this system is that, in addition to being a natural laboratory for the measurement of general-relativistic effects, it also provides the opportunity to study the structure of the pulsars' magnetospheres. The reason is that the plane of the orbital motion is inclined only 3° to the line of sight, so that we can observe the propagation of the radio waves from one pulsar through the magnetosphere of the other. Periodic eclipses of both pulsars are observed in the radio [1, 2]. It is important that the eclipses of the more powerful millisecond pulsar A are brief (about 27 s), which corresponds to an eclipsing region near pulsar B of about 1.9×10^9 cm, appreciably smaller than B's magnetosphere, $R_{LB} = cP_B/2\pi = 1.3 \times 10^{10}$ cm. The quantity R_L (in this case, R_{LB}) is the radius of the light surface where the corotation speed of the plasma is comparable to the speed of light c. This is due to the fact that the neutron-star binary system is so close that the plasma wind coming from the millisecond pulsar A penetrates and strongly distorts the magnetosphere of pulsar B. The total rate of rotational energy loss by pulsar A is

$$\dot{E}_A = 5.8 \times 10^{33} \left(\frac{J}{10^{45}\ \mathrm{g\ cm^2}} \right)\ \mathrm{erg/s},$$

since $\dot{E} = 4\pi^2 J(\dot{P}/P^3)$, $\dot{P}_A = 1.7 \times 10^{-18}$, and $J \simeq 10^{45}$ g cm^2 is the standard value for the moment of inertia of a neutron star. The distance between the stars is $d = 8.5 \times 10^{10}$ cm. With this separation, the flux of stellar-wind energy from pulsar A in the vicinity of pulsar B is $F_A = \dot{E}_A/4\pi d^2 = 6.4 \times 10^{10}$ erg cm^{-2} s^{-1}. Assuming the wind is relativistic, we can find the energy density in the stellar wind from pulsar A acting on the magnetosphere of pulsar B: $W_A = F_A/c = 2.1$ erg/cm^3. The magnetic field for which the stellar-wind energy would be comparable to the magnetic-field energy ($B^2/8\pi = W_A$) is $B_a = 7.3$ G. This value corresponds to the magnetic field inside the magnetosphere of pulsar B. Even if this is the value at the magnetosphere boundary, $r = R_{LB}$, the magnetic field at the surface of the neutron star should be 1.6×10^{13} G, in strong contradiction to its rotational energy losses.

The flux of charged particles in A's wind compresses the magnetic field in the magnetosphere of pulsar B where the wind encounters and flows past B, stretching out the field on the opposite side of B (forming a so-called magnetospheric "tail"). Thus, the structure of B's magnetosphere is strongly distorted by the wind from A, and resembles the shape acquired by the magnetosphere of the Earth due to the action of the solar wind flowing past its dipolar magnetic field. In this case, the rotational energy losses of pulsar B should differ markedly from those for a "classical" ordinary radio pulsar. The figure shows a schematic of the interaction between radio pulsars A and B. It is usually assumed that a rotating dipolar magnetic field leads to the radiation of

1063-7729/05/4906-0446$26.00

electromagnetic radiation at a frequency equal to the rate of rotation of the neutron star, $\omega = \Omega = 2\pi/P$. This is called magnetic-dipole radiation, and is the mechanism via which a neutron star loses rotational energy. To order of magnitude, these magnetic-dipole losses are equal to

$$\dot{E}_{MD} \simeq B^2\Omega^4 R^6 c^{-3},$$

where B is the dipolar magnetic field at the surface of the star and R is the star's radius. These losses are also proportional to the square of the sine of the inclination of the dipole axis to the rotational axis; since we do not know this quantity for pulsar B, we will take it to be of order unity. On this basis, we estimate the magnetic field at the surface of the radio pulsar using the relation $\dot{E}_{MD} = J\Omega\dot{\Omega}$: $B \simeq (P\dot{P}_{-15})^{1/2} \times 10^{12}$ G, where $\dot{P}_{-15}$ is the deceleration of the rotation in units of 10^{-15} s/s. We find for pulsar B $\dot{P}_B = 0.88 \times 10^{-15}$ and the corresponding magnetic field $B_B \simeq 1.6 \times 10^{12}$ G [1].

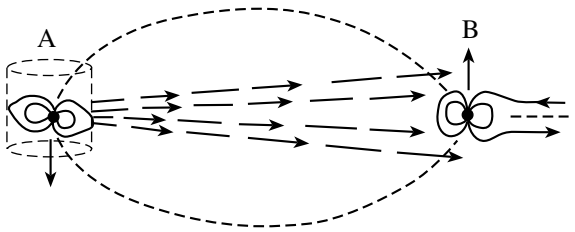
Magnetic-dipole radiation arises during the rotation of a dipolar field in a vacuum, but in reality, the pulsar magnetosphere is filled with relativistic plasma. In this case, the magnetic-dipole radiation is screened [3], and the rotational energy losses are associated with electric currents flowing in the magnetosphere in the vicinity of open field lines, which close at the surface of the neutron star in a region with size $r_0 \simeq R(R\Omega/c)^{1/2}$ near the poles. The rotational energy losses are associated with radiation by the relativistic wind generated in the magnetosphere. To order of magnitude, the radiated energy is

$$\dot{E}_c \simeq B^2\Omega^4 R^6 c^{-3} i,$$

where $i = j/j_{GJ}$ is the dimensionless electrical current flowing in the magnetosphere in units of the so-called Goldreich–Julian current $j_{GJ} = B\Omega/2\pi$. The current flowing along the neutron-star surface at the polar cap creates a braking torque $K \simeq \pi j B r_0^4/c$, which leads to the energy losses $\dot{E}_c = K\Omega$. We can see that, when the current is $j \simeq j_{GJ}$, the magnetic-dipole energy losses and energy losses associated with currents are comparable.

2. THE MAGNETOSPHERE OF PULSAR B

The distorted structure of the magnetic field in the magnetosphere of pulsar B should strongly influence its energy losses. We denote the distance from the center of pulsar B where a shock dividing the wind from pulsar A and the magnetosphere of B forms r_a (the Alfven radius). Within $r < r_a$, the magnetic field of B is close to dipolar and the magnetosphere corotates with the pulsar at the angular speed Ω_B.



Schematic of the interaction between pulsars A and B. The wind from the stronger pulsar A penetrates the light surface of pulsar B, leading to the formation of an elongated tail in the magnetosphere of B.

Thus, $B_B(R/r_a)^3 = B_a = (8\pi W_A)^{1/2} = 7.3$ G. The projection of the region with $r \simeq r_a$ onto the stellar surface along the magnetic field in the vicinity of the magnetic poles is $r_0 \simeq R(R/r_a)^{1/2}$. We can see that the size of the polar cap is appreciably larger than in the case of an ordinary pulsar ($r_a \ll c/\Omega_B = R_{LB}$) due to the flow of the companion's wind past the pulsar. Furthermore, the electric current j flowing in the polar region is also increased. The maintenance of a magnetic field in the magnetosphere tail $B \simeq B_a$ requires a current $j = B_a c/2\pi r_a \simeq j_{GJ}(c/\Omega_B r_a) \gg j_{GJ}$. The electrical current j closing in the polar region at the stellar surface creates a braking torque that leads to the rotational energy losses

$$\dot{E}_B = B_B^2 \Omega_B r_a^3 \left(R/r_a\right)^6,$$

which exceed the magnetic-dipole losses by a factor of $(c/\Omega_B r_a)^3$. This increase in the loss rate is associated with the more efficient action of the rotating magnetic field of the star as a unipolar inductor [4]. The potential difference arising at the contacts of the inductor is $U = \Omega\Delta f/c$, where Δf is the difference of the magnetic fluxes at the two contacts. For an ordinary pulsar, $\Delta f = B r_0^2 \simeq BR^2(R\Omega/c)$, while, in our case, the magnetosphere tail closes a region $r \simeq r_a$ with the light surface $r = R_L$, $\Delta f = BR^2(R/r_a) \gg BR^2(R\Omega/c)$. The magnitude of the electric current flowing in the magnetosphere, $I = \pi j r_a^2$, likewise increases. This leads to an increase in the rotational losses of the star:

$$\dot{E}_B = UI = B_B^2 R^6 \Omega_B r_a^{-3}.$$

The electrical potential U acts on the charged particles in the magnetosphere, giving rise to a relativistic wind from pulsar B.

Using the two relations

$$\dot{E}_B = B_B^2 R^6 \Omega_B r_a^{-3}, \quad 2\dot{E}_A = cB_B^2 \left(\frac{R}{r_a}\right)^6 d^2,$$

we can independently determine the values of r_a and B_B:

$$r_a = \left(\frac{cd^2}{2\Omega_B}\frac{\dot{E}_B}{\dot{E}_A}\right)^{1/3},$$

$$B_B = \dot{E}_B\left(\frac{cd}{\Omega_B R^3}\right)\left(\frac{1}{2c\dot{E}_A}\right)^{1/2}.$$

Substituting the observed values for d, Ω_B, $\dot{E}_A$, and $\dot{E}_B = 1.6\times10^{30}$ erg/s, we obtain

$$r_a = 2.4\times10^9 \text{ cm}, \quad B_B = 10^{11}\left(\frac{R}{10^6 \text{ cm}}\right)^{-3} \text{ G}. \tag{1}$$

Note that this value for r_a—the size of pulsar B's magnetosphere—coincides with the size of the region that eclipses the radio emission from pulsar A, 1.9×10^9 cm. This indicates that there is, indeed, an interface separating the wind from pulsar A and the magnetosphere of pulsar B. This interface lies in the region where the pressures of the wind and of the magnetosphere magnetic field are equal, and has the form of a shock and contact discontinuity at the front and lateral surfaces.

This raises the question of the nature of the radio eclipse at $r \simeq r_a$. Various opinions about this have been expressed in the literature.

1. Kaspi *et al.* [2] and Lyutikov [5] have proposed that the eclipses are due to cyclotron absorption of the radio waves from pulsar A by relativistic particles in the wind from pulsar B near the shock interface. However, the magnetic field near the magnetosphere boundary of pulsar B, $B \simeq B_a = 7.3$ G, implies a modest electron-cyclotron frequency in this region, $\nu_{ca} \simeq 10^7$ s^{-1}, which is appreciably lower than the radio frequencies at which the eclipses are observed, $\nu_{ca} \ll \nu \simeq 10^9$ s^{-1}. This hinders the realization of an electron-cyclotron resonance that is capable of strongly absorbing the radio emission of pulsar A.

2. Another mechanism that can explain the eclipses is the linear transformation of electromagnetic waves frozen in the pulsar wind into longitudinal waves in the strong gradient of the electron density in the vicinity of the shock [6]. The transformation coefficient has an exponential dependence on the frequency of the waves, $\exp\{-\nu/\nu_{cr}\}$, which can explain the absence of a frequency dependence for the eclipse characteristics at frequencies $\nu < \nu_{cr}$, as is observed for J0737−3039, if $\nu_{cr} > 10^9$ s^{-1}.

Our estimate of the magnetic field at the surface of pulsar B ($B_B = 10^{11}$ G) is an order of magnitude lower than estimates based on magnetic-dipole losses (1.6×10^{12} G [1]). Essentially, a neutron star with $B_B = 10^{11}$ G and $P_B = 2.77$ s should not be a radio pulsar, and pulsar B would not be a radio pulsar without a second neutron star as a close companion. Pulsar B lies below the "death line" for ordinary neutron stars on the $B-P$ diagram. The relativistic wind from pulsar A flowing past its magnetosphere creates the conditions required for the generation of an electron–positron plasma in the polar regions, thereby also leading to the generation of radio emission. It follows that the radio emission should arise in inner regions of the pulsar magnetosphere ($r < r_a$), where the magnetosphere corotates with the star. This inner region of the magnetosphere with size r_a is at least an order of magnitude smaller than the light surface, $R_L = 1.3\times10^{10}$ cm.

It is important that the size of the surface from which the relativistic wind flows from pulsar B can be estimated by equating the energy fluxes from the two pulsars:

$$S = 4\pi d^2\left(\dot{E}_B/\dot{E}_A\right).$$

This area corresponds to a radius of $r = 2d(\dot{E}_B/\dot{E}_A)^{1/2} = 2.8\times10^9$ cm. The closeness of this radius to $r_a = 2.4\times10^9$ cm indicates that all the wind from pulsar B flows out through its magnetosphere tail. The divergence angle of the tail 2θ is determined by the divergence angle of the wind from pulsar A at a distance r_a from the center of pulsar B: $2\theta = 2r_a/d = 5.6\times10^{-2}$ rad $= 3.2°$. This is close to the angle between the line of sight and the orbital plane for the binary system. This indicates that we can observe the magnetospheric tail, and therefore observe radio eclipses of pulsar A. The tail will extend to large distances l, until its magnetic field $B = B_a(1+l/d)^{-2}$ becomes comparable to the interstellar field. It is clear, however, that the effective size of the tail is equal to the distance between the pulsars, $d = 8.5\times10^{10}$ cm. Indeed, the energy stored in the tail's magnetic field is

$$\epsilon_m = \int \frac{B^2}{8\pi}\pi r_a^2\left(1+\frac{l}{d}\right)^{-2} dl = \frac{B_a^2}{8\pi}\pi r_a^2 d.$$

In our case, the stored magnetic energy is $\epsilon_m = 3.3\times10^{30}$ erg.

3. FLARES

The magnetic field in the tail is directed in different directions on opposite ends of the tail—a configuration that is unstable to magnetic reconnection. The characteristic time for reconnection is given by the ratio of the transverse size of the tail r_a to the Alfven speed v_a. Inside the magnetosphere, the energy density of the particles is lower than the energy density

of the magnetic field, so that $v_a \simeq c$. Thus, the decay time for current sheets in the magnetosphere tail is roughly $\tau \simeq 0.1$ s. The magnetic energy stored in the tail should be transformed into the energy of accelerated particles on this time scale, with the subsequent radiation of electromagnetic radiation. Electrons and positrons in the magnetosphere will be accelerated by the electric field arising due to the annihilation of the magnetic field, $E \simeq B(r_a/c\tau) = 7.3$ cgs $= 2.2 \times 10^3$ V/cm. The maximum energy that can be acquired by these electrons and positrons is $\mathcal{E}_{\max} = Er_a = 5 \times 10^{11}$ eV. On the other hand, the particles cannot achieve very high energies due to synchrotron losses. The energy balance for the particles is determined by the equation

$$\frac{d\mathcal{E}}{dt} = ecE - \frac{2}{3}\frac{e^4B^2}{m_e^2c^3}\gamma^2,$$

where $\gamma = \mathcal{E}/m_ec^2$ is the Lorentz factor of the particles. In our case, when $E \simeq B$, the characteristic limiting Lorentz factor determined by the synchrotron losses is

$$\gamma_c = (3c/2r_e\omega_c)^{1/2} \simeq 3.5 \times 10^7.$$

Here, r_e is the classical radius of the electron and ω_c is the electron-cyclotron frequency, $\omega_c = eB/m_ec \simeq 1.3 \times 10^8$ s^{-1}. Thus, we can see that during the reconnection, particles can be accelerated to maximum Lorentz factors of the order of

$$\gamma_{\max} = \mathcal{E}_{\max}/m_ec^2 \simeq 10^6.$$

The characteristic frequency for synchrotron radiation by accelerated particles with $\gamma \simeq \gamma_{\max}$ is $\omega \simeq \omega_c\gamma_{\max}^2 \simeq 1.3 \times 10^{20}$ s^{-1}. The energy of the corresponding synchrotron photons is $\mathcal{E}_{ph} = \hbar\omega \simeq 10^{-7}$ erg $\simeq 10^5$ eV $= 100$ keV. The synchrotron-loss time is $\tau_r = \gamma_c^2\gamma_{\max}\omega^{-1} \simeq 10$ s. All the energy stored in the magnetic tail, $\epsilon \simeq \epsilon_m = 3.3 \times 10^{30}$ erg, will go into gamma-ray radiation over this time scale. Since the accelerated particles will move along the magnetic field in the tail at relativistic speeds ($\gamma \simeq \gamma_{\max} \simeq 10^6$), the directivity of the synchrotron radiation will be determined by the divergence angle of the magnetic field in the tail, $2\theta = 3.2^\circ \gg \gamma_{\max}^{-1}$. With this divergence angle, the flux of gamma-ray energy in the direction of the Earth, which is at a distance $D \simeq 0.6$ kpc from the binary system, will be equal to

$$F_{ph} = \epsilon_m/\pi\theta^2D^2\tau_r \simeq 4 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}.$$

Thus, flares of gamma-ray radiation with energies near 100 keV, durations of about 10 s, and fluxes of 4×10^{-11} erg cm^{-2} s^{-1} may be observed during the radio eclipses of pulsar A. These flares should be irregular, as is the case for magnetic substorms in the Earth's magnetosphere, due to the release of energy during the reconnection of magnetic field lines in the "tail" that forms due to the action of the solar wind.

In conclusion, we note that Istomin and Komberg [7, 8] have considered a model for the origin of cosmological gamma-ray bursts in which the energetics of the process is provided by the reconnection of magnetic-field lines in the narrow magnetospheric tail of a neutron star or white dwarf, as in the case of the binary pulsar considered here. However, in this model, the tail arises due to the action of the shock from a supernova that occurs in a close binary containing a compact magnetized stellar object. In this context, gamma-ray observations of PSR J0737−3039 could shed light on this possible mechanism for the radiation of "classical" gamma-ray bursts.

ACKNOWLEDGMENTS

The author thanks B.V. Komberg for useful discussions. This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16762) and the Program of Support for Leading Scientific Schools (grant no. NSh-1503.2003.2).

REFERENCES

1. A. G. Lyne, M. Burgay, M. Kramer, *et al.*, astro-ph/0401086 (2004).
2. V. M. Kaspi, S. M. Ransom, D. C. Backer, *et al.*, astro-ph/0401614 (2004).
3. V. S. Beskin, A. V. Gurevich, and Ya. N. Istomin, *Physics of Pulsar Magnetosphere* (Cambridge Univ. Press, Cambridge, 1993).
4. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon Press, Oxford, 1960), p. 303.
5. M. Lyutikov, astro-ph/0403076 (2004).
6. Ya. N. Istomin, in *Pulsar Astronomy—2000 and Beyond*, Ed. by M. Kramer, N. Wex, and N. Wielebinvi (Astron. Soc. Pac., San Francisco, 2000), ASP Conf. Ser., Vol. 202, p. 533 (2000).
7. Ya. N. Istomin and B. V. Komberg, Astron. Zh. **79**, 1008 (2002) [Astron. Rep. **46**, 908 (2002)].
8. Ya. N. Istomin and B. V. Komberg, New Astron. **8**, 209 (2003).

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 450–458. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 505–513.

Boundary Conditions for Simulations of the Thermal Outburst of a Type Ia Supernova

M. V. Popov, S. D. Ustyugov, and V. M. Chechetkin

Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya pl. 4, Moscow, 125047 Russia

Received December 1, 2004; in final form, December 3, 2004

Abstract—We present a technique to calculate the boundary conditions for simulations of the development of large-scale convective instability in the cores of rotating white-dwarf progenitors of type Ia supernovae. The hydrodynamical equations describing this situation are analyzed. We also study the impact of the boundary conditions on the development of the thermal outburst. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Hydrodynamical processes in astrophysical objects are usually simulated within restricted domains. For this reason, the modeling of the relevant physical processes is influenced by the selection of boundary conditions. This is particularly important in models of large-scale convection, which require long computational times. In these cases, perturbations developing in the system expand over the entire computational domain, and their interaction with the boundary affects the convection pattern. Various types of boundary conditions for convection problems have been suggested and studied. Using the method of characteristics, Hossain and Mullan [1] found two types of boundary conditions for an open computational domain, which proved to be very suitable for studies of the convective zone in the Sun. It was shown that, when the domain is closed and pressure averaging is applied at the boundary, this results in qualitative changes in the convection pattern and the type of acoustic oscillations that arise. Stein and Nordlund [2] fixed the energy and introduced a buffer layer in which the density and velocity fluctuations had finite amplitude. A zero velocity gradient was adopted at the lower boundary, and the pressure was selected taking into account the density variations. Gadun [3] used the condition of the absence of an average mass flux through the upper and lower boundaries of the domain.

Here, we use the local method of characteristics to study the impact of various boundary conditions on the development of convective processes in type Ia supernova outbursts. The simulations of convection in a supernova presented in [4] indicate the complexity of the flow structure: in particular, there are regions at the boundary where matter flows in or out at various rates. Some restrictions must therefore be imposed on the selection of boundary conditions, in order to make them suitable for this particular problem. Traditional boundary conditions used in previous models for convection in supernovae implied that the hydrodynamical parameters of the matter in boundary cells were constant, which is inappropriate when the large-timescale evolution of the matter flow is considered.

2. FORMULATION OF THE PROBLEM, INITIAL CONDITIONS, AND BASIC FORMULAS

In the course of its evolution, a star with a mass lower than $8M_\odot$ leaves the main sequence and turns into a red giant. A core consisting of a mixture of carbon and oxygen produced by thermonuclear burning in the inner layers of the star begins to form at its center. The core contains strongly ionized plasma, which can be described by the equation state for degenerate matter. Essentially, this is a white dwarf (WD) in the process of formation; under certain conditions, it may become a predecessor to a type Ia supernova. During its formation, the mass and temperature of the WD gradually increase. When the critical temperature $\sim 3 \times 10^8$ K has been reached, the CO mixture in the center of the WD ignites, and another core consisting of nuclei of iron-peak elements begins to form inside the WD. This process may result in a violation of thermal and mechanical equilibrium in the star and in the development of large-scale hydrodynamical instabilities. The thermonuclear burning of the degenerate matter in the WD occurs in a deflagration mode, as was shown in [5].

We will consider a rotating WD in which thermal instability has started to develop. According to current evolutionary concepts [6], the mass of the CO layer together with the "iron" core is roughly

one and a half solar masses; i.e., close to the Chandrasekhar limit. We will assume that the carbon and oxygen have equal masses, the radius of the star is $R_0 = 1.5 \times 10^8$ cm, and its central density is $\rho_0 = 2 \times 10^9$ g/cm^3. The equation of state $p = p(\rho, S)$ will be taken in a tabulated form appropriate for totally ionized matter, with the electron–positron component described by Fermi–Dirac statistics applying various asymptotics and the ion component described in an ideal-gas approximation [7, 8].

If the stellar matter is taken to be a compressible, nonviscous fluid, the hydrodynamical equations in the Euler variables can be written in the form

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{m} = 0, \\ \dfrac{\partial m_i}{\partial t} + \dfrac{\partial \Pi_{ik}}{\partial x_k} = \rho g_i, \\ \dfrac{\partial (\rho S)}{\partial t} + \operatorname{div}(\rho S \mathbf{v}) = 0, \end{cases} \tag{1}$$

where $\Pi_{ik} = p\,\delta_{ik} + \rho v_i v_k$, $\mathbf{m} = \rho \mathbf{v}$ is the momentum, g_i are the components of the gravitational acceleration, p is the pressure, and S is the entropy. We chose spherical coordinates in which $(x_1, x_2, x_3) = (r, \theta, \phi)$, and the Lame coefficients are $(h_1, h_2, h_3) = (1, r, r\sin\theta)$. The formulas for the divergence of a vector and tensor in two dimensions in curvilinear coordinates are given in the Appendix to [4]. It is convenient to present the system (1) in divergent form. To this end, we introduce the vector $\mathbf{w}$, which consists of conservative variables and enters into the partial time derivative, as well as the flux vectors $\mathbf{F}(\mathbf{w})$, $\mathbf{G}(\mathbf{w})$, and $\mathbf{H}(\mathbf{w})$, which enter into the partial derivatives with respect to the spatial variables. We also introduce the source vector $\mathbf{S}(\mathbf{w})$, which enters into the right-hand side of the vector equation:

$$\mathbf{w}_t + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \mathbf{F}\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \mathbf{G}\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\mathbf{H} = \mathbf{S}, \tag{2}$$

where

$$\mathbf{w} = \left(\rho, m_r, m_\theta, m_\phi, \rho S\right)^T,$$

$$\mathbf{F} = \left(m_r, p + \rho v_r^2, \rho v_\theta v_r, \rho v_\phi v_r, m_r S\right)^T,$$

$$\mathbf{G} = \left(m_\theta, \rho v_r v_\theta, p + \rho v_\theta^2, \rho v_\theta v_\phi, m_\theta S\right)^T,$$

$$\mathbf{H} = \left(m_\phi, \rho v_r v_\phi, \rho v_\theta v_\phi, p + \rho v_\phi^2, m_\phi S\right)^T,$$

$$\mathbf{S} = \left\{0, \rho g_r + \frac{1}{r}\left[2p + \rho\left(v_\theta^2 + v_\phi^2\right)\right], \rho g_\theta + \frac{1}{r}\left(p + \rho v_\phi^2\right)\cot\theta - \frac{1}{r}\rho v_r v_\theta, \rho g_\phi - \frac{\rho v_\phi}{r}\left(v_r + v_\theta \cot\theta\right), 0\right\}^T.$$

3. CHARACTERISTIC ANALYSIS

The system (2) is hyperbolic and describes the propagation of waves either entering or leaving the computational domain at its boundary, with velocities corresponding to various characteristics. A set of characteristics exists for each direction in the three-dimensional space; in spherical coordinates, these directions are r, θ, and ϕ. The waves emerging from the computational domain are totally specified by the solution (2) inside this domain; therefore, there is no need to specify any boundary condition to describe these waves. On the contrary, the propagation of waves entering the computational domain from the outside is specified by the solution (2) outside the domain and, thus, requires the specification of some particular boundary conditions. The number of boundary conditions can vary with time from zero to five, depending on the velocity of the matter at the boundary. All possible cases of the boundary conditions for a hyperbolic system of equations are presented in [9].

To describe the behavior of the waves and determine the boundary conditions, the characteristics of the system (2) must be analyzed. In spherical coordinates, we can carry out this analysis only for the r direction. Let us rewrite the system (2) in the form

$$\mathbf{w}_t + \frac{\partial \mathbf{F}}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta \mathbf{G}) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\mathbf{H} - \mathbf{S} + \frac{2\mathbf{F}}{r} = 0. \tag{3}$$

We write the derivative $\partial \mathbf{F}/\partial r$ in the form

$$\frac{\partial \mathbf{F}}{\partial r} = \frac{\partial \mathbf{F}}{\partial \mathbf{w}}\frac{\partial \mathbf{w}}{\partial r} = A(\mathbf{w})\frac{\partial \mathbf{w}}{\partial r}, \tag{4}$$

where $A(\mathbf{w})$ is the Jacobian matrix with the elements

$$a_{ij} = \frac{\partial F_i}{\partial w_j}.$$

The Jacobian of a hyperbolic system of equations has a complete set of right-handed and left-handed eigenvectors corresponding to the real eigenvalues. Hence, $A(\mathbf{w})$ can be presented in the form

$$A(\mathbf{w}) = R_{\mathbf{w}} \Lambda R_{\mathbf{w}}^{-1}, \tag{5}$$

where $R_{\mathbf{w}}$ is the matrix whose columns are the right-handed eigenvectors, $R_{\mathbf{w}}^{-1}$ is the inverse matrix, whose rows are the left-handed eigenvectors, and Λ is the diagonal matrix of the eigenvalues: $\Lambda_{ij} = 0$ for $i \neq j$, $\Lambda_{ij} = \lambda_i$ for $i = j$, where λ_i are the solutions for the characteristic equation

$$\det|\lambda E - A| = 0. \tag{6}$$

The solution of (6) is presented by the eigenvalues $\lambda_1 = v_r + c$, $\lambda_2 = v_r - c$, and $\lambda_3 = \lambda_4 = \lambda_5 = v_r$,

where $c = \sqrt{(p_\rho)_s}$ is the sound speed. The right-handed eigenvectors of the Jacobian $A(\mathbf{w})$ were calculated in [4]. Let us write the matrix of right-handed eigenvectors $R_\mathbf{w}$:

$$R_\mathbf{w} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ v_r + c & v_r - c & v_r & 0 & 0 \\ v_\theta & v_\theta & 0 & 1 & 0 \\ v_\phi & v_\phi & 0 & 0 & 1 \\ S & S & -\dfrac{\rho}{\xi} + S & 0 & 0 \end{pmatrix}.$$

Here, for convenience, we denote $\xi = (p_s)_\rho/(p_\rho)_s$. The inverse matrix of the left-handed eigenvectors $R_\mathbf{w}^{-1}$ is

$$R_\mathbf{w}^{-1} = \begin{pmatrix} \dfrac{1}{2} - \dfrac{v_r}{2c} - \dfrac{S\xi}{2\rho} & \dfrac{1}{2c} & 0 & 0 & \dfrac{\xi}{2\rho} \\ \dfrac{1}{2} + \dfrac{v_r}{2c} - \dfrac{S\xi}{2\rho} & -\dfrac{1}{2c} & 0 & 0 & \dfrac{\xi}{2\rho} \\ \dfrac{S\xi}{\rho} & 0 & 0 & 0 & -\dfrac{\xi}{\rho} \\ -v_\theta + v_\theta \dfrac{S\xi}{\rho} & 0 & 1 & 0 & -v_\theta \dfrac{\xi}{\rho} \\ -v_\phi + v_\phi \dfrac{S\xi}{\rho} & 0 & 0 & 1 & -v_\phi \dfrac{\xi}{\rho} \end{pmatrix}.$$

Using (4) and (5) and multiplying (3) by $R_\mathbf{w}^{-1}$ from the left, we obtain

$$R_\mathbf{w}^{-1}\mathbf{w}_t + \Lambda R_\mathbf{w}^{-1}\frac{\partial \mathbf{w}}{\partial r} + R_\mathbf{w}^{-1}\left\{\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\mathbf{G}) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{H} - \mathbf{S} + \frac{2\mathbf{F}}{r}\right\} = 0. \quad (7)$$

The term $R_\mathbf{w}^{-1}\mathbf{w}_t$ represents the time derivatives of the amplitudes of waves propagating along the characteristics:

$$\mathbf{a}_t = R_\mathbf{w}^{-1}\mathbf{w}_t.$$

For example, with accuracy to within a factor of $\rho/2c$, the first component,

$$(a_1)_t = \left(R_\mathbf{w}^{-1}\mathbf{w}_t\right)_1 = \frac{1}{2}\left(\xi\frac{\partial S}{\partial t} + \frac{\partial \rho}{\partial t} + \frac{\rho}{c}\frac{\partial v_r}{\partial t}\right) = \frac{\rho}{2c}\left(\frac{1}{\rho c}\frac{\partial p}{\partial t} + \frac{\partial v_r}{\partial t}\right),$$

is the time derivative of the amplitude of the wave propagating along the characteristic that corresponds to the velocity $v_r + c$ (see, for example, [10]). Let us introduce the vector

$$\mathbf{L} = \Lambda R_\mathbf{w}^{-1}\frac{\partial \mathbf{w}}{\partial r}, \quad (8)$$

and write (7) in the form

$$\mathbf{a}_t + \mathbf{L} + R_\mathbf{w}^{-1}\left[\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\mathbf{G}) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{H} - \mathbf{S} + \frac{2\mathbf{F}}{r}\right] = 0. \quad (9)$$

The selection of the boundary conditions depends on the way in which the $\mathbf{L}$ are determined. According to (8), they can be written in the form

$$L_1 = \frac{1}{2}(v_r + c)\left\{\xi\frac{\partial S}{\partial r} + \frac{\partial \rho}{\partial r} + \frac{\rho}{c}\frac{\partial v_r}{\partial r}\right\}, \quad (10)$$

$$L_2 = \frac{1}{2}(v_r - c)\left\{\xi\frac{\partial S}{\partial r} + \frac{\partial \rho}{\partial r} - \frac{\rho}{c}\frac{\partial v_r}{\partial r}\right\},$$

$$L_3 = -v_r\xi\frac{\partial S}{\partial r},$$

$$L_4 = v_r\rho\frac{\partial v_\theta}{\partial r} - v_r v_\theta \xi\frac{\partial S}{\partial r},$$

$$L_5 = v_r\rho\frac{\partial v_\phi}{\partial r} - v_r v_\phi \xi\frac{\partial S}{\partial r}.$$

In numerical simulations, formulas (10) should be used only for waves that are leaving the computational domain; i.e., for characteristics whose eigenvalues λ_i are positive. For the remaining characteristics, the boundary conditions should be constructed based on physical reasoning.

4. NONREFLECTING BOUNDARY CONDITIONS

When modeling the development of large-scale convective instability in the core of a pre-supernova, nonreflecting boundary conditions can be used, making it possible to obtain a stable solution even for a rapidly rotating star. The use of this type of boundary condition was first suggested by Hedstrom [11]. Nonreflecting boundary conditions assume that no perturbations enter the computational domain from the outside. This means that the amplitudes of waves arriving from the direction r normal to the boundary (for which $\lambda_i < 0$) do not vary with time, $\mathbf{a}_t = 0$. It follows from (9) that this condition is satisfied if

$$\mathbf{L} = -R_\mathbf{w}^{-1}\left\{-\mathbf{S} + \frac{2\mathbf{F}}{r}\right\}.$$

We have omitted the terms

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\mathbf{G}) \quad \text{and} \quad \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{H},$$

since they describe variations of hydrodynamical values due to the fluxes in the θ and ϕ directions, and thus cannot affect perturbations arriving normal to the boundary. The amplitudes of waves propagating along the considered characteristics depend only on r and t.

In the given problem, the suppression of arriving perturbations is justified, since the evolution of a star and the development of hydrodynamical instability are essentially specified by processes in the central region of the star, while processes in its outer layers only slightly affect the thermal outburst. Thus, we obtain for L_i corresponding to the condition $\lambda_i < 0$

$$L_1 = \frac{1}{2}\left(-1 + \frac{v_r}{c} + \frac{S\xi}{\rho}\right) C_1 - \frac{1}{2c} C_2 - \frac{\xi}{2\rho} C_5, \tag{11}$$

$$L_2 = \frac{1}{2}\left(-1 - \frac{v_r}{c} + \frac{S\xi}{\rho}\right) C_1 + \frac{1}{2c} C_2 - \frac{\xi}{2\rho} C_5,$$

$$L_3 = -\frac{S\xi}{\rho} C_1 + \frac{\xi}{\rho} C_5,$$

$$L_4 = \left(1 - \frac{S\xi}{\rho}\right) v_\theta C_1 - C_3 + \frac{\xi}{\rho} v_\theta C_5,$$

$$L_5 = \left(1 - \frac{S\xi}{\rho}\right) v_\phi C_1 - C_4 + \frac{\xi}{\rho} v_\phi C_5,$$

where C_i are the components of the vector

$$\mathbf{C} = -\mathbf{S} + \frac{2\mathbf{F}}{r}.$$

Let us multiply the system (7) by $R_{\mathbf{w}}$ from the left:

$$\mathbf{w}_t + R_{\mathbf{w}}\mathbf{L} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\mathbf{G}) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{H} - \mathbf{S} + \frac{2\mathbf{F}}{r} = 0.$$

In open form, this system of equations has the appearance

$$\begin{cases}
\dfrac{\partial\rho}{\partial t} + L_1 + L_2 + L_3 + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta m_\theta) + \dfrac{1}{r\sin\theta}\dfrac{\partial m_\phi}{\partial\phi} + \dfrac{2m_r}{r} = 0, \\
\dfrac{\partial m_r}{\partial t} + (v_r + c) L_1 + (v_r - c) L_2 + v_r L_3 + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta\rho v_r v_\theta) \\
+ \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\phi}(\rho v_r v_\phi) - \rho g_r - \dfrac{\rho}{r}\left(v_\theta^2 + v_\phi^2\right) + \dfrac{2\rho v_r^2}{r} = 0, \\
\dfrac{\partial m_\theta}{\partial t} + v_\theta (L_1 + L_2) + L_4 + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}\left[\sin\theta\left(p + \rho v_\theta^2\right)\right] + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\phi}(\rho v_\theta v_\phi) \\
- \rho g_\theta - \dfrac{1}{r}\left(p + \rho v_\phi^2\right)\cot\theta + \dfrac{3}{r}\rho v_r v_\theta = 0, \\
\dfrac{\partial m_\phi}{\partial t} + v_\phi (L_1 + L_2) + L_5 + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta\rho v_\theta v_\phi) + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\phi}\left(p + \rho v_\phi^2\right) \\
- \rho g_\phi + \dfrac{\rho v_\phi}{r}(3v_r + v_\theta\cot\theta) = 0, \\
\dfrac{\partial(\rho S)}{\partial t} + S(L_1 + L_2) + \left(-\dfrac{\rho}{\xi} + S\right) + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta m_\theta S) + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\phi}(m_\phi S) + \dfrac{2m_r S}{r} = 0.
\end{cases} \tag{12}$$

We use either (10) or (11) to calculate the L_i, depending on the sign of the corresponding eigenvalues. Namely, four cases of matter flow at the boundary are possible.

1. Subsonic outflow: $|v_r| < c$, $v_r \geq 0$; L_1, L_3, L_4, L_5 are calculated using (10), and L_2 using (11).

2. Subsonic inflow: $|v_r| < c$, $v_r < 0$; L_1 is calculated using (10), and L_2, L_3, L_4, and L_5 using (11).

3. Supersonic outflow: $|v_r| \geq c$, $v_r > 0$; all the L_i are calculated using (10).

4. Supersonic inflow: $|v_r| \geq c$, $v_r < 0$; all the L_i are calculated using (11).

5. BOUNDARY CONDITIONS THAT ARE STATIONARY WITH RESPECT TO THE RADIUS

Another type of boundary condition that can be applied for simulations is conditions that are stationary with respect to the radius. To derive such boundary conditions, we will rewrite the system (12)

in physical variables, i.e., we transform the conservative variables $(\rho, m_r, m_\theta, m_\phi, \rho S)$ into the variables $(\rho, v_r, v_\theta, v_\phi, S)$. We obtain after simple manipulation

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + L_1 + L_2 + L_3 + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta \rho v_\theta) \\ + \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\phi}(\rho v_\phi) + \dfrac{2\rho v_r}{r} = 0, \\ \dfrac{\partial v_r}{\partial t} + \dfrac{v_\theta}{r}\dfrac{\partial v_r}{\partial\theta} + \dfrac{v_\phi}{r\sin\theta}\dfrac{\partial v_r}{\partial\phi} - \dfrac{c}{\rho}(L_2 - L_1) \\ - g_r - \dfrac{v_\theta^2 + v_\phi^2}{r} = 0, \\ \dfrac{\partial v_\theta}{\partial t} + \dfrac{v_\theta}{r}\dfrac{\partial v_\theta}{\partial\theta} + \dfrac{v_\phi}{r\sin\theta}\dfrac{\partial v_\theta}{\partial\phi} + \dfrac{1}{\rho r}\dfrac{\partial p}{\partial\theta} + \dfrac{1}{\rho}L_4 \\ - \dfrac{v_\theta}{\rho}L_3 - g_\theta + \dfrac{v_r v_\theta}{r} - \dfrac{v_\phi^2}{r}\cot\theta = 0, \\ \dfrac{\partial v_\phi}{\partial t} + \dfrac{v_\theta}{r}\dfrac{\partial v_\phi}{\partial\theta} + \dfrac{v_\phi}{r\sin\theta}\dfrac{\partial v_\phi}{\partial\phi} + \dfrac{1}{\rho r\sin\theta}\dfrac{\partial p}{\partial\phi} \\ + \dfrac{1}{\rho}L_5 - \dfrac{v_\phi}{\rho}L_3 - g_\phi + \dfrac{v_\phi}{r}(v_r + v_\theta\cot\theta) = 0, \\ \dfrac{\partial S}{\partial t} + \dfrac{v_\theta}{r}\dfrac{\partial S}{\partial\theta} + \dfrac{v_\phi}{r\sin\theta}\dfrac{\partial S}{\partial\phi} - \dfrac{1}{\xi}L_3 = 0. \end{cases} \tag{13}$$

Boundary conditions that are stationary with respect to the radius assume that the physical parameters at the boundary can vary only due to fluxes along the θ and ϕ directions. Therefore, the combined contribution from the other terms must be zero. This results in the conditions

$$\begin{cases} L_1 + L_2 + L_3 + \dfrac{2\rho v_r}{r} = 0, \\ -\dfrac{c}{\rho}(L_2 - L_1) - g_r - \dfrac{v_\theta^2 + v_\phi^2}{r} = 0, \\ \dfrac{1}{\rho}L_4 - \dfrac{v_\theta}{\rho}L_3 - g_\theta + \dfrac{v_r v_\theta}{r} - \dfrac{v_\phi^2}{r}\cot\theta = 0, \\ \dfrac{1}{\rho}L_5 - \dfrac{v_\phi}{\rho}L_3 - g_\phi + \dfrac{v_\phi}{r}(v_r + v_\theta\cot\theta) = 0, \\ -\dfrac{1}{\xi}L_3 = 0. \end{cases}$$

Hence, we derive the expressions for the L_i:

$$L_1 = \frac{\rho}{2c}\left(g_r + \frac{v_\theta^2 + v_\phi^2}{r}\right) - \frac{\rho v_r}{r}, \tag{14}$$

$$L_2 = -\frac{\rho}{2c}\left(g_r + \frac{v_\theta^2 + v_\phi^2}{r}\right) - \frac{\rho v_r}{r},$$

$$L_3 = 0,$$

$$L_4 = \rho g_\theta - \frac{\rho v_r v_\theta}{r} + \frac{\rho v_\phi^2}{r}\cot\theta,$$

$$L_5 = \rho g_\phi - \frac{\rho v_\phi}{r}(v_r + v_\theta\cot\theta).$$

As in the case of nonreflecting boundary conditions, formulas (14) are used only for negative eigenvalues, and formulas (10) should be used for positive eigenvalues.

Note that any extrapolation of the physical variables, for example, v_r, will result in instability in these variables in the r direction. In other words, the steady-state condition cannot be reduced, for example, to the condition

$$\frac{\partial v_r}{\partial r} = 0. \tag{15}$$

To prove this, let us multiply the expression for L_1 from (10) by $2/(v_r + c)$, and the expression for L_2 by $2/(v_r - c)$, then subtract the second from the first. (We are considering the case $v_r \neq \pm c$.) This yields

$$\frac{\partial v_r}{\partial r} = \frac{c}{\rho}\left(\frac{L_1}{v_r + c} - \frac{L_2}{v_r - c}\right).$$

If we use the condition (15) to calculate the velocity component v_r at the boundary, i.e., if we assume that $(v_r)_N = (v_r)_{N-1}$, where N is the number of a boundary cell, we will obtain

$$\frac{L_1}{v_r + c} = \frac{L_2}{v_r - c}.$$

However, another condition must be satisfied for a stationary flow:

$$L_1 + L_2 = -\frac{2\rho v_r}{r},$$

which follows from (14).

All attempts to extrapolate physical variables at the boundary during simulations result in numerical instabilities.

6. DIFFERENCE SCHEME

We can construct a difference grid by manipulating the fourth equation of (12), bringing the term $1/r\ \rho v_\phi v_\theta \cot\theta$ inside the derivative:

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta \rho v_\theta v_\phi) + \frac{\rho v_\theta v_\phi \cot\theta}{r} = \frac{1}{r\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin^2\theta \rho v_\theta v_\phi\right).$$

We number the cells of the difference grid with the indices i, j and k, which vary in the r, θ, and ϕ directions, respectively. Integer indices are related to the cell centers, and half-integer indices to the cell

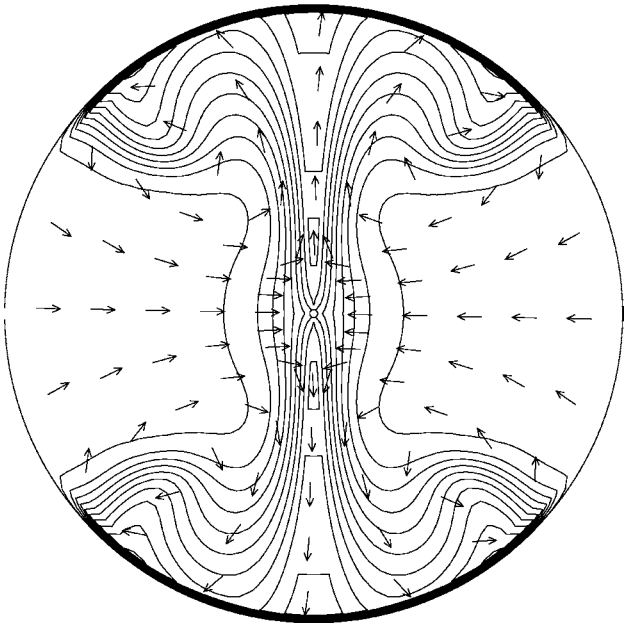


Fig. 1. Meridional cross section of the isentropic surfaces for the case of historical boundary conditions at $t = 0.15$ s. The arrows indicate the direction of the matter flow.

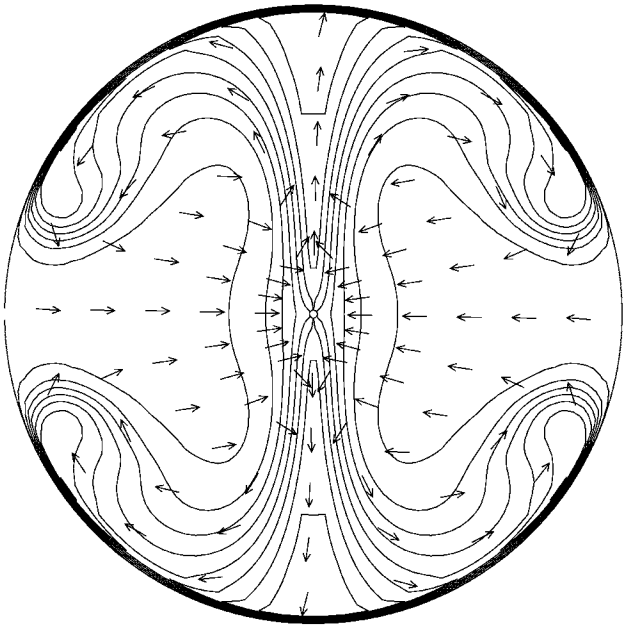


Fig. 2. Same as Fig. 1 for $t = 0.20$ s.

boundaries. To write the equations of (12) in difference form, we replace the derivatives with the differential relations

$$\frac{\partial \mathbf{w}}{\partial t} = \frac{\mathbf{w}^{n+1} - \mathbf{w}^n}{\tau},$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\mathbf{G}\right) = \frac{\sin\theta_{j+1/2}\mathbf{G}_{j+1/2} - \sin\theta_{j-1/2}\mathbf{G}_{j-1/2}}{r_i\left(\cos\theta_{j-1/2} - \cos\theta_{j+1/2}\right)},$$

$$\frac{1}{r\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin^2\theta\mathbf{G}\right) = \frac{\sin^2\theta_{j+1/2}\mathbf{G}_{j+1/2} - \sin^2\theta_{j-1/2}\mathbf{G}_{j-1/2}}{r_i\sin\theta_j\left(\cos\theta_{j-1/2} - \cos\theta_{j+1/2}\right)},$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{H} = \frac{\mathbf{H}_{k+1/2} - \mathbf{H}_{k-1/2}}{r_i\sin\theta_j\left(\phi_{k+1/2} - \phi_{k-1/2}\right)},$$

where $\mathbf{w}^n$ is the value of the vector $\mathbf{w}$ at the nth time step and τ is the time step. In reality, all values on a three-dimensional grid are denoted with three indices, but we have omitted the repeating indices for simplicity.

Note that, if we expand the derivative with respect to θ in the third equation of (12), then the term $p\cot\theta/r$ will appear, which will cancel out this same term in the free term. To ensure that this also occurs in the differential approximation, we must assume that, in the corresponding term of the free term,

$$\cot\theta = \frac{\sin\theta_{j+1/2} - \sin\theta_{j-1/2}}{\cos\theta_{j-1/2} - \cos\theta_{j+1/2}}.$$

We must assume that $\cot\theta = \cot\theta_j$ in the term $pv_\phi^2\cot\theta/r$. All remaining values in (12) are also related to the centers of cells. In the expressions (10) for L_i, the derivatives with respect to r are approximated by "reverse" differences. For example, L_1 will be written in the form

$$(L_1)_i = \frac{1}{2}\left((v_r)_i + c_i\right) \times\left\{\xi_i\frac{S_i - S_{i-1}}{r_i - r_{i-1}} + \frac{\rho_i - \rho_{i-1}}{r_i - r_{i-1}} + \frac{\rho_i}{c_i}\frac{(v_r)_i - (v_r)_{i-1}}{r_i - r_{i-1}}\right\}.$$

The method used to calculate the fluxes $\mathbf{G}$ and $\mathbf{H}$ at the cell boundaries is described in [4].

If we wish to determine the components of the gravitational acceleration, we must solve for the equilibrium configuration of a rotating gaseous sphere [4]. Due to the axial symmetry of the problem, $g_\phi = 0$, and only g_r and g_θ are present. In the calculations, we fixed the values of these components in tabulated form and did not vary them with time.

The system of equations (12) is used only at the cell boundaries. In the remainder of the domain, the computations must be carried out using the initial system (2), which can be solved in an analogous fashion (described in [4]).

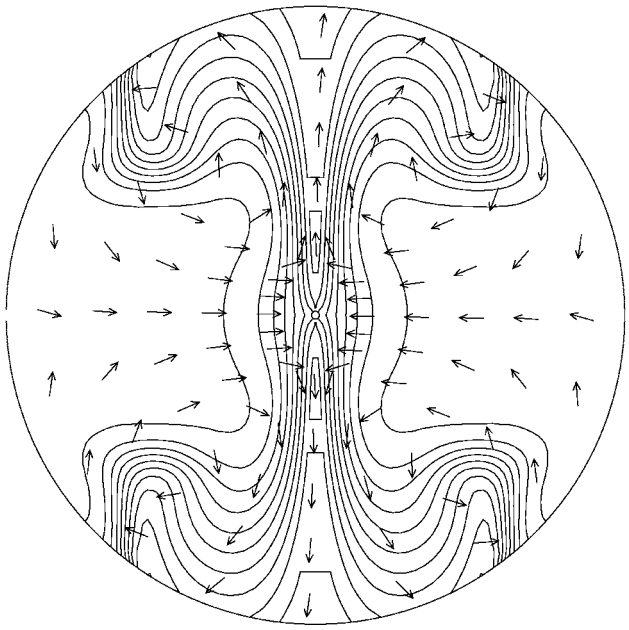


Fig. 3. Same as Fig. 1 for the case of nonreflecting boundary conditions.

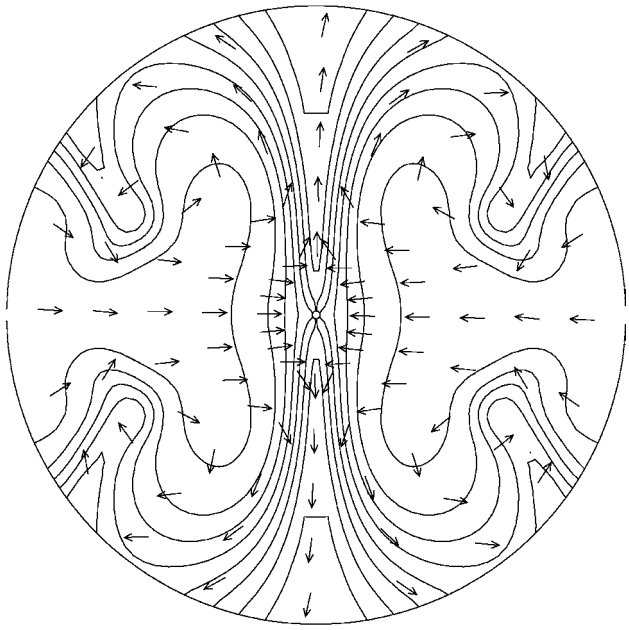


Fig. 4. Same as Fig. 3 for $t = 0.20$ s.

The time step τ is specified by the Courant condition, and is calculated at each step n using the formula

$$\tau = C_{Cour} \min_{i,j,k} \left\{ \frac{r_{i+1/2} - r_{i-1/2}}{|v_r| + c}, \right.$$
$$\left. \frac{r_i \left(\theta_{j+1/2} - \theta_{j-1/2}\right)}{|v_\theta| + c}, \frac{r_i \sin\theta_j \left(\phi_{k+1/2} - \phi_{k-1/2}\right)}{|v_\phi| + c} \right\},$$

where $C_{Cour} = \text{const}$ is the Courant number ($0 < C_{Cour} < 1$). The minimum is taken over the total computational domain.

7. RESULTS FOR DIFFERENT BOUNDARY CONDITIONS

We modeled a thermal outburst in a rotating star using three types of boundary conditions:

(1) historical;

(2) nonreflecting;

(3) stationary with respect to the radius for inflowing and nonreflecting for outflowing matter.

The computations were carried out on a $\{N_\phi \times N_\theta \times N_r\} = \{40 \times 80 \times 40\}$ grid with Courant number $C_{cour} = 0.8$ in the approximation of rigid-body rotation. The ratio of the rotational T and gravitation W energies was taken to be $T/|W| = 0.01$, which corresponds to an angular velocity for the rotation of $\Omega_0 = 2.0732\ \text{s}^{-1}$.

Figures 1 and 2 present a meridional cross section of the isentropic surfaces for the case of historical boundary conditions for times $t = 0.15$ s and $t = 0.20$ s. We can see a densification of the contours of constant entropy in the vicinity of the boundary of the computational domain, due to the reflection of perturbations reaching the boundary. This results in variations in the matter flow. The matter that has been burned in the center, which possesses increased entropy, starts to move at the inner surface of the computational domain. The boundary conditions do not allow a free outflow of matter along the axes of rotation, and prevent the formation of a regular jet structure. All the convective motion is strictly limited to the region within the computational domain.

Figures 3 and 4 present the results of the computations with nonreflecting boundary conditions for the same times. In this case, no reflection of the perturbations from the boundary is visible. The type of convective motion is maintained, and matter freely passes through the boundary.

Figures 5 and 6 present the results of the computations with boundary conditions that are stationary with respect to the radius for inflowing and nonreflecting for outflowing matter. This mixed type of boundary condition was selected based on the following physical reasoning. The velocity of the matter flow along the rotational axis approach supersonic values; consequently, only a small fraction of the matter can be reflected from the boundary and return to the computational domain. Therefore, nonreflecting boundary conditions were used for the outflowing matter. This essentially reduces to computing only L_2 using (11) for the case of subsonic outflow, since only L_2 corresponds to a negative eigenvalue for $\lambda_2 = v_r - c$. Thus,

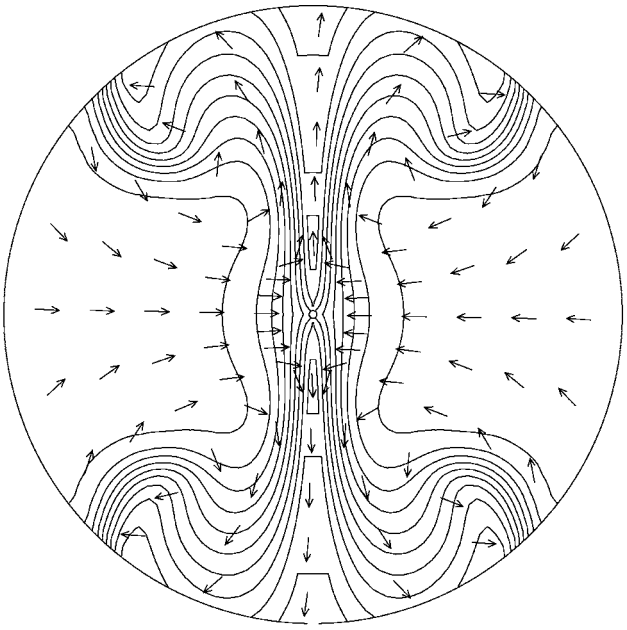


Fig. 5. Same as in Fig. 1 for the case of boundary conditions that are stationary with respect to the radius for inflowing and nonreflecting for outflowing matter.

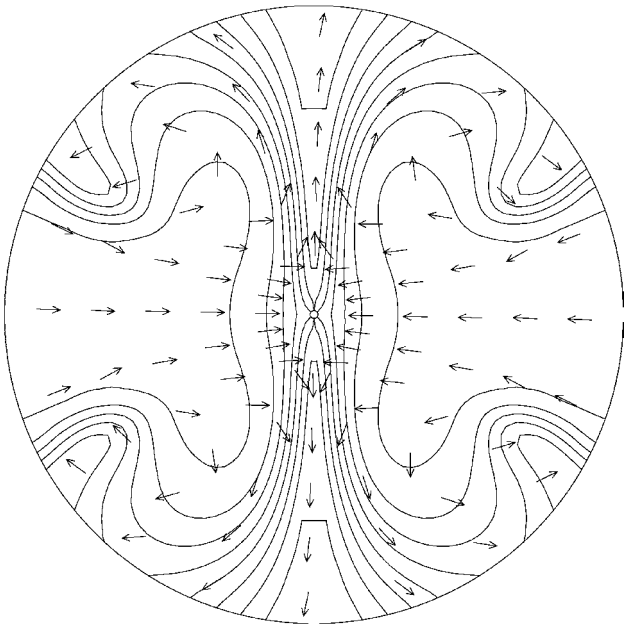


Fig. 6. Same as Fig. 5 for $t = 0.20$ s.

supersonic outflow does not require the determination of boundary conditions. Cool thermonuclear fuel from the outer layers of the star should flow in along the equatorial plane. Here, we set up steady-state inflow conditions.

The results obtained with this type of boundary condition differ very little from those obtained with the nonreflecting boundary conditions. The evolution of the isentropic surfaces is similar in both cases (Figs. 4, 6), with matter flowing along the equatorial plane from the boundary toward the center. In the case of nonreflecting boundary conditions (Fig. 4), this flow captures some of the burned matter with increased entropy and transports it back toward the center, which does not happen when the boundary conditions for the inflowing matter are stationary with respect to the radius (Fig. 6). However, this process does not substantially affect the thermal outburst, since the mass fraction of the burned matter that is transported back toward the center is extremely small. Nonetheless, we prefer the use of boundary conditions that are stationary with respect to the radius for inflowing matter and nonreflecting for outflowing matter, since, as can be seen from Figs. 5 and 6, the influence of the boundary is lowest in this case.

We also studied a model with more rapid rotation: $T/|W| = 0.05$, $\Omega_0 = 4.2880$ s^{-1}. In this case, the differences become appreciable. With historical boundary conditions, reflection from the boundary destroys the convective process and results in numerical instabilities in the vicinity of the boundary. With nonreflecting boundary conditions and steady-state radial inflow, the boundary does not affect the type and velocity of the convective motion, and no numerical problems occur. Note that, with this rotation, the type of motion at the boundary in the equatorial plane is changed when the convection reaches this region. Initially subsonic inflow becomes supersonic, and the number of characteristics that must be determined, depending on the type of boundary condition, will change.

8. CONCLUSION

We have shown that, on large timescales with an existing convection pattern, nonreflecting boundary conditions for the outgoing flow and the condition of a steady-state flow for the incoming flow should be used. Nonreflecting boundary conditions for incoming perturbations yield similar results for all cases of the flow at the boundary. The traditional use of historical boundary conditions, as well as extrapolating the calculated values at the boundary, are inappropriate and result in numerical instabilities. These boundary conditions can be applied only in early stages of the development of convection in the star, when the influence of boundary conditions can be neglected.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 03-02-16548)

and INTAS (grant no. 01-491). We also acknowledge support from the Federal Program in Support of Leading Scientific Schools of the Russian Federation (grant no. NSh-1029.2003.2).

REFERENCES

1. M. Hossain and D. J. Mullan, Astrophys. J. **416**, 733 (1993).
2. R. F. Stein and A. Nordlund, Astrophys. J. Lett. **342**, L95 (1989).
3. A. S. Gadun, Preprint No. 86-106R, ITF UAN (Inst. Theor. Phys. Ukrainian Acad. Sci., Kiev, 1986).
4. M. V. Popov, S. D. Ustyugov, and V. M. Chechetkin, Astron. Zh. **81**, 1011 (2004) [Astron. Rep. **48**, 921 (2004)].
5. V. S. Imshennik, N. L. Kal'yanova, A. V. Koldoba, *et al.*, Pis'ma Astron. Zh. **25**, 250 (1999) [Astron. Lett. **25**, 206 (1999)].
6. V. S. Imshennik and D. K. Nadezhin, *Rez. Nauki Tekh.* (VINITI, Moscow, 1982), Vol. 21.
7. D. K. Nadezhin, Nauch. Inf. Astron. Sovet Akad. Nauk SSSR **32** (1974).
8. D. K. Nadezhin, Nauch. Inf. Astron. Sovet Akad. Nauk SSSR **33** (1974).
9. W. Kevin and J. Thompson, Comput. Phys. **89**, 439 (1990).
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
11. L. D. Hedstrom, J. Comput. Phys. **30**, 222 (1979).

Translated by K. Maslennikov

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 459–462. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 514–518.

The X-ray Emission of Magnetars

I. F. Malov[1] and G. Z. Machabeli[2]

[1] *Pushchino Radio Astronomy Observatory, Astro Space Center of the Lebedev Physical Institute, Pushchino, Russia*

[2] *Abastumani Astrophysical Observatory, Abastumani, Tbilisi, Georgia*

Received August 5, 2004; in final form, December 3, 2004

Abstract—It is shown that cyclotron radiation by electrons near the surface of a neutron star with a magnetic field of $\sim 10^{12}$ G can easily provide the observed quiescent radiation of magnetars (Anomalous X-ray Pulsars and Soft Gamma-ray Repeaters). Pulsed emission is generated by the synchrotron mechanism at the periphery of the magnetosphere. Short-time-scale cataclysms on the neutron star could lead to flares of gamma-ray radiation with powers exceeding the power of the X-ray emission by a factor of $2\gamma^2$, where γ is the Lorentz factor of the radiating particles. It is shown that an electron cyclotron line with an energy of roughly 1 MeV should be generated in the magnetar model. The detection of this line would serve as confirmation of the correctness of this model. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Interest in magnetars—anomalous X-ray pulsars (AXPs) and soft gamma-ray repeaters (SGRs)—has risen recently due to the detection of cyclotron absorption lines with energies of about 5–8 keV for several of these sources [1–3]. In the most popular theory [4], these objects are neutron stars with anomalously strong surface magnetic fields B_s of the order of 10^{14}–10^{15} G (two to three orders of magnitude higher than the fields of ordinary pulsars). The magnetic-field strength can be derived from the observed parameters of the cyclotron lines. If it is assumed that the observed lines are emitted at the electron-cyclotron frequencies, the implied magnetic-field strengths do not exceed 10^{12} G; i.e., they correspond to the values characteristic of ordinary radio pulsars.

The proponents of the standard magnetar model suggest that the objects can still possess fields of 10^{14}–10^{15} G if the observed absorption lines correspond to proton-cyclotron frequencies. However, in this case, it is unclear why we do not observe electron-cyclotron lines formed in more distant regions of the magnetosphere. Such lines should be observed, since the magnetic field falls off with distance from the surface (according to an inverse cubic law in the case of a dipolar field). Doubt is also cast on the proton-cyclotron hypothesis by the results of [5], which demonstrate that, in fields exceeding the critical value $B_{cr} = 2\pi m^2 c^3/eh = 4.4 \times 10^{13}$ G, the proton cyclotron lines should be smeared by effects associated with the polarization of the vacuum, and should therefore not be observed. Accordingly, increasing importance is being acquired by alternative models that do not require fields of the order of 10^{14}–10^{15} G.

In [6], we proposed a model based on the theory developed by Machabeli and collaborators over more than 20 years (see, for example, [7–9]). The idea behind this model is as follows. Electromagnetic waves are generated in the electron–positron plasma filling the magnetosphere. These waves propagate along the pulsar's magnetic-field lines, and quasilinear diffusion develops due to the influence of these waves on the particles. This diffusion changes the particle-distribution function from one- to two-dimensional; i.e., the resonance particles of the plasma acquire orthogonal momenta. This, in turn, brings the synchrotron mechanism into action, which generates high-frequency optical or X-ray synchrotron radiation, depending on the parameters of the plasma [10, 11]. The wavelength of this high-frequency radiation λ is much smaller than the "mean" distance between the particles, $n^{-1/3}$ (where n is the plasma density), so that this radiation essentially does not interact with the ambient medium.

In addition to the waves propagating along the field lines, a drift wave that moves across the magnetic field is generated in the plasma in association with the drift of particles due to the curvature of the pulsar field [8]. This wave encircles the magnetosphere, and gives rise to variations in the radius of curvature of the surrounding magnetic-field lines. The high-frequency synchrotron radiation propagates tangent to the field lines in the region where it is generated.

1063-7729/05/4906-0459$26.00

Therefore, in the presence of variations in the curvature of the field lines, radiation will be directed toward the observer with a period equal to the period of the drift waves. In our model, it is precisely this period that determines the spacing between the X-ray pulses in AXPs and SGRs.

We present here a qualitative explanation for the observed X-ray quiescent radiation, and discuss a possible origin of gamma-ray bursts. Our model does not require the introduction of unusually strong magnetic fields, and explains the observations fully in the framework of standard theories for the radiation of ordinary radio pulsars.

2. GENERATION OF THE X-RAY BACKGROUND

We must separate the radiation generated in the magnetosphere into two components: eigenmode radiation and radiation in a one-particle approximation. The eigenmodes result from the radiation of the ensemble of particles making up the plasma in the pulsar magnetosphere. In the second type of radiation, each particle in this ensemble is treated as an independent source of radiation. The former case is realized when the wavelength λ is larger than the "mean" distance between particles in a plasma with density n ($\lambda > n^{-1/3}$), while the second case is realized when $\lambda < n^{-1/3}$.

In the standard theory [12, 13], the rotation of a conducting star with a frozen-in magnetic field induces an electric field that has a component along the magnetic field. This electric field tears charges from the surface of the neutron star, and can lead to the creation of electron–positron pairs. This process will occur only under certain conditions. First, the energy of the gamma-ray photon involved ε_γ must exceed $2m_ec^2$, where m_e is the mass of the electron and c is the speed of light. Second, the angle between the propagation of the gamma-ray and the magnetic field must be large enough to satisfy the inequality [14]

$$B_\perp \varepsilon_\gamma > 10^{18}\ \text{G eV}. \tag{1}$$

The gamma-rays are emitted along tangents to the curved magnetic-field lines. The angle between the direction of propagation of the gamma-ray and the magnetic field increases as the radiation travels from the place where it is generated, and condition (1) will begin to be satisfied when this angle becomes sufficiently large. At the same time, the gamma-rays will begin to be absorbed due to the development of the conversion process

$$\gamma + B \to e^+ + e^- + B + \gamma'. \tag{2}$$

The electrons e^- and positrons e^+ created at the surface of the neutron star should have fairly large transverse momenta $p_\perp$ (or pitch angles, $\psi = \arctan p_\perp/p_{||}$), which are radiated away due to synchrotron losses on a time scale $t \sim 10^{-15}$–10^{-14} s. Simultaneously, the electrons and positrons in the powerful magnetic-field of the pulsar ($B_s \sim 10^{11}$–10^{12} G) "settle" into Landau quantum levels.

Let us consider the range of frequencies at which radiation will be generated near the surface of the neutron star.

It is known (see, for example, [15]), that a frequency ν in the frame of the observer is associated with a frequency ν_0 in the co-moving frame (in which $V_{||} = 0$) by the relation

$$\nu = \nu_0 \frac{(1 - V^2/c^2)^{1/2}}{1 - V\cos\alpha/c}, \tag{3}$$

where α is the angle between the particle velocity and the line of sight toward the observer.

If the Lorentz factor of the radiating particles is $\gamma = (1 - V^2/c^2)^{-1/2} \gg 1$ and the angle α is small, expression (3) can be written

$$\nu = \frac{2\nu_0}{1/\gamma + \alpha^2\gamma}, \tag{4}$$

and, when $\alpha^2\gamma \ll 1/\gamma$,

$$\nu \approx 2\nu_0\gamma. \tag{5}$$

In the case of large values of α,

$$\nu \approx \nu_0/\gamma. \tag{6}$$

If

$$1 \lesssim \alpha^2\gamma \lesssim 10, \tag{7}$$

and $B_s \sim 10^{12}$ G, then the electron-cyclotron frequency,

$$\nu_0 = \frac{eB_s}{2\pi mc} \tag{8}$$

falls into the soft X-ray range (1–10 keV) in the observer's frame, and this X-ray radiation can escape into interstellar space after first crossing the magnetosphere filled with electron–positron plasma. Since the scatter in the angles α can be substantial and the distribution of Lorentz factors for the particles will not be monoenergetic, the resulting spectrum should be broad. Here, we consider only approximate estimates, and will carry out a more detailed calculation of the resulting spectrum in a future paper.

Since the dipolar magnetic field of the neutron star falls off with distance in accordance with an inverse cubic law, only the frequency in the generation region

can coincide with frequencies of harmonics corresponding to transitions between Landau levels [16]:

$$\varepsilon_m - \varepsilon_n = (p_{\perp m}^2 - p_{\perp n}^2)/2m_e = h\nu_0 S, \qquad (9)$$
$$S = (m-n) = \pm 1, \pm 2, \ldots$$

As we noted above, lines corresponding to these harmonics have been observed [1]. However, they have been interpreted as cyclotron lines associated with the absorption of non-relativistic protons in magnetic fields of $\sim 10^{14}$–10^{15} G [17]. If this interpretation is correct, electron-cyclotron lines should be observed near 1 MeV. The detection of these lines would provide clear confirmation of the existence of magnetic fields with $B > B_{cr}$; however, no such observational data have been reported.

In our model, protons do not play an important role, and it is electron-cyclotron lines that have been observed, so that such lines should not be observed near 1 MeV.

According to the theory of radiation by individual relativistic particles, the observer will receive radiation within a cone with opening angle $\theta \sim 1/\gamma$ (see, for example, [18]).

We will suppose that the radiation described by our model makes the dominant contribution to the X-ray quiescent radiation due to magnetars.

Another source of radiation along tangents to the magnetic field is present at distances comparable to the radius of the light cylinder [6], which creates the pulsed component, along with a possible additional contribution to the magnetar X-ray quiescent radiation. This radiation is likewise concentrated in a cone, so that the observer should observe a pattern of two cones, as is shown schematically in Fig. 1.

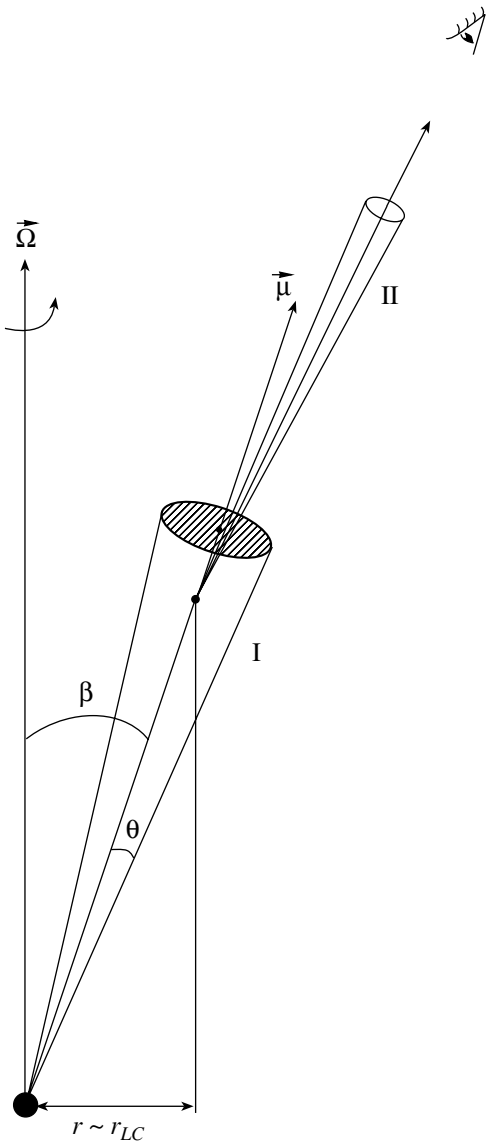


Fig. 1. Schematic of the two cones of radiation: cone I contains quasicontinuous X-ray radiation, while cone II contains the pulsed component.

3. IRREGULAR BURSTS OF POWERFUL GAMMA-RAY RADIATION

As we noted above, if the angle between the direction of propagation of the radiation and the line of sight is such that $1 < \alpha^2\gamma < 10$, the observer receives soft X-ray radiation. However, for various reasons (such as a starquake), the cone for the background radiation component can become tilted, such that the angle α becomes very small ($\alpha^2\gamma^2 \lesssim 1$). In this case, in accordance with formula (3), the frequency can be shifted into the gamma-ray range ($\nu \sim 2\gamma\nu_0$). This frequency depends strongly on the Lorentz factor of the radiating particles. It seems reasonable to suppose that particles with various energies should participate in this process, so that the resulting spectrum should be broad.

The transformation of the power of the radiation into the observer's frame is described by the relation

$$P_\nu = P_{\nu 0}\frac{1}{1 - V\cos\alpha/c}. \qquad (10)$$

When $\alpha \to 0$, the power P_ν grows sharply, and can reach values

$$P_\nu \approx 2P_{\nu 0}\gamma^2. \qquad (11)$$

We again emphasize that, in the quiescent state ($1 < \alpha^2\gamma < 10$), the observer will detect soft X-ray radiation directed along the corresponding field lines. At times when $\alpha < 1/\gamma$, the formation of powerful flares of gamma-ray radiation is possible. According

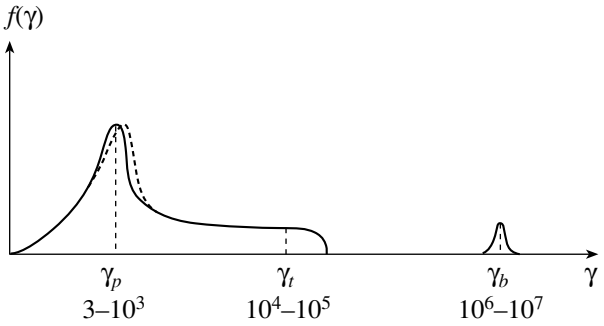


Fig. 2. Distribution of Lorentz factors for the relativistic plasma in the pulsar magnetosphere. The dashed curve shows the distribution for positrons.

to formula (11), this gamma-ray radiation can be a factor of $2\gamma^2$ stronger than the X-ray background radiation. If the X-ray power is $\sim 10^{36}$ erg/s, Lorentz factors $\gamma \sim 10^4$ are required to provide gamma-ray powers of 10^{44} erg/s. In the traditional model for radio pulsars, such energies are possessed by secondary electrons and positrons in the high-energy tail of the distribution (Fig. 2) [19].

4. DISCUSSION AND CONCLUSIONS

1. We have shown the fundamental possibility of explaining the X-ray background radiation of AXPs and SGRs at 1–10 keV using a model with surface magnetic fields for the neutron stars of $\sim 10^{12}$ G. The spectrum is formed primarily due to the cyclotron mechanism acting near the surface of the star, and should be fairly broad.

An analogous model for the formation of the directional beam and spectrum of the X-ray radiation of an anisotropic distribution of relativistic particles using the accreting pulsar Her X-1 as an example was considered in [20, 21].

2. Short-time-scale cataclysmic events on the neutron star that lead to a temporary coincidence between the axis of the radiation cone and the line of sight toward the Earth could give rise to bursts of gamma-ray emission.

3. The power of the gamma-ray flares should be a factor of $2\gamma^2$ higher than the power of the X-ray background emission. The flare spectrum should also be fairly broad.

4. The detection of electron-cyclotron lines near 1 MeV in AXPs and SGRs would provide support for the magnetar model. The absence of these lines in the observed spectra of these sources, on the contrary, provides indirect evidence in support of our model.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-02-16509) and the National Science Foundation (grant no. 00-98685).

REFERENCES

1. N. Rea, G. L. Izrael, and L. Stella, astro-ph/0309402.
2. A. I. Ibrahim *et al.*, astro-ph/0210513.
3. A. I. Ibrahim, J. H. Swank, and W. Parke, Astrophys. J. Lett. **584**, L17 (2003).
4. R. Duncan and C. Thompson, Astrophys. J. Lett. **392**, L19 (1992).
5. W. C. G. Ho, D. Lai, A. Y. Potekhin, and G. Chabrier, astro-ph/0309261.
6. I. F. Malov, G. Z. Machabeli, and V. M. Malofeev, Astron. Zh. **81**, 258 (2003) [Astron. Rep. **47**, 232 (2003)].
7. G. Z. Machabeli and V. V. Usov, Pis'ma Astron. Zh. **5**, 445 (1979) [Sov. Astron. Lett. **5**, 238 (1979)].
8. A. Z. Kazbegi, G. Z. Machabeli, and G. I. Melikidze, Aust. J. Phys. **44**, 573 (1987).
9. A. Z. Kazbegi, G. Z. Machabeli, and G. I. Melikidze, in *IAU Colloq. No. 128: Magnetospheric Structure and Emission Mechanisms of Radio Pulsars*, Ed. by T. H. Hankins, J. M. Rankin, and J. A. Gil (Pedagogical Univ. Press, 1992), p. 232.
10. I. F. Malov and G. Z. Machabeli, Astrophys. J. **554**, 587 (2001).
11. I. F. Malov and G. Z. Machabeli, Astron. Zh. **79**, 755 (2002) [Astron. Rep. **46**, 684 (2002)].
12. P. Goldreich and W. H. Julian, Astrophys. J. **157**, 869 (1969).
13. P. A. Sturrock, Astrophys. J. **164**, 529 (1971).
14. T. Erber, Rev. Mod. Phys. **38**, 626 (1966).
15. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Fizmatgiz, Moscow, 1967; Pergamon, Oxford, 1975).
16. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Fizmatgiz, Moscow, 1963; Pergamon, Oxford, 1977).
17. S. Zane, R. Turrola, L. Stella, and A. Treves, astro-ph/0103316.
18. G. Bekefi, *Radiation Processes in Plasmas*, Ed. by G. Bekefi (Wiley, New York, 1966; Mir, Moscow, 1971).
19. J. Arons, in *Plasma Physics*, Ed. by T. D. Guyenne (European Space Agency, Paris, 1981), p. 273.
20. G. S. Bisnovatyi-Kogan, Astron. Zh. **50**, 902 (1973) [Sov. Astron. **17**, 574 (1973)].
21. A. N. Baushev and G. S. Bisnovatyi-Kogan, Astron. Zh. **76**, 283 (1999) [Astron. Rep. **43**, 241 (1999)].

Translated by D. Gabuzda

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 463–469. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 519–526.

The Formation of Radio-Pulsar Spectra

G. Z. Machabeli and G. T. Gogoberidze

Center for Plasma Astrophysics, Abastumani Astrophysical Observatory, A. Kazbegi 2a, Tbilisi, 380060 Georgia

Received July 13, 2004; in final form, December 3, 2004

Abstract—It is shown that scattering of electromagnetic waves by Langmuir waves taking into account the electrical drift motion of the particles is the most efficient nonlinear process contributing to a radio pulsar's spectrum. If an inertial interval exists, stationary spectra with spectral indices of -1.5 or -1 can be formed, depending on the wave excitation mechanism. The obtained spectra are in satisfactory agreement with observational data. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Our aim here is to consider possible nonlinear processes that could contribute to the radiation emitted in a pulsar magnetosphere. The observed spectra of pulsars have been studied in detail (see, e.g., [1–3]). Pulsars typically have steeper spectra than other cosmic objects. The radio emission of various pulsars displays a wide range of spectral indices α ($I \sim \nu^{\alpha}$, where I is the observed intensity and ν is the radiated frequency), from $+1.4$ to -3.8 [1]. However, in the overwhelming majority of pulsars, $-1.5 > \alpha > -2$ [2]. Another characteristic feature of the spectra is a low-frequency cutoff (an abrupt change in the sign of the spectral index), observed in some pulsars from 39 to 400 MHz [4], and a high-frequency break (a sharp increase in the spectral index), observed from 400 MHz to 9 GHz [5].

We assume that the pulsar radiation is generated and its spectrum formed in the magnetosphere. The radiation leaves the magnetosphere, escapes into interstellar space, and reaches the observer, conserving the properties of the natural modes of the magnetospheric plasma. This approach has been able to explain a number of the main observed properties of pulsar radiation, such as its polarization properties [6], nulling [7], micropulses [8], mode switching [9], etc.

There is no doubt that the spectrum of any radiating object, including a pulsar, is one of its most important characteristics. However, there are virtually no theoretical models for the formation of a pulsar's spectrum. Successful modeling obviously requires studies of the turbulent state of the pulsar magnetosphere. There is no general theory of turbulence, and we must make do with the methods of weak-turbulence theory. In our case, such an approximation is justified by the actual situation: in a pulsar magnetosphere, we are dealing with a flow of relativistic electron–positron plasma that is driven beyond the light cylinder, after which the conditions change abruptly, so that there may not be enough time for the development of strong turbulence.

We consider here nonlinear processes that, in our opinion, are responsible for the formation of the characteristic spectrum of a pulsar's radio emission. The development of the theory of nonlinear processes in plasmas has resulted in the methods that we will use to analyze the plasma turbulence (see, e.g., [10, 11]).

Before proceeding to an analysis of nonlinear processes, we briefly describe the considered model for the pulsar magnetosphere in Section 2. Section 3 examines the natural modes of an electron–positron plasma and reviews linear mechanisms for their generation. In Section 4, we analyze the relevant nonlinear processes, and obtain stationary solutions for the spectra and present a physical analysis of the results in Section 5.

2. THE PLASMA OF A PULSAR MAGNETOSPHERE

We adhere to the generally accepted model of a pulsar magnetosphere [12, 13], in which a relativistic flow of electron–positron plasma pierced by an electron beam is moving along open force lines of the pulsar magnetic field $\mathbf{B}_0$. We will assume a dipolar pulsar magnetic field. We will take "open" field lines to be those intersecting the surface of the light cylinder (this is a hypothetical cylinder, on whose surface the linear velocity of a magnetic-field line reaches the speed of light c in the case of rigid-body rotation). We suppose that an electrostatic field is generated at the pulsar surface along the open field lines [12], which pulls out charges from the surface of the neutron star, and that it is the electrons accelerated in this field that form the primary beam. The primary beam moves along curved field lines, generating gamma-rays, which, in turn, propagate along tangents to the

1063-7729/05/4906-0463$26.00

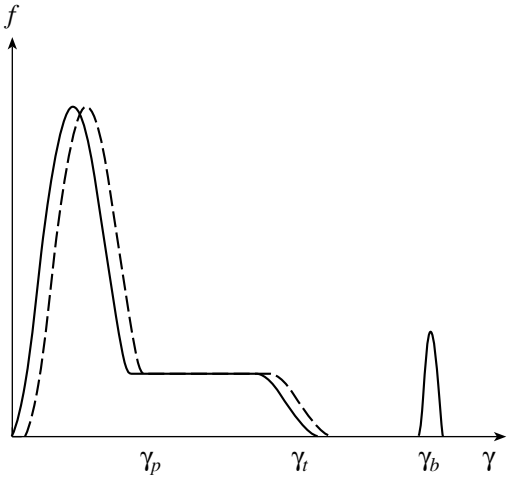


Fig. 1. Distribution function f of the pulsar-magnetosphere plasma. The solid and dashed curves show the distribution functions for electrons and positrons, respectively.

field lines, producing electron–positron pairs at energies $\epsilon_\gamma > 2m_e c^2$ (where m_e is the electron mass) [13]. This initiates a cascade process, which ceases when the shielding of the accelerating electrostatic field is full. In the generally accepted theory, the distribution function f has the form shown in Fig. 1. This distribution function can be represented

$$f = f_p + f_t + f_b, \tag{1}$$

where f_p describes the bulk of the plasma, f_t the long tail, which extends in the direction of the plasma motion, and f_b the beam of primary particles.

The distribution function (1) is one-dimensional, and extends asymmetrically toward the light cylinder. The solid and dashed curves in Fig. 1 show the distribution functions of the electrons and positrons, respectively. The shift of the distribution functions is due to the fact that the primary beam consists of particles of only one sign.

It is assumed that

$$n_p\gamma_p \approx n_t\gamma_t \approx \frac{1}{2}n_b\gamma_b, \tag{2}$$

where γ_p, γ_t, and γ_b are the Lorentz factors and n_p, n_t, and n_b the densities of the particles of the bulk, tail, and primary beam of the plasma, respectively.

For typical pulsars, $\gamma_b \sim 10^6$–10^7, $n_b^0 \sim 10^{11}$ cm^{-3}, $\gamma_t \sim 10^4$, $n_t^0 \sim 10^{13}$–10^{14} cm^{-3}, $\gamma_p \sim 10$, and $n_p^0 \sim 10^{17}$ cm^{-3}. The subscript 0 denotes values at the pulsar surface. With distance from the surface, the densities decrease as $n = n_0(r_0/r)^3$, where $r_0 \approx 10^6$ cm is the radius of the star and r is the distance from the pulsar center. Note that the dipole magnetic field of the pulsar also obeys a cubic law: $|B_0| = |B_0^0|(r_0/r)^3$, where $|B_0^0| \sim 10^{12}$ G is the magnetic induction at the surface.

It follows from the condition of quasi-neutrality that

$$\Delta\gamma = \gamma_+ - \gamma_- = \int f_+\gamma d^3p - \int f_-\gamma d^3p. \tag{3}$$

Here, the subscripts $\pm$ refer to positrons and electrons, respectively, and the values of $f_\pm$ are normalized such that

$$\int f_\pm d^3p = 1. \tag{4}$$

The value of $\Delta\gamma$ is small, but this quantity plays a key role in explaining the polarization properties of pulsars.

The radio emission of most pulsars is 2–10% circularly polarized at their intensity peaks. According to [6], this circular polarization can exist only within a small angle θ between the wave vector $\mathbf{k}$ and the pulsar magnetic field $\mathbf{B}_0$:

$$\theta^2 \lesssim \frac{\omega}{\omega_B}\Delta\gamma \ll 1. \tag{5}$$

This means that the observer receives emission that is formed within a small angle near the magnetic-field lines, and the field lines are directed toward the observer when the waves leave the plasma. It is, therefore, natural for us to consider waves generated along tangents to the slightly curved magnetic-field lines. This is consistent with the linear theory of electron–positron plasmas: waves are generated essentially along lines of force [6, 14].

3. LINEAR WAVES AND THEIR GENERATION

Due to the absence of gyrotropy, the spectrum of a magnetized electron–positron plasma is rather simple. It consists of three branches [14]: a purely transverse t wave and two potential–nonpotential lt waves. The t wave has no exact counterpart in a common electron–ion plasma. The electric vector of this wave is perpendicular to the plane containing the wave vector $\mathbf{k}$ and the magnetic field. During field-aligned propagation, the spectrum of these waves is

$$\omega_t = kc(1 - \delta), \tag{6}$$

where $\delta \equiv \omega_p^2/(4\omega_B^2\langle\gamma^3\rangle)$,

$$\omega_p \equiv \left(\frac{8\pi e^2 n_p}{m_e}\right)^{1/2}, \quad \omega_B \equiv \frac{e|B_0|}{m_e c} \tag{7}$$

are the plasma and cyclotron frequency, respectively, the angular brackets denote averaging over the distribution function, and $\langle\gamma\rangle \approx \gamma_p$.

The lt waves are divided into two branches: lt_1 and lt_2. The electric vectors of these waves lie in the plane containing the vectors $\mathbf{k}$ and $\mathbf{B}_0$. During field-aligned propagation ($\mathbf{k}||\mathbf{B}_0$), the lt_2 wave has sub-light speed ($\omega < kc$) and merges with the t wave into an electromagnetic wave with the spectrum (6) and its electric vector in some arbitrary direction perpendicular to $\mathbf{B}_0$.

During field-aligned propagation, the upper (lt_1) branch is longitudinal in its lower part and represents a Langmuir wave (l); its dispersion in the super-light-speed range ($\omega \gtrsim kc$) is [15, 16]

$$\omega_l^2 = \omega_p^2\langle\gamma^{-3}\rangle + 3k^2c^2\left(1 - \frac{\langle\gamma^{-5}\rangle}{\langle\gamma^{-3}\rangle}\right), \qquad (8)$$

and, for $\omega \approx kc$,

$$\omega_l \approx c\left[k - \alpha(k - k_0)\right], \qquad (9)$$

where $k_0^2 = 2\langle\gamma\rangle\omega_p^2/c^2$ and $\alpha = \langle\gamma\rangle/2\langle\gamma^3\rangle$.

It follows from (6) and (9) that for

$$k \approx \bar{k} \equiv k_0\left(1 + \frac{1}{2}\frac{\omega_p^2}{\omega_B^2}\langle\gamma^3\rangle\right), \qquad (10)$$

the frequency of the Langmuir wave becomes equal to the frequency of the electromagnetic wave. The dispersion curves of the waves propagating along the magnetic field in an electron–positron plasma are shown in Fig. 2.

Mechanisms for the generation of waves in the relativistic electron–positron plasma of the pulsar magnetosphere have been widely discussed in the literature (see, e.g., [17]). We favor the maser mechanism [18]. The magnetosphere plasma is nonequilibrium for two reasons: the presence of a high-energy beam, and the asymmetric and one-dimensional nature of the distribution function. If the Cherenkov resonance condition is fulfilled, the drift Cherenkov resonance wave-generation mechanism is possible when the drift motion of particles in the slightly curved magnetic field is taken into account [6]. The linear stage was studied in detail in [19], and the quasi-linear stage in [20]. Here, we will consider instead another mechanism: resonance at the abnormal Doppler effect, for which the necessary condition is [6]

$$\omega - k_{||}v_{||} - k_\perp u_\perp^m + \frac{\omega_B}{\gamma_r} \approx 0. \qquad (11)$$

Here, $k_{||}$ and $k_\perp$ are the projections of the wave vector along and across the magnetic field, $u_\perp^m \equiv cv_{||}\gamma_r/\omega_B R_B$ is the drift velocity of the particles due to the curvature of the magnetic field, R_B is the radius of curvature of the magnetic field lines, and γ_r is the Lorentz factor of the resonant particles.

A detailed analysis demonstrates that wave generation is possible on fast particles of the tail and beam [21] if $2\gamma_r^2\delta > 1$. Substituting (2) into the resonance condition (11), we find for $k_\perp = 0$

$$\omega_t \approx k_{||}c \approx \frac{\omega_B}{\gamma_r\delta}. \qquad (12)$$

For typical pulsar parameters [22, 23], it follows from (12) that the frequency of the generated t waves is below the cyclotron frequency:

$$\omega_t \ll \omega_B. \qquad (13)$$

It is important to emphasize that the transverse waves are strongly damped at the cyclotron resonance by particles of the plasma bulk and relatively slow particles of the tail. The condition for the cyclotron resonance is

$$\omega - k_{||}v_{||} - k_\perp u_\perp^m - \frac{\omega_B}{\gamma_r} \approx 0. \qquad (14)$$

Analysis of this equation demonstrates that cyclotron damping is efficient when

$$\omega_t \gtrsim 2\gamma_p\omega_B. \qquad (15)$$

The linear mechanisms for the generation of waves via the abnormal Doppler-effect resonance or drift-Cherenkov resonance excite transverse electromagnetic t waves. In turn, nonlinear interactions redistribute the energy of the generated waves among various modes and scales; if certain conditions discussed below are fulfilled, this results in the formation of a stationary turbulence spectrum.

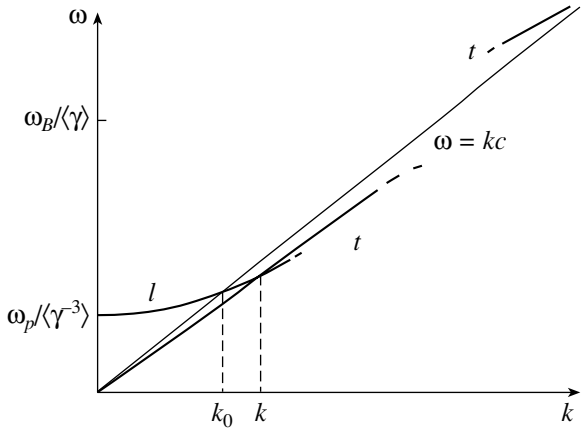


Fig. 2. Dispersion curves of waves propagating along the magnetic field in an electron–positron plasma.

4. NONLINEAR PROCESSES

We will use the weak-turbulence parameter to construct a theory for the turbulent processes occurring in a plasma. For a relativistic electron–positron plasma, this parameter is

$$\frac{|E|^2}{m_e c^2 n_p \gamma_p} \ll 1, \qquad (16)$$

where **E** is the electric field of the wave.

Condition (16) means that we describe the development of turbulence for which the energy density of the waves is less than the mean energy density of the particles.

In addition to the small parameter (16), there are two other small parameters for the magnetosphere plasma, whose existence considerably simplifies the subsequent analysis. The first is $\omega_p^2/\omega_B^2 \ll 1$. The second,

$$\gamma_p^2 \frac{(\omega - k_{||}v_{||})^2}{\omega_B^2} \ll 1, \tag{17}$$

enables us to proceed to the so-called drift approximation.

In common laboratory plasma for which the weak-turbulence condition is fulfilled, three-wave resonant decay processes should be the most intense [25]. There are additional limitations to realizing the resonant decay conditions in an electron–positron plasma. In particular, since the particles in an electron–positron plasma have identical mass and equal and opposite charges, second- and higher-orders terms that are proportional to odd powers of the charge do not contribute to the nonlinear currents, since the contributions from the electrons and positrons completely cancel. In addition, as we noted above, the observer receives waves that propagate within a small angle ($\theta \ll 1$) relative to the magnetic field. It also follows from the resonant condition (11) that the generated waves propagate nearly along the magnetic-field lines [14].

The decay of a potential Langmuir l wave into two transverse t waves was considered in [26], and all possible three-wave decays for nearly field-aligned wave propagation were studied in [27]. As we noted above, these studies focused on possible decays of the l wave, since it was assumed that, as in a common plasma, the most unstable and most readily excited waves are Langmuir waves. However, in a plasma with a strong relativistic beam, the so-called kinetic-instability approximation is violated when the resonance width $\omega - k_{||}v_{||}$ is smaller than the increment. In this case, the number of particles involved in the resonance is small, and the instability does not develop.

The presence of a resonance at the abnormal Doppler effect, which is a powerful source of t waves, changes the situation radically. The question of possible decays of the t waves that may be responsible for the observed spectral parameters of the pulsar radiation becomes much more interesting.

It is convenient to use quantum-mechanical concepts to study the nonlinear processes. We introduce the amplitude a_k^λ and the plasmon occupation number N_k^λ (see, e.g., [11, 24, 25, 28]):

$$\omega_k N_k^\lambda \equiv \omega_k \langle a_k^\lambda a_k^{\lambda *} \rangle = \frac{1}{\omega_k} \left(\frac{\partial}{\partial \omega} \omega^2 \varepsilon_k^\lambda \right)_{\omega = \omega_k^\lambda} \frac{|E_k|^2}{4\pi}. \tag{18}$$

Here,

$$\varepsilon_k^\lambda \equiv \frac{e_i^{\lambda *} e_j^\lambda}{\omega^2} \left(k_i k_j c^2 + \omega^2 \varepsilon_{ij} \right), \tag{19}$$

$\mathbf{e}^\lambda$ is the polarization vector for waves of type $\lambda(lt, l, t)$, $E_j^\lambda \equiv e_j^\lambda E$, ε_{ij} is the permittivity tensor, an asterisk denotes complex conjugation, and angular brackets denote averaging over phases.

It follows from Maxwell's equations that

$$\partial_t a_k + i\omega_k^\lambda = \sum_{\lambda_1 \lambda_2} \int d\mathbf{k}_1 d\mathbf{k}_2 d\omega_1 d\omega_2 a_k^{\lambda_1} a_k^{\lambda_2 *} \times \delta(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2) \delta(\omega - \omega_1 - \omega_2) V_{\lambda|\lambda_1\lambda_2}, \tag{20}$$

where the matrix element of the decay $\lambda \to \lambda_1 + \lambda_2$ is

$$V_{\lambda|\lambda_1\lambda_2} = 4\pi \frac{\omega_{k_1}^{\lambda_1} \omega_k^\lambda}{\omega_k^\lambda} (\omega_{k_1}^{\lambda_1} - \omega_{k_2}^{\lambda_2}) \times (\partial_\omega \omega^2 \varepsilon^\lambda)_{\omega = \omega_k^\lambda}^{-1/2} (\partial_\omega \omega^2 \varepsilon^\lambda)_{\omega = \omega_{k_1}^{\lambda_1}}^{-1/2} \times (\partial_\omega \omega^2 \varepsilon^\lambda)_{\omega = \omega_{k_2}^{\lambda_2}}^{-1/2} \sigma_{\lambda|\lambda_1\lambda_2}, \tag{21}$$

and $\sigma_{\lambda|\lambda_1\lambda_2}$ is the nonlinear conductance tensor.

For identical polarizations of all three waves, the process $t \to t' + t''$ at $k_\perp = 0$ is forbidden, because the second-order current is proportional to the third power of the charge ($j^{(2)} \sim e^3$) and the contributions from electrons and positrons completely cancel. This process is formally possible in the case of propagation at a small angle to the magnetic field, but the matrix element of the interaction [27]

$$V_{t|t't''} \sim \omega_p^3/\omega_B^3 \ll 1 \tag{22}$$

is negligible.

It remains to consider processes involving l waves. In an electron–ion laboratory plasma, the most efficient nonlinear process is the decay of Langmuir waves to Langmuir and ion-acoustic waves [29]. This process is accompanied by a "reddening" of the Langmuir plasmon; i.e., a transfer of energy to longer-wavelength perturbations.

There is no exact counterpart for this process in an electron–positron plasma. The process $l \to l' + t$,

where we have a transverse t wave instead of a longitudinal potential ion–acoustic wave, can be considered to be the most analogous. In the case of strictly field-aligned propagation, the electric vectors of the l waves are directed along the magnetic field ($\mathbf{E}^l||\mathbf{E}^{l'}||\mathbf{B}_0$), whereas the electric field of the t wave is $\mathbf{E}^t \perp \mathbf{B}_0$. Therefore, such a decay is impossible in this case. If the wave propagates at a small angle θ to the magnetic field, the matrix element is [27]

$$V_{l|l't} \approx \frac{e}{m_e c}\frac{\omega_p^2}{\omega_B \omega_l}\sqrt{\frac{\omega_l}{\omega_{l'}}}\frac{k_l c}{\sqrt{\omega_t}}\theta. \qquad (23)$$

Since the maximum value of θ does not exceed δ, we conclude that the matrix element of this process is also quite small.

Another possible three-wave process, $l + l' \to t$, was considered in [30]; the matrix element of this process is also described by (23), and, accordingly, is small.

The decay $l \to t + t'$, considered in [26], requires the propagation of t waves in opposite directions. Otherwise, it is impossible to satisfy the resonant-decay conditions

$$\omega_l = \omega_t + \omega_{t'}, \quad \mathbf{k}_l = \mathbf{k}_t + \mathbf{k}_{t'}. \qquad (24)$$

Naturally, this process cannot play any role in the formation of the pulsar radio emission at frequencies exceeding that of the Langmuir waves.

The process $t \to t' + l$ without including the drift motion of plasma particles was considered in [27]. In this case, the interaction involves near-light-speed Langmuir waves with $\omega_l \approx k_l c$. For field-aligned propagation, the matrix element of the interaction is

$$V_{t|lt'} \approx \frac{e}{m_e c}\frac{\omega_p^2}{\omega_B \omega_l}\sqrt{\frac{\omega_t \omega_{t'}}{\omega_l}}. \qquad (25)$$

Comparison of (22), (23), and (25) demonstrates that, even without the drift motion of the particles, the process $t \to t' + l$ is the dominant three-wave process. Taking into account the drift motion of plasma particles makes this process even more efficient, due to the presence of the strong external magnetic field $\mathbf{B}_0$.

For simplicity of presentation, we will consider the case of strictly field-aligned wave propagation. For t waves, the electric vector is $\mathbf{E}^t \perp \mathbf{B}_0$. Therefore, drift motion along $\mathbf{B}_0$ is possible due to oscillations of the t and t' waves. For concreteness, we will take the fields $\mathbf{E}^t$ and $\mathbf{E}^{t'}$ to be directed along the x and y axes, respectively (Fig. 3). The magnetic field $\mathbf{B}_0$ is directed along the z axis. In this case, E_x^t leads to drift motion of the particles along y:

$$u_y^d = c\frac{E_x^t}{B_0}, \qquad (26)$$

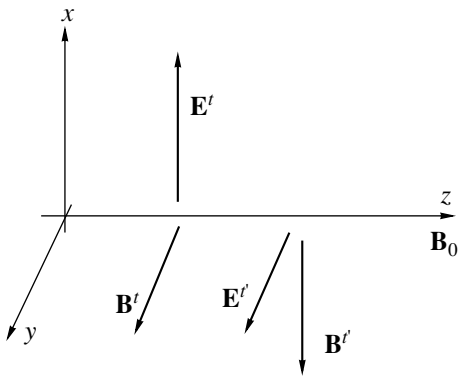


Fig. 3. Scheme for the calculation of nonlinear drift currents.

which, in turn, crossing the magnetic field of the t' wave, generates a nonlinear longitudinal field:

$$\mathbf{E} = \frac{1}{c}\left[\mathbf{u}^d, \mathbf{B}'\right]. \qquad (27)$$

We especially emphasize that the electrical drift velocity (26) is the same for electrons and positrons. Therefore, the drift of the particles does not result in the generation of a current in the linear approximation.

If the resonant conditions

$$\omega_t = \omega_l + \omega_{t'}, \quad \mathbf{k}_t = \mathbf{k}_l + \mathbf{k}_{t'} \qquad (28)$$

are fulfilled, the beating of the t and t' waves generating the longitudinal electrical field (27) resonates with the slightly sub-light-speed Langmuir oscillations.

Since the Langmuir waves are longitudinal, in the presence of particles with velocities equal to the phase velocity of the wave, they can be subject to collisionless Landau damping. In this case, instead of the process $t \to t' + l$, we should consider nonlinear scattering of t waves by particles of the plasma—the so-called wave–particle–wave interaction [24, 25]. The condition for this resonance is

$$\omega_t - \omega_{t'} = (k_{t||} - k_{t'||})v_{||}, \qquad (29)$$

where $v_{||}$ is the longitudinal velocity of the resonant particle.

Combining (6) and (29) and using the fact that $v_{||} \approx c[1 - 1/(2\gamma^2)]$, we readily find that the resonance can take place for particles with γ factors

$$\gamma_r \approx \sqrt{1/2\delta}. \qquad (30)$$

According to the theory developed in [6, 19, 22], the waves are generated near the light cylinder, at distances of the order of 10^9 cm from the surface of the star. Consequently,

$$\gamma_r \sim 10^2 - 10^3. \qquad (31)$$

Scattering could be rather efficient if the Langmuir waves satisfying conditions (28) are strongly damped. The Langmuir waves could be damped on particles of the bulk plasma. However, due to the violation of the applicability of the kinetic approximation in a relativistic plasma, efficient collisionless damping is hindered (though possible). Here, we will consider the case when the Langmuir waves are weakly damped, so that the dominant process is the decay $t \to t' + l$.

Maxwell's equations describing the evolution of the fields in the case of exactly field-aligned wave propagation can be reduced in the standard manner to

$$\left(\frac{\partial^2}{\partial t^2} - c^2\frac{\partial^2}{\partial z^2}\right)\mathbf{E}_\perp + 4\pi\frac{\partial \mathbf{j}_\perp^{NL}}{\partial t} = 0, \qquad (32)$$

$$\frac{\partial E_z}{\partial t} + 4\pi j_z^{NL} = 0. \qquad (33)$$

The nonlinear currents $\mathbf{j}_\perp^{NL}$ and j_z^{NL} in (32), (33) can readily be calculated in a simplified scheme; substituting (26) to (27) and using Faraday's law, we obtain

$$j_z^{NL} = -\frac{i}{4\pi}\left(\frac{k_z^t}{\omega^t}E_x E'_y + \frac{k_z^{t'}}{\omega^{t'}}E'_x E_y\right). \qquad (34)$$

The transverse current can be computed as follows:

$$\mathbf{j}_\perp^{NL} = \mathbf{u}_\perp^d \rho_e, \qquad (35)$$

where the perturbation of the charge density ρ_e is determined from the Poisson equation $\mathrm{div}\mathbf{E} = 4\pi\rho_e$, and E_z is determined from (26).

Substituting into (32) and (33) electric fields in the form

$$E_{z,\perp}(t) = E_{z,\perp k}(t)e^{i(\omega t - kz)} + \text{c.c.}, \qquad (36)$$

and using (32)–(35) together with the linear dispersion equations (6) and (9), we obtain for the amplitudes slowly varying with time the equations

$$\frac{\partial E_{xk_t}}{\partial t} = \frac{i}{2}\frac{c\omega_l}{B_0}\left(\frac{k^{t'}}{\omega^{t'}} + \frac{k^l}{\omega^l}\right)E_{zk_l}E^*_{yk_{t'}}, \qquad (37)$$

$$\frac{\partial E_{yk_{t'}}}{\partial t} = -\frac{i}{2}\frac{c\omega_l}{B_0}\left(\frac{k^t}{\omega^t} + \frac{k^l}{\omega^l}\right)E_{zk_l}E^*_{xk_t}, \qquad (38)$$

$$\frac{\partial E_{zk_l}}{\partial t} = -\frac{i}{2}\frac{c\omega_l}{\langle\gamma\rangle B_0}\left(\frac{k^t}{\omega^t} + \frac{k^{t'}}{\omega^{t'}}\right)E_{xk_t}E^*_{yk_{t'}}. \qquad (39)$$

Supposing that there are many waves with random phases, and introducing the occupation numbers

$$N_{x,yk} = \frac{|E_{x,yk}|^2}{4\pi\omega_t}, \quad N_{zk} = \frac{\langle\gamma^2\rangle|E_{zk}|^2}{8\pi\omega_l}, \qquad (40)$$

in the standard manner [24, 25], we obtain from (37)–(39) the matrix element for the interaction

$$V_{t|lt'} \approx \frac{c}{B_0}(\omega_t k_{t'} + \omega_{t'}k_t)\sqrt{\frac{\omega_l}{\omega_t\omega_{t'}\langle\gamma^2\rangle}}. \qquad (41)$$

Comparison of (25) and (41) demonstrates that taking into account the drift of the particles increases the matrix element, i.e., intensifies the considered process.

5. STATIONARY SPECTRA

Typical of the case we have considered is the situation when $\omega_t, \omega'_t \gg \omega_l$, and, consequently, the step for the transfer of the high frequency t waves is small. In addition, we assume that the spectrum forms after the escape of the radiation from the linear-generation region; i.e., there is no pumping in the energy redistribution region. As we noted above, damping of waves with frequencies below $2\gamma_p\omega_B$ can also be neglected. Thus, we come to the classical inertial interval, in which a Kolmogorov-type spectrum is formed.

According to [10], the spectra formed in the Kolmogorov (inertial) interval, where there are no sources or sinks of waves, are completely determined by the matrix element for the interaction (41) and the dispersion characteristics of interacting waves, which are described by (6) and (9). This problem admits two stationary solutions for the occupation numbers N_k, which have the form [10]

$$N_{k1} = A_1 k^{-s-d}, \quad N_{k2} = A_2 k^{-s-d+1/2}, \qquad (42)$$

where d is the dimension of the problem and s is the index for the degree of homogeneity of the matrix element of the interaction. The first stationary solution, $N_{k1} = A_1 k^{-5/2}$, corresponds to a constant energy flux from large-scale to small-scale irregularities, while the second solution, $N_{k1} = A_2 k^{-2}$, corresponds to a constant flux of the plasmon number toward large-scale perturbations.

In the case considered, $d = 1$ and $s = 3/2$; therefore, possible stationary spectra are $N_{k1} = A_1 k^{-5/2}$ and $N_{k2} = A_2 k^{-2}$. Taking into account the linear character of the dispersion of the t waves and the quasi-one-dimensional character of the problem, we obtain for the energy spectrum $\mathcal{E}_\omega$, which is proportional to the observed emission intensity I_ω:

$$I_{\omega 1} \sim \mathcal{E}_{\omega 1} \sim \omega^{-3/2}, \quad I_{\omega 2} \sim \mathcal{E}_{\omega 2} \sim \omega^{-1}. \qquad (43)$$

Which of the spectra is realized in a particular pulsar depends on the linear mechanisms for the wave excitation and absorption. If, as is supposed in Section 3, waves are excited via a resonance at the abnormal Doppler effect, which is especially efficient

when $\omega_t \ll \omega_B$, while damping at the cyclotron resonance is efficient when $\omega_t \gtrsim 2\gamma_p\omega_B$, the spectrum $\mathcal{E}_{\omega 1}$ should be realized.

As we noted above, the radio emission of most pulsars has spectral indices $-1.5 > \alpha > -2$. Therefore, our result, $\alpha = -1.5$, is quite satisfactory. However, we should bear in mind the important restrictions we have imposed to simplify the analysis, in particular, the assumption that there exists an inertial interval and the requirement that the spectrum be stationary.

6. CONCLUSION

We have considered the formation of spectra in the weakly turbulent electron–positron plasma of a pulsar magnetosphere. All available data about pulsars confirm the presence of a very strong magnetic field, $|B_0| \sim 10^{12}$ G, at the stellar surface. This compels us to consider the theory of plasma turbulence for strong magnetic fields, where, in our opinion, the determining role should be played by the electrical drift of the particles, which results in the efficient nonlinear generation of a longitudinal electrical field during interactions between two t waves. The presence of a strong external magnetic field simplifies the problem and makes it quasi-one-dimensional. Applying well-known methods of the theory of weak turbulence, we have calculated the spectral indices of stationary spectra that can be formed in the inertial interval. Analysis of the linear mechanisms for wave generation and damping shows that a situation with wave generation resulting from resonance at the abnormal Doppler effect and damping taking place at the cyclotron resonance is most probable. In this case, a spectrum with an index of -1.5 should be realized. This result is in satisfactory agreement with the observations: the observed spectral indices for most pulsars are $-1.5 > \alpha > -2$.

REFERENCES

1. V. M. Malofeev, Astron. Soc. Pac. Conf. Ser. **105**, 271 (1996).
2. O. Maron, J. Kijak, M. Kramer, and R. Wielebinski, Astron. Astrophys. **147**, 195 (2000).
3. D. R. Lorimer, J. A. Yates, A. G. Lyne, and M. D. Gould, Mon. Not. R. Astron. Soc. **273**, 411 (1995).
4. V. A. Izvekova, A. D. Kuzmin, V. M. Malofeev, and Y. Shitov, Astrophys. Space Sci. **78**, 45 (1981).
5. A. G. Lyne, F. G. Smith, and D. A. Graham, Mon. Not. R. Astron. Soc. **153**, 337 (1971).
6. A. Z. Kazbegi, G. Z. Machabeli, and G. I. Melikidze, Mon. Not. R. Astron. Soc. **253**, 377 (1991).
7. A. Kazbegi, G. Machabeli, G. Melikidze, and C. Shukre, Astron. Astrophys. **309**, 515 (1996).
8. G. Machabeli, D. Khechinashvili, G. Melikidze, and D. Shapakidze, Mon. Not. R. Astron. Soc. **327**, 984 (2001).
9. I. F. Malov, V. M. Malofeev, G. Machabeli, and G. Melikidze, Astron. Zh. **74**, 303 (1997) [Astron. Rep. **41**, 262 (1997)].
10. V. E. Zakharov and E. A. Kuznetsov, Zh. Éksp. Teor. Fiz. **75**, 904 (1978) [Sov. Phys. JETP **48**, 458 (1978)].
11. V. E. Zakharov, V. S. L'vov, and G. Falkovich, *Kolmogorov Spectra of Turbulence I* (Springer-Verlag, Berlin, 1992).
12. P. Goldreich and W. H. Julian, Astrophys. J. **157**, 869 (1969).
13. P. A. Sturrock, Astrophys. J. **164**, 529 (1971).
14. A. S. Volokitin, V. V. Krasnosel'skikh, and G. Z. Machabeli, Fiz. Plazmy **11**, 531 (1985) [Sov. J. Plasma Phys. **11**, 310 (1985)].
15. V. P. Silin, Zh. Éksp. Teor. Fiz. **38**, 1577 (1960) [Sov. Phys. JETP **11**, 1136 (1960)].
16. V. N. Tsytovich, Zh. Éksp. Teor. Fiz. **40**, 1775 (1961) [Sov. Phys. JETP **13**, 1249 (1961)].
17. D. B. Melrose, *Instabilities in Space and Laboratory Plasmas* (Cambridge Univ. Press, Cambridge, 1986).
18. V. L. Ginzburg and V. V. Zheleznyakov, Ann. Rev. Astron. Astrophys. **13**, 511 (1975).
19. M. Lutikov, G. Machabeli, and R. Blandford, Astrophys. J. **512**, 804 (1999).
20. D. Shapakidze, G. Machabeli, G. Melikidze, and D. Khechinashvili, Phys. Rev. E **67**, 026407 (2003).
21. Dzh. Lominadze, A. V. Mikhailovskii, and G. Z. Machabeli, Fiz. Plazmy **5**, 1337 (1979) [Sov. J. Plasma Phys. **5**, 748 (1979)].
22. G. Z. Machabeli and V. V. Usov, Pis'ma Astron. Zh. **5**, 445 (1979) [Sov. Astron. Lett. **5**, 238 (1979)].
23. G. Z. Machabeli and V. V. Usov, Pis'ma Astron. Zh. **15**, 910 (1989) [Sov. Astron. Lett. **15**, 393 (1989)].
24. V. N. Tsitovich, *Nonlinear Effects in Plasma* (Plenum, New York, 1970).
25. R. Z. Sagdeev and A. A. Galeev, *Nonlinear Plasma Theory* (Benjamin, New York, 1969).
26. A. V. Mikhailovskii, Fiz. Plazmy **6**, 613 (1980) [Sov. J. Plasma Phys. **6**, 336 (1980)].
27. M. E. Gedalin and G. Z. Machabeli, Fiz. Plazmy **9**, 1015 (1983) [Sov. J. Plasma Phys. **9**, 592 (1983)].
28. R. C. Davidson, *Methods in Nonlinear Plasma Theory* (Academic, New York, 1972).
29. V. N. Oraevskii and R. Z. Sagdeev, Zh. Tekh. Fiz. **32**, 1291 (1962) [Sov. Phys. Tech. Phys. **7**, 955 (1963)].
30. P. G. Mamradze, G. Z. Machabeli, and G. I. Melikidze, Fiz. Plazmy **6**, 1293 (1980) [Sov. J. Plasma Phys. **6**, 707 (1980)].

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 470–476. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 527–534.

Optimal Choice of the Softening Length and Time Step in N-body Simulations

S. A. Rodionov and N. Ya. Sotnikova

Sobolev Astronomical Institute, St. Petersburg State University, Universitetskiĭ pr. 28, Petrodvorets, St. Petersburg, 198904 Russia

Received July 20, 2004; in final form, December 3, 2004

Abstract—A criterion for the choice of optimal softening length ϵ for the potential and the choice of time step dt for N-body simulations of a collisionless stellar system is analyzed. Plummer and Hernquist spheres are used as models to analyze how changes in various parameters of an initially equilibrium stable model depend on ϵ and dt. These dependences are used to derive a criterion for choosing ϵ and dt. The resulting criterion is compared to Merritt's criterion for choosing the softening length, which is based on minimizing the mean irregular force acting on a particle with unit mass. Our criterion for choosing ϵ and dt indicate that ϵ must be a factor of 1.5–2 smaller than the mean distance between particles in the densest regions to be resolved. The time step must always be adjusted to the chosen ϵ (the particle must, on average, travel a distance smaller than $0.5\,\epsilon$ during one time step). An algorithm for solving N-body problems with adaptive variations of the softening length is discussed in connection with the task of choosing ϵ, but is found not to be promising. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The evolution of gravitating systems on time intervals shorter than the time scale for two-body relaxation is described by the collisionless Boltzmann equation. The number of particles in N-body simulations of such systems is usually several orders of magnitude smaller than the number of stars in real systems. This approach usually involves solving the collisionless Boltzmann equation using Monte Carlo methods, and then using particles in the N-body simulations to produce a representation of the density distribution in the system. In practice, integrating the equations of motion of the particles involves smoothing the pointwise potential of each individual particle, e.g., by substituting a Plummer sphere potential for the actual potential. This procedure modifies the law governing the interactions between particles at small distances:

$$\mathbf{F}_{ij}^{\rm real} = Gm_i m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3} \qquad (1)$$

$$\rightarrow \mathbf{F}_{ij}^{\rm soft} = Gm_i m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{(|\mathbf{r}_j - \mathbf{r}_i|^2 + \epsilon^2)^{3/2}},$$

where $\mathbf{F}_{ij}^{\rm real}$ and $\mathbf{F}_{ij}^{\rm soft}$ are the real and softened forces, respectively, on a particle of mass m_i located at point $\mathbf{r}_i$ produced by a particle of mass m_j located at point $\mathbf{r}_j$, and ϵ is the softening length for the potential. Other methods of the softening length for the potential are possible.

Potential smoothing is used in N-body simulations for two reasons.

First, attempts to solve the equations of motion of gravitating points using purely Newtonian potentials and simple constant-step integration always lead to problems during close pair encounters (the integration scheme diverges). Correct modeling of such systems requires the use of either variable-step integrating algorithms or sophisticated regularization methods, which lead to unreasonably long CPU times. No such problems arise when a softened potential is used.

Second, smoothing the potential reduces the "graininess" of the particle distribution, thereby making the potential of the model system more similar to that of a system with a smooth density distribution, i.e., a system described by the collisionless Boltzmann equation.

It is obvious that ϵ cannot be too large: this would result in substantial distortion of the potential, and would also impose strong constraints on the spatial resolution of structural features of the system. It is important to have an objective criterion for choosing the softening length in N-body simulations. In this paper, we derive such a criterion by analyzing the time variations of the distribution functions for spherically symmetric models with different values of ϵ.

2. MERRITT'S CRITERION FOR CHOOSING ϵ

Merritt [1] proposed a criterion for choosing the softening length based on the minimization of the

1063-7729/05/4906-0470$26.00

mean irregular force acting on a particle. He introduced the mean integrated square error (MISE), which characterizes the difference between the force produced by a discrete set of N bodies and the force acting in a system with a continuous density distribution $\rho(\mathbf{r})$:

$$MISE(\epsilon) = E\left(\int \rho(\mathbf{r})|\mathbf{F}(\mathbf{r},\epsilon) - \mathbf{F}_{\rm true}(\mathbf{r})|^2 d\mathbf{r}\right). \quad (2)$$

Here, E denotes averaging over many realizations of the system, $\mathbf{F}(\mathbf{r},\epsilon)$ is the force acting on a particle with unit mass located at the point $\mathbf{r}$ produced by N particles with softened potentials, and $\mathbf{F}_{\rm true}(\mathbf{r})$ is the force acting on the same particle in a system with a continuous density distribution $\rho(\mathbf{r})$.

Two factors contribute to the MISE.

(1) Fluctuations of the discrete density distribution. These are very important at small ϵ, decrease with ϵ (see the left-hand branch of the curve in Fig. 1) and for a given ϵ, decrease with increasing N [1].

(2) Errors in the representation of the potential (the difference between the softened potential and a point-mass potential). These errors, on the contrary, are very large for large ϵ, decrease with decreasing ϵ, and do not depend on N (the right-hand branch of the curve in Fig. 1).

The MISE reaches its minimum at some $\epsilon = \epsilon_m$ (Fig. 1). Merritt suggested that this ϵ_m be adopted as the optimal choice of the softening length in N-body simulations. This quantity depends on N and, e.g., in the case of a Plummer model, can be approximated by the dependence $\epsilon_m \approx 0.58\,N^{-0.26}$ in the interval $N = 30$–$300\,000$ (in virial units, where $G = 1$, the total mass of the model is $M_{\rm tot} = 1$, and the total energy of the system is $E_{\rm tot} = -1/4$) [2].

3. FORMULATION OF THE PROBLEM

Merritt's criterion is based on minimizing the mean error in the representation of the force in an equilibrium system at the initial time. It is of interest to derive a criterion for choosing the softening length based directly on the dynamics of the system. If a stable equilibrium configuration is described by a discrete set of N bodies, due to various errors, the parameters of the system (e.g., the effective radius) will deviate from their initial values with time. The smaller these deviations are, the more adequately the model reproduces the dynamics of the system. It is reasonable to suppose that these variations will be minimized for some ϵ.

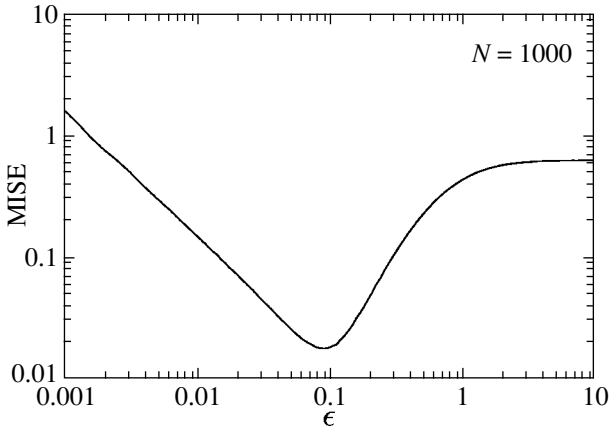


Fig. 1. Dependence of the MISE on the softening length ϵ for a Plummer model with $N = 1000$.

We analyzed two equilibrium models: a Plummer-sphere model,

$$\rho(r) = \frac{3M_{\rm tot}}{4\pi}\frac{a_{\rm P}^2}{\left(r^2 + a_{\rm P}^2\right)^{5/2}}, \quad (3)$$

as an example of a model with a nearly uniform density distribution, and a Hernquist-sphere model,

$$\rho(r) = \frac{M_{\rm tot}}{2\pi}\frac{a_{\rm H}}{r(r + a_{\rm H})^3}, \quad (4)$$

with an isotropic velocity distribution [3], as an example of a model with a strongly nonuniform density distribution. In (3) and (4), $a_{\rm P}$ and $a_{\rm H}$ are the scale lengths of the density distributions for the Plummer and Hernquist models (the Plummer and Hernquist models contain about 35 and 25% of their total masses inside the radii $a_{\rm P}$ and $a_{\rm H}$, respectively). We use here virial units ($G = 1$, $M_{\rm tot} = 1$, $E_{\rm tot} = -1/4$), for which $a_{\rm P} = 3\pi/16 \approx 0.59$ and $a_{\rm H} = 1/3$. We computed the gravitation force using the TREE method [4]. In the computations, we set the parameter θ, which is responsible for the accuracy of the computation of the force in the adopted algorithm, equal to 0.75. Test simulations made with higher ($\theta = 0.5$) and lower ($\theta = 1.0$) accuracy showed that our conclusions are independent of θ within the indicated interval. Hernquist *et al.* [5] showed that the errors in the representation of the force in the TREE method are smaller than the errors due to smoothing of the potential, so that the above approach can be used to solve the formulated problem. We integrated the equations of motion of the particles using a standard leap-frog scheme with second order of accuracy in time step. The number of particles was $N = 10\,000$, and the computations were carried out in the NEMO [6] software package.

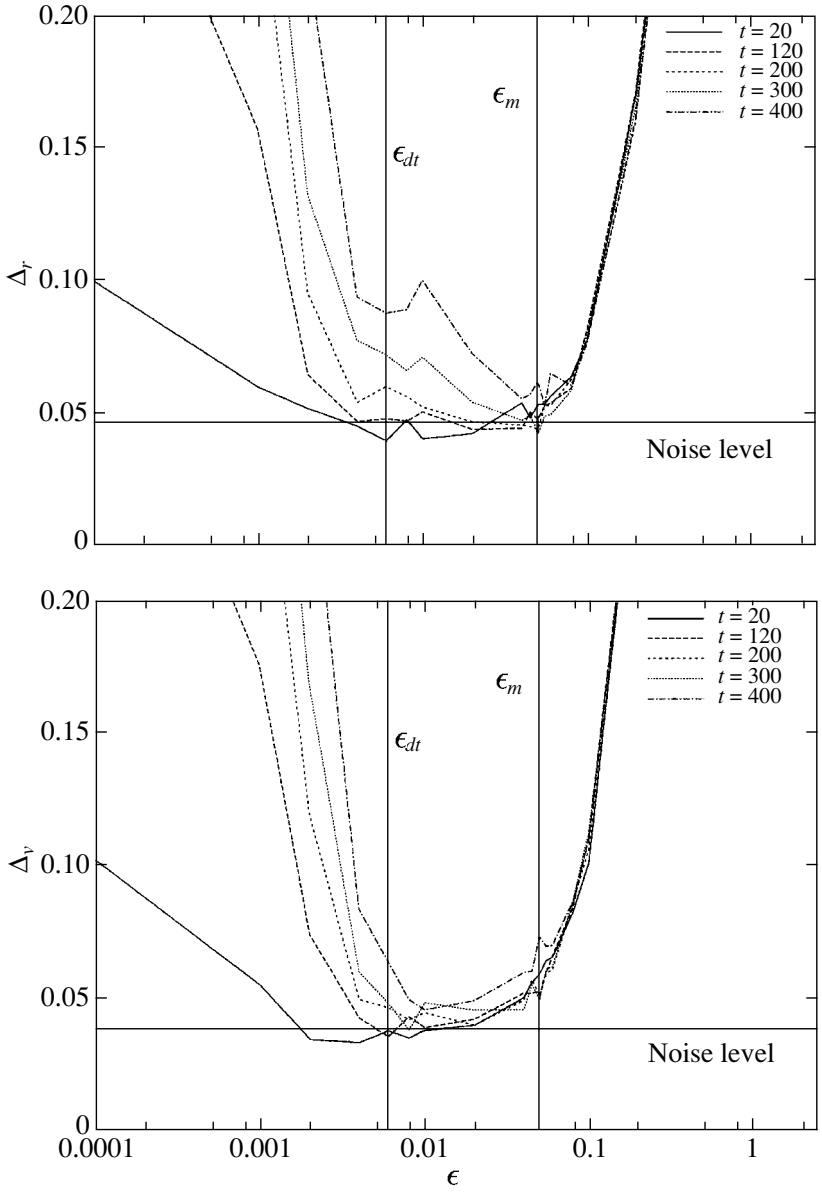


Fig. 2. Dependence of Δ_r and Δ_v on ϵ for different times for the Plummer-sphere model. The horizontal line shows the natural noise levels for Δ_r and Δ_v. The vertical lines indicate the position of ϵ_m—the optimal softening length according to the criterion of Merritt (for $N = 10\,000$, $\epsilon_m = 0.05$ in virial units), and the minimum ϵ_{dt} to which the time step was adjusted ($\epsilon_{dt} = 0.006$). Averaging was performed over several models for each ϵ. The time step is $dt = 0.01$.

We investigated two methods for smoothing the potential. The first is based on substituting the Plummer potential for the point-mass potential, and modifies the force of interaction between the particles according to the scheme (1). The second method uses a cubic spline [7] to smooth the potential. We report here results only for the first method, but all our conclusions remain unchanged in the case of spline smoothing.

We analyzed how various global parameters of the system are conserved, in particular, the variation of the particle distribution function in space and, separately, the variation of the particle-velocity distribution.

We characterized the deviation of the particle distribution in space at time t from the initial distribution using the quantity Δ_r, which was computed as follows. We subdivided the model into spherical layers and computed for each layer the difference between the number of particles at time t and the number at the initial time. We then computed Δ_r as the sum of the absolute values of these differences and normalized

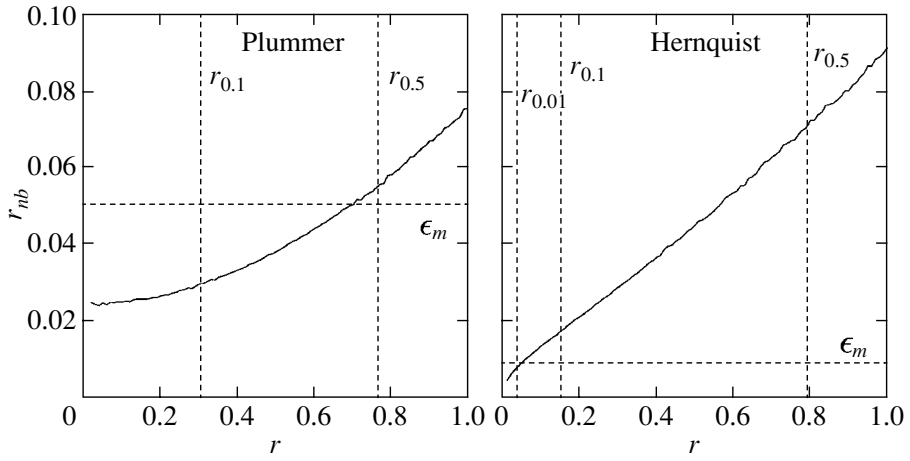


Fig. 3. Average distance to the nearest neighbor at various radii for the Plummer and Hernquist models with $N = 10\,000$. The vertical lines correspond to the radii of spheres containing 0.01, 0.1, and 0.5 of the total mass of the system. The horizontal lines shows the optimal ϵ_m according to the criterion of Merritt.

it to N, the total number of particles in the system. The thickness of the layers was 0.1 and the maximum radius was equal to 2 (each of the spherical layers in both models contained a statistically significant number of particles, always greater than 50). Note the following (very important) point: Δ_r will not be equal to zero for two random realizations of the same model. This quantity has appreciable natural noise due to the finite number of particles considered. We estimated the level of this natural noise to be the mean Δ_r averaged over a large number of pairs of random realizations of the model.

We computed the parameter Δ_v for the distribution of particles in velocity space in a similar way. For the results reported here, the thickness of the layer was 0.1 and the maximum velocity was 1.5 (this choice of maximum velocity ensured that each spherical layer contained a statistically significant number of particles for both models—more than 30).

We also computed the two-body relaxation time for the models as a function of ϵ. We determined this timescale as the time over which the particles deviate significantly from their initial radial orbits (the deviation criterion and method used to estimate it were similar to those employed by Athanossoula *et al.* [8]). The estimation method can be briefly described as follows. We constructed a random realization of the system (the Plummer or Hernquist sphere). In this system, we chose a particle moving toward the center in an almost radial orbit and located about one effective radius from the center. The orbit of this particle was adjusted to a strictly radial one. We followed the evolution of the entire system until this particle traveled a distance equal to 1.5 times its initial distance from the center. Knowing the time t_p that has elapsed since the start of the motion and the angle of deflection of the particle from its initial orbit, α_p, we can compute the time required for the particle to deviate from its initial direction by one radian. We then average the resulting time over a large number of similar particles to estimate the two-body relaxation time

$$\frac{1}{t_\mathrm{relax}} = \left\langle \frac{\alpha_\mathrm{p}}{\sqrt{t_\mathrm{p}}} \right\rangle^2, \qquad (5)$$

where t_relax is the two-body relaxation time and $\langle \ldots \rangle$ denotes averaging over a large number of test particles. Formula (5) is valid in the diffusion approximation, i.e., assuming small angles for individual scattering acts.

4. RESULTS OF NUMERICAL SIMULATIONS

4.1. The Plummer Sphere

Figures 2–4 illustrate the results of the computations for the Plummer sphere. Our main conclusions are the following.

(1) Deviations of the softened potential from the Newtonian potential become important when $\epsilon > \epsilon_m$. The system adjusts to the new potential by changing its distribution function. The adjustment is almost immediate, and occurs on time scales $t \ll 20$ (one time unit corresponds approximately to the crossing time for the core, a_P). Such a strongly smoothed system that has evolved far from its initial state shows hardly any further changes (Fig. 2). This result agrees well with the behavior of the fractional error of the total energy of the system for large ϵ, discussed by Athanossoula *et al.* [2, Fig. 7].

(2) Close pair encounters are computed incorrectly when the particle travels a distance greater than ϵ

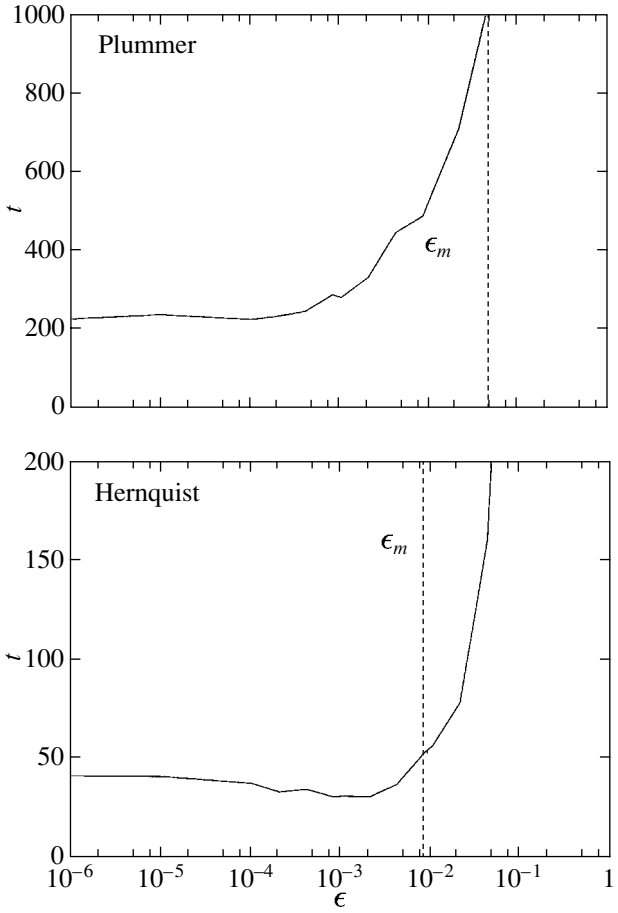


Fig. 4. Relaxation times for the Plummer and Herquist models as functions of ϵ. The vertical line indicates the optimum ϵ_m according to the criterion of Merritt. The time step was $dt = 0.001$. Averaging was performed over 2500 and 1000 test particles for each ϵ for the Plummer and Hernquist models, respectively.

during one time step dt. This introduces an error, which rapidly accumulates and results in substantial changes in Δ_r and Δ_v (Fig. 2). This happens when $\epsilon < \epsilon_{dt} = \sigma_v dt$, where σ_v is the mean velocity dispersion in the system ($\sigma_v \approx 0.7$ in the adopted virial units for the Plummer model).

(3) The evolution of the system outside the interval $\epsilon_{dt} < \epsilon < \epsilon_m$ is simulated incorrectly.

(4) A comparison of the relaxation time of the system (Fig. 4) with the time scale for the variations of the density distribution in the system (Δ_r) shows that, in the interval $\epsilon_{dt} < \epsilon < \epsilon_m$, Δ_r varies only insignificantly on time intervals a factor of two to three shorter than the relaxation time. When $\epsilon = \epsilon_m$, the system preserves its structure for the longest time, since the relaxation time is maximum for this ϵ (upper panel in Fig. 2).

(5) At the same time, it is clear from Fig. 3 that, for the model considered, ϵ_m exceeds the mean distance between particles in the central regions, so that we expect strong modification of the potential in a substantial fraction of the system at $\epsilon = \epsilon_m$. As they adjust to the changed potential, the particles rapidly change their dynamical characteristics (Δ_v; lower panel in Fig. 2). It follows that ϵ_m is by no means the best choice from the viewpoint of conserving the initial velocity distribution function, although Δ_v remains approximately constant after the adjustment has ended.

(6) The effect of the two factors affecting the MISE (the "graininess" of the potential and its modification as a result of smoothing) are separated in time. The first factor acts on time scales comparable to the relaxation time (for a given ϵ), whereas systematic errors in the representation of the potential appear almost immediately. It follows that if the systematic errors are small for the chosen ϵ, then the N-body model represents the initial system very accurately, but only over a time interval that is a factor of 2 or 3 shorter than the relaxation time for this ϵ.

(7) Choosing ϵ to be a factor of 1.5–2 smaller than the mean distance between particles in the densest region yields the optimal solution from the viewpoint of conserving both Δ_r and Δ_v. For the Plummer sphere with $N = 10\,000$, this corresponds to $\epsilon \approx 0.01-0.02$ (in virial units). We can see from Fig. 2 that the simulation of the system is almost optimal when $\epsilon \approx 0.01-0.02$, but this is achieved by restricting the evolution time of the system ($t \leq 100$). When $\epsilon = \epsilon_m \approx 0.05$, the system is simulated satisfactorily on time scales $t \leq 300$, but we must pay for this increased evolution time: the system is not correctly represented by the model from the very start.

4.2. The Hernquist Sphere

Figures 3–5 illustrate the computation results for the Hernquist-sphere model. Our main conclusions are the following.

(1) We can see from Fig. 3 that, with the Hernquist-sphere model, ϵ_m is comparable to the mean distance between particles in the very central regions of the model (which contain less than 1% of all particles). This means that the MISE can be affected even by a small number of particles with strongly modified potentials.

(2) Figure 5 shows that the change in the particle distributions in the configuration and velocity space (Δ_r and Δ_v) are indistinguishable from the noise even for $\epsilon > \epsilon_m$, right to $\epsilon \approx 0.02$. This is explained by the fact that, when $\epsilon = 0.02$, only 10% of the particles have ϵ values greater than the mean distance between particles, and only this small number of particles have

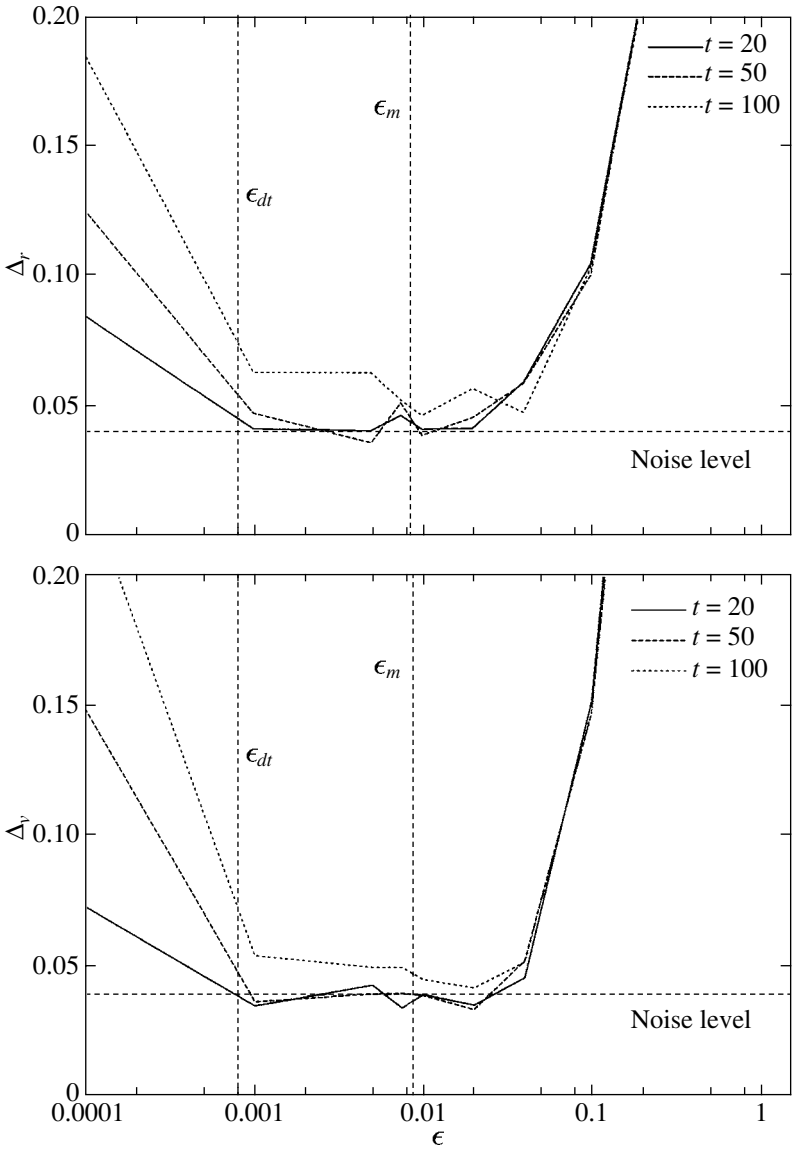


Fig. 5. Same as Fig. 2 for the Hernquist-sphere potential. The time step is $dt = 0.001$. In this case, we have for $N = 10\,000$ and this time step: $\epsilon_m = 0.0087$ $\epsilon_{dt} = 0.0008$.

strongly modified potentials. We simply do not observe the adjustment of these central regions to the new potential.

(3) We showed in the previous section that, for nearly uniform models, ϵ should be about a factor of 1.5–2 smaller than the mean distance between particles in the densest regions. This criterion cannot be directly applied to significantly nonuniform models, since the central density peaks would be washed out due to the finite number of particles. For such models, ϵ must be about a factor of 1.5–2 smaller than the mean distance between particles in the regions to be resolved. In the case of a Hernquist sphere with $N = 10\,000$, the N-body model with $\epsilon = 0.01$ adequately represents the system everywhere except for the most central regions, which contain 2–3% of all the particles (Fig. 3). When $\epsilon = 0.02$, even the regions containing 10% of all particles are not simulated correctly.

(4) Unlike the Plummer sphere, in the case of the Hernquist sphere, significant changes in Δ_r and Δ_v occur on times of the order of two "relaxation times" $t = 100$ (Figs. 4 and 5). We put the term relaxation time in quotes here, because this concept is not applicable in the Hernquist model. The relaxation time is, by definition, a local parameter, and has different values in the central region and periphery for the strongly nonuniform systems we analyze here.

5. DO N-BODY SIMULATIONS NEED ADAPTIVE CODE?

A number of authors have discussed the possibility of developing a code for N-body simulations with an adaptively adjustable softening length; i.e., with ϵ varying as a function of the mean distance between particles at a given point of the system [1, 9]. Dehnen [9] showed, based on the formalism of Merritt [1], that the use of such an adjustable softening length significantly decreases the role of the irregular forces averaged over the entire system. It follows from the results reported here that the only possible advantage of such a code might be a reduction of computing time when modeling very nonuniform models. This follows from the fact that one can choose larger softening lengths, and thereby longer time steps, in less dense regions.

We encountered a problem when implementing such a code, however. The change of ϵ for individual particle results in an asymmetric change of the total energy of the system. When ϵ is increased ($\epsilon_{\rm small} \to \epsilon_{\rm mean}$), particles located near a radius of about $2\epsilon_{\rm mean}$ pass from a "harder" to a "softer" potential, i.e., the absolute value of the potential energy decreases. A decrease in ϵ ($\epsilon_{\rm big} \to \epsilon_{\rm mean}$) results in the opposite effect, but, in this case, on average, the change in the potential influences the greater number of particles located near the radius $2\epsilon_{\rm big}$, so that the absolute value of the total potential energy of the system gradually increases with time. Our N-body simulations showed significant increases in the total energy. Various artificial methods can be suggested to compensate for this energy change, but it remains unclear how this will affect the evolution of the system as a whole.

6. CONCLUSIONS

The results of our study indicate that the softening length ϵ in N-body simulations should be a factor of 1.5–2 smaller than the mean distance between particles in the densest regions to be resolved. In this case, the time step must be adjusted to the chosen ϵ (on average, the particle must travel a distance smaller than one-half ϵ during one time step).

The use of code with variable softening lengths appears to be of limited value, and does not show particular promise.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-12-17152), the Federal Astronomy Program (project no. 40.022.1.1.1101), and by the President of the Russian Federation Program for Support of Leading Scientific Schools of Russia (grant no. NSh-1088.2003.2).

REFERENCES

1. D. Merritt, Astron. J. **111**, 2462 (1996).
2. E. Athanassoula, A. Bosma, J.-C. Lambert, *et al.*, Mon. Not. R. Astron. Soc. **293**, 369 (1998).
3. L. Hernquist, Astrophys. J. **356**, 359 (1990).
4. J. Barnes and P. Hut, Nature **324**, 446 (1986).
5. L. Hernquist, P. Hut, and J. Makino, Astrophys. J. **402**, L85 (1993).
6. P. J. Teuben, Astron. Soc. Pac. Conf. Ser. **77**, 398 (1995).
7. L. Hernquist and N. Katz, Astrophys. J., Suppl. Ser. **70**, 419 (1989).
8. E. Athanassoula, Ch. L. Vozikis, and J. C. Lambert, Astron. Astrophys. **376**, 1135 (2001).
9. W. Dehnen, Mon. Not. R. Astron. Soc. **324**, 273 (2001).

Translated by A. Dambis

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 477–484. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 535–543.

Cyclic Variation in the Spatial Distribution of the Coronal Green Line Brightness

O. G. Badalyan[1], V. N. Obridko[1], and J. Sýkora[2]

[1] *Institute of Terrestrial Magnetism, Ionosphere, and Radiowave Propagation, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*

[2] *Astronomical Institute, Slovak Academy of Sciences, Tatranská Lomnica, 05960 Slovak Republic*

Received December 1, 2004; in final form, December 3, 2004

Abstract—The spatial and temporal brightness distributions of the Fe XIV 530.3 nm coronal green line (CGL) and cyclic variations of these distributions are analyzed for a long time interval covering more than five 11-year cycles (1943–2001). The database of line brightnesses is visually represented in the form of a movie. Substantial restructuring of the spatial distribution of the CGL brightness occur over fairly short time intervals near the so-called reference points of the solar cycle; such points can be identified based on various sets of solar-activity indices. Active longitudes are observed in the CGL brightness over 1.5–3 yr. Antipodal and "alternating" active longitudes are also detected. The movie can be used to compare the CGL brightness data with other indicators of solar activity, such as magnetic fields. The movie is available at http://helios.izmiran.rssi.ru/hellab/Badalyan/green/. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The intensity of the solar FeXIV λ530.5 nm coronal emission line, which is the brightest coronal line in the optical, is a highly informative index of solar activity. A long series of systematic observations of the green line covering more than five past activity cycles is now available. An important advantage of this index is that it refers to all heliographic latitudes simultaneously, thus providing uniform data for studying solar activity over the entire surface of the Sun, in contrast to Wolf numbers, for example, which characterize low-latitude activity or polar faculae appearing at high latitudes.

The coronal green line (CGL) originates in the lower corona at temperatures of ~2 MK, which are most favorable for the production of the FeXIV ion. Calculations of the ionization equilibrium in the corona [1] indicate that the ionization curve of FeXIV is fairly narrow, so that the abundance of FeXIV ions decreases dramatically and the line weakens with either decreases or substantial increases in the temperature of the emission region. On the other hand, since the intensity of an emission line is proportional to the square of the density, the regions that are brightest in the CGL are dense loops and clusters of loops in the inner corona. The existence of such regions is related to and controlled by the coronal magnetic fields. Thus, studies of the spatial and temporal distributions of the CGL brightness can be used to trace the evolution of the coronal magnetic fields. It is important that the CGL brightness is a directly measurable index, in contrast to the strength of the coronal magnetic field, which can only be calculated using photospheric observations and adopting certain assumptions. Another important point is that the time interval covered by CGL observations is much longer than the interval for which systematic data on the photospheric magnetic fields are available.

The long series of CGL brightness observations supplements and complements currently available extra-atmospheric observations of the corona. The coronal images obtained with the Yohkoh, SOHO, TRACE, and CORONAS spacecraft cover the period since 1991. Such observations enable us to compare coronal images in various UV and X-ray lines with daily maps of the observed magnetic field. The radiation in the FeXII λ195 Å line recorded by instruments on spacecraft is emitted in nearly the same regions as the CGL. Images taken in the λ195 Å line show that the coronal emission is enhanced over active regions and weaker over coronal holes. However, since the instruments on various spacecraft are not identical, the observational data can be very nonuniform. This, along with the relatively short duration of the period covered by the extra-atmospheric observations, hinder the use of these observations for studies of long-period and cyclic variations in the physical conditions in the inner corona.

Comparisons of the spatial and temporal CGL brightness distributions with the strength of the magnetic field and its components represent a promising method for studying solar-activity variations and

1063-7729/05/4906-0477$26.00

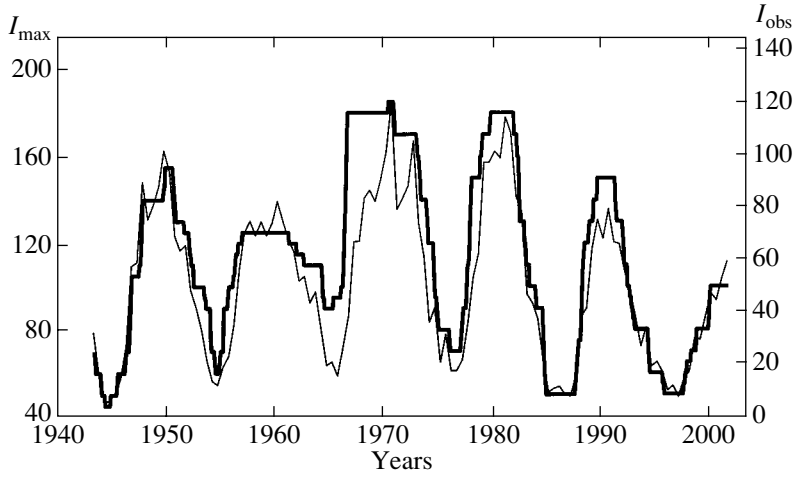


Fig. 1. Cyclic variation in the brightness of the coronal green line, I_{max}, adopted for the movie (heavy curve) compared to the brightness variations in I_{obs} observed in the equatorial zone $\pm 20°$ (light curve).

mechanisms for coronal heating. In [2, 3], we considered the relationship between cyclic variations in the CGL brightness and the coronal magnetic fields. We compared synoptic maps of the CGL brightness distribution and of the magnetic field at the level of the green-line emission calculated in a potential approximation based on observations of the photospheric field. This indicated some degree of correlation between the green-line intensity and the strength of the coronal magnetic field.

Long-term CGL brightness monitoring observations are also important for studies of so-called active longitudes. Investigation of this phenomenon requires long series of data, and numerous studies have been dedicated to this problem [4–8]. However, the available results are not always trustworthy, and are even contradictory in some cases. This stems from various facts. First, the rotational speeds of some tracers appear to differ from the Carrington velocity. Moreover, it has been shown that the solar plasma that entrains these tracers exhibits two different angular rotational speeds. Second, as is well known, the rate of rotation depends on the latitude, and this dependence may be different for different tracers. Finally, the technique of using tracers has the fundamental drawback that the tracers do not form a continuous numerical field. Therefore, to derive such a field for, e.g., sunspots, the frequency of sunspot emergence at a given point of the solar surface must be taken into account. For example, spots emerge only rarely at latitudes above 30°, making the results less reliable there. Thus, the CGL brightness data have the real advantage that they from a uniform and continuous numerical field over the entire solar disk over a long timescale. A short-timescale, CGL brightness investigation of active longitudes was carried out in [9–11].

Regions of faint green-line emission are of considerable interest. Waldmeier [12, 13] termed them coronal holes (Löcher in German). Subsequently, they were identified with regions of reduced EUV and X-ray brightness [14–16]. A detailed atlas of synoptic CGL brightness maps for 1947–1976 [17] was used to identify regions of faint CGL emission with properties similar to those of coronal holes. A comprehensive study of the sizes and locations of regions of faint CGL emission revealed a relationship between these regions and the solar wind and geomagnetic activity [18]. Such regions may be associated statistically with regions of reduced magnetic field at the same latitude.

We study here the spatial and temporal CGL brightness distributions during cycles 18–23. We have used a sequence of synoptic maps of the CGL brightness, with each map representing an average over six Carrington rotations, to trace cyclic variations in the spatial distribution of the CGL brightness, reveal active longitudes and their evolution, and identify alternating and antipodal active longitudes. These maps were used to compose a movie, which is available at http://helios.izmiran.rssi.ru/hellab/Badalyan/green/, where color illustrations for this paper can also be found.

2. BRIEF DESCRIPTION OF THE MOVIE

To obtain a photometrically uniform database of FeXIV 530.3 nm coronal emission line intensities, we compiled monitoring coronagraphic measurements that had been systematically carried out by a relatively small network of coronal stations (see [19, 20] for details concerning the compilation of the database; a brief description of the database can also be found

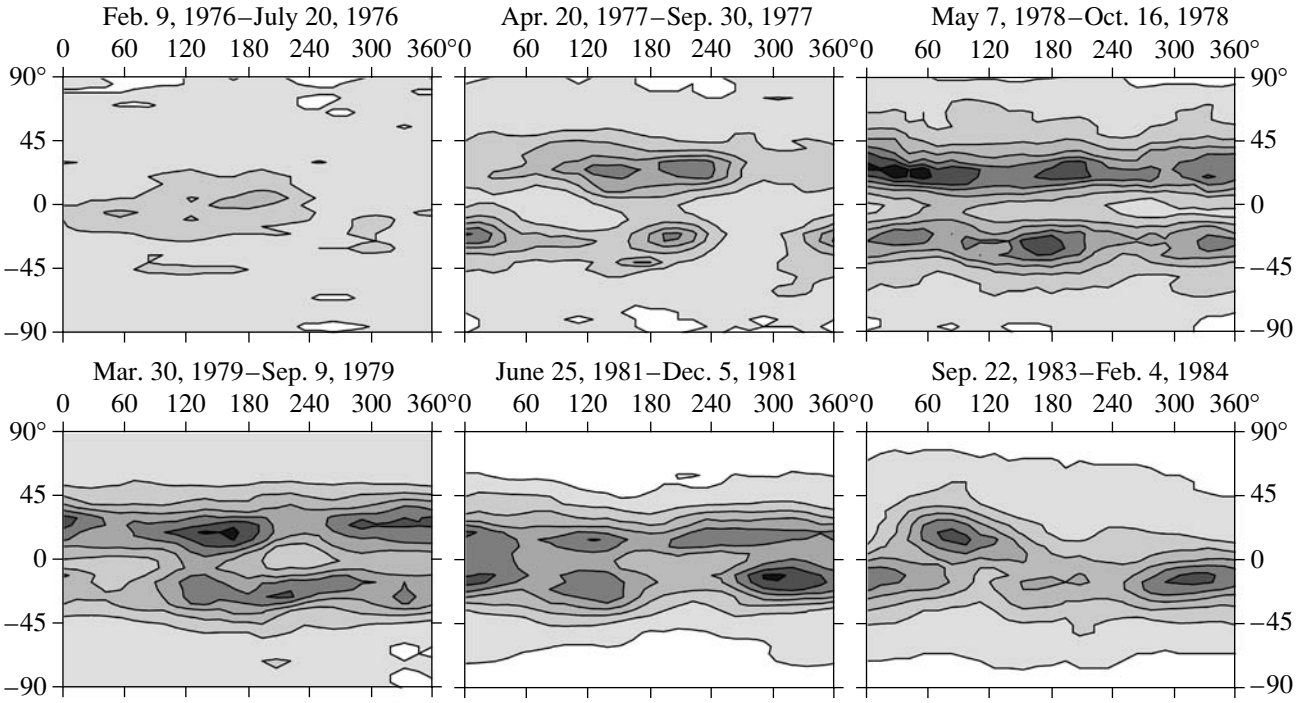


Fig. 2. Typical synoptic maps of the coronal green line for cycle 21. Each map is an average over six Carrington rotations; the time interval is indicated at the top of each map. The full brightness range for each map is divided into eight gradations, with the maximum brightness shown in black and the minimum brightness in white. Heliographic longitude and latitude are plotted as the horizontal and vertical coordinates. The steps ΔI for the successive maps are 7.5, 11, 15, 20, 22.5, and 10 absolute coronal units.

in [3, 21, 22]). The data obtained at various observatories were reduced to a height of 60″ over the limb. The spatial resolution is ∼13° in solar longitude (one observation per day) and 5° in solar latitude. The intensity of the green line is represented in absolute coronal units (acu, one millionth of the emission of the solar-disk center within 1 Å in the continuum next to the line). The original measurements refer to the eastern and western limbs, and the final data are reduced to the central meridian. It is these data that we used to construct the movie and all the figures given here. The CGL brightness database covers the period 1939–2001 and can be used to study time variations in the CGL brightness for individual latitude zones. This database has already been employed to analyze spatial and temporal changes and cyclic variations in the CGL brightness [20, 23–25].

We studied the evolution of the distribution of the CGL brightness using our movie, which contains 784 frames and covers the period from 1943 to 2001 (cycles 18–22 and half of cycle 23). Each frame is a CGL-brightness synoptic map averaged over six consecutive Carrington rotations. The time shift between successive frames is equal to one rotation. In each frame (and in the figures presented here), time increases from right to left and, accordingly, the Carrington longitude increases from left to right. The time indicated at the lower right corner of each frame corresponds to the beginning of the fourth of the six rotations within the averaging interval used for the map; i.e., to the center of the time interval considered. This time corresponds to the rightmost point of the horizontal axis. The curve in the lower left part of the frame represents the annual mean Wolf number, and the red point indicates the Wolf number at the beginning of the fourth rotation in the averaging interval; this time is indicated in the lower right part of the frame. Averaging over six Carrington rotations is best suited to reveal long-lived features on relatively large scales.

The brightness of the corona is color coded in each frame. The entire brightness range for the given frame is divided into eight gradations, with the maximum brightness shown in yellow and the minimum brightness in cyan. By the maximum brightness, I_{max}, we mean the highest contour level, which was chosen separately for each frame so as to satisfy the following conditions.

(1) The relative changes in the brightnesses of individual structures should be traceable from frame to frame. To this end, we tried to successively choose the maximum brightnesses so that they did not differ too strongly from one another, and all the main features of the brightness distributions were best represented

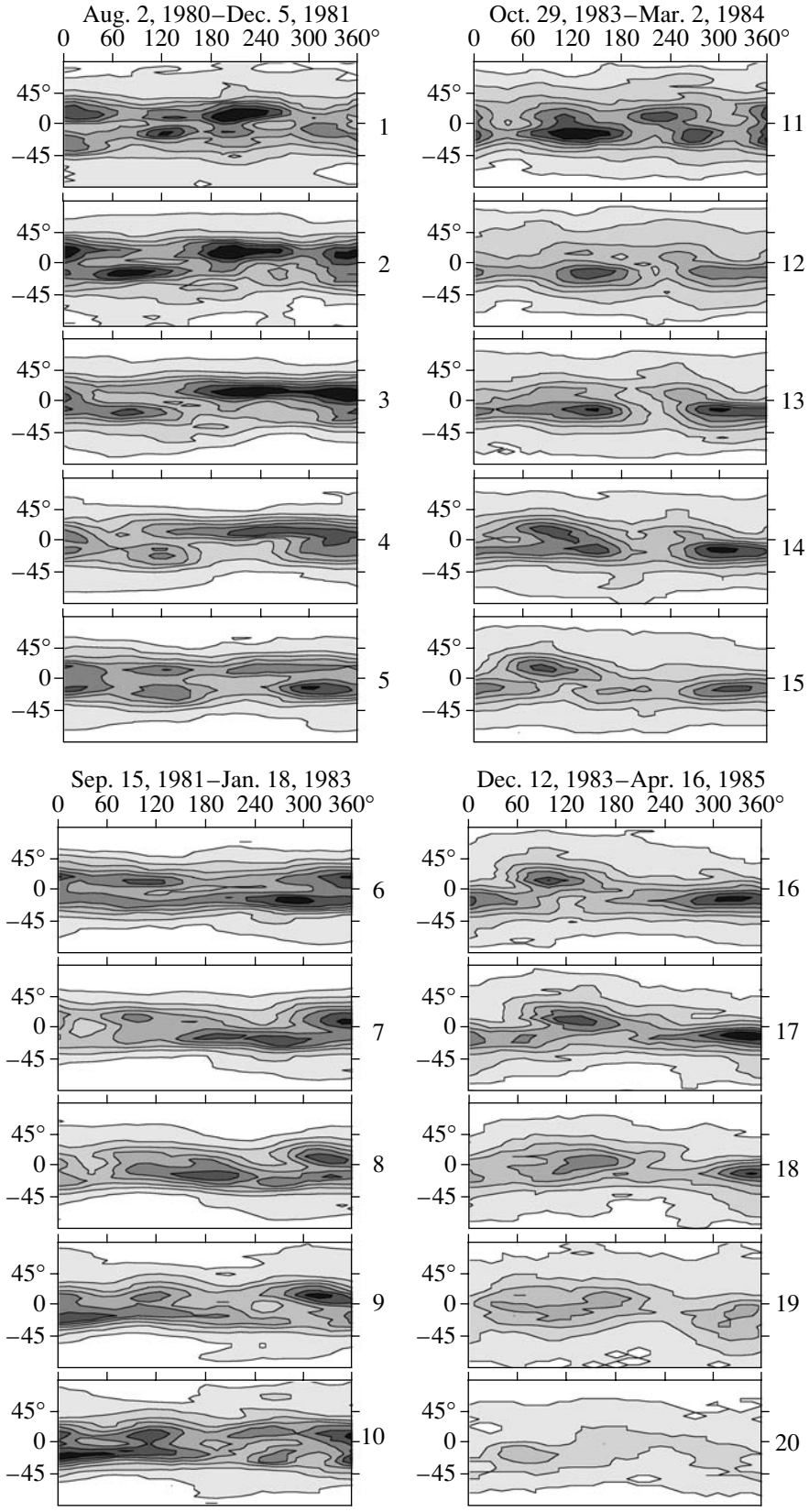


Fig. 3. Sequence of synoptic maps for the descending branch of cycle 21. Each map is an average over six rotations, and the time interval between maps corresponds to three rotations. The numbers on the right indicate the time sequence. The corresponding time interval is indicated at the top and in the middle of each column. The steps ΔI for the successive maps are 20, 20, 22.5, 22.5, 22.5, 20, 20, 17.5, 16.25, 15, 11.25, 11.25, 11.25, 10, 10, 8.75, 8.75, 7.5, 5, and 5 absolute coronal units.

in the frames. At the same time, the successive maximum brightnesses should change monotonically.

(2) The time variation of the maximum brightnesses over successive frames should, by and large, follow the cyclic variation in the CGL brightness.

The parameters of each frame are given in a table at the website indicated above. Figure 1 shows the time variations of the maximum brightness over a sequence of frames (heavy curve). The variation in the chosen maximum brightness $I_{\max}$ agrees very well with the cyclic variation in the CGL brightness $I_{\rm obs}$ in the equatorial region. When the two conditions stated above are met, not all possible brightness gradations can always be represented in a frame. The reason for this is that, in reality, the coronal brightness does not always vary with time monotonically, and periods of temporary decreases or increases in the overall brightness are noted fairly frequently. Furthermore, to better represent the general decrease in the brightness of the corona near the activity minima in the movie, we intentionally raised the maximum brightness adopted for the corresponding frame. For this reason, only three or even two gradations can be seen in some frames (see, e.g., the first map in Fig. 2 below, where four gradations are present).

The maps for 1954 should be noted separately. We had to change the lowest brightness value for that epoch and use $I_{\min} = 4$ acu instead of zero, so that the corresponding frames (a total of 16) do not visually fall out of the general sequence of frames. This situation may result from calibration errors associated with the stations that operated during that period.

Subsequent analyses demonstrated that appreciable restructurings in the general brightness distribution of the CGL become more pronounced when a long sequence of synoptic maps is considered. Constructing special map tables containing 15–20 frames each (similar to that shown below in Fig. 3) proved to be very convenient for determining the epochs of such restructurings.

3. SOME PROPERTIES OF THE SPATIAL BRIGHTNESS DISTRIBUTION OF THE GREEN LINE DURING THE ACTIVITY CYCLE

The long series of synoptic maps for the CGL brightness enables us to trace gradual changes in the brightness distribution during the solar-activity cycle. Epochs of specific restructurings in the synoptic map were identified, and compared with reference points in the cyclic curve, which can be determined using sets of solar-activity indices, primarily sunspots [26–30]. The reference points define the epochs of fundamental changes in the spatial and temporal organization of solar activity more precisely than do jumps in the time derivatives of various indices. The concept of reference points is important for understanding the nature of solar activity and forecasting of this activity. The essence of reference points is that t_{mA} and t_{AM} correspond to the beginning and end of the growth phase of the activity cycle, respectively, while t_{MD} and t_{Dm} represent the beginning and end of the decline phase. Usually, the points m and M—the epochs of minimum and maximum of the activity cycle—are added to these reference points.

Figure 2 illustrates the evolution of the CGL brightness distribution during the 21st activity cycle as an example. A specific brightness range from zero to $I_{\max}$ is chosen for each map, with the darkest color corresponding to the maximum CGL brightnesses and white to the minimum brightnesses. As in the movie, we chose the step ΔI in accordance with the cyclic variations in the CGL brightness, in order to better represent the transition from the minimum to the maximum of the cycle. Thus, each map has its own brightness scale. The maximum number of color gradations is eight; however, some maps do not contain all colors—see, e.g., maps 1, 2, and 6 in Fig. 2 (see Section 2 above). The maps in Fig. 2 illustrate the restructuring of the corona at the reference points. The heading of each map contains the six-Carrington-rotation time interval over which the averaging was done. Recall that time increases from right to left in this and the following figure.

The following scenario for the evolution of the CGL brightness can be described, using cycle 21 as an example (Fig. 2). The green line is faint almost everywhere on the Sun at the cycle minimum, and only isolated brightened regions are observed. Near the reference point t_{mA} (the beginning of the growth phase), two "rivers" with isolated, brighter "islands" appear. By t_{AM} (the beginning of the cycle-maximum phase), the two rivers become stable and brighten considerably. During the activity maximum, the rivers begin to approach and touch each other; by t_{MD} (the beginning of the decline phase), they merge into one river with two well-defined "streams." Finally, one narrower river forms by the reference point t_{Dm} (the beginning of the minimum phase), and begins to break up into separate bright islands. This evolutionary pattern can be traced during all five activity cycles.

The evolution of the spatial and temporal CGL brightness distribution can be traced in more detail using successive synoptic maps. The evolution of the large-scale structure of the green corona for the decline phase of cycle 21 is shown in Fig. 3. As in Fig. 2, each map represents an average over six rotations. The time interval between successive maps is three rotations. Gradual changes in the CGL brightness distribution and the presence of active longitudes can

Times of spatial restructurings in the CGL brightness distribution

Cycle	t_{mA}		t_{AM}		M		t_{MD}		t_{Dm}	
18	1230	1229	1244	1242	1251	1256	1276	1282	1313	1306
19	1360	1363	1376	1379	1390	1393	1424	1419	1444	1440
20	1497,	1499,	1525	1520,	1540	1540	1560	1558	1596	1598,
	1514	1519		1529						1615
21	1657	1656,	1677	1680	1684	1692	1698,	1715	1749	1759,
		1665					1714			1748
22	1791	1781,	1815	1812	1820	1818	1834	1840	1877	1875
		1804								
23	1924	1929,	1954	1955	1956	1961				
		1940								

be seen over 1.5–3 years. For example, an activity enhancement at longitudes ~210° that gradually moves to 240° can be noted in maps 1–5. This active longitude is most pronounced in the northern hemisphere during this period. Maps 6 and 7 already exhibit a brightening at these longitudes in the southern hemisphere, which later gradually disappears. In contrast, a depression of brightness can be seen at these longitudes in map 12 and the following maps. In many cases, two pronounced brightenings are observed at nearly the same longitudes in both the northern and southern hemispheres (see, e.g., maps 10, 11, and 14) as a single activity complex existing simultaneously in both hemispheres. As the activity minimum is approached, a general decrease in the green-line brightness is clearly visible in the maps.

The table presents the epochs of restructurings of the spatial brightness distribution of the CGL determined for the above scenario for the cyclic CGL brightness variations. The activity-cycle number is given in the first column of the table. For each reference point, the number of the Carrington rotation that corresponds to a CGL brightness restructuring is indicated in the left column, together with the rotation number for the reference points determined from the Wolf numbers and large-scale magnetic fields (see [30]). In some cases, the reference points determined from the Wolf numbers and large-scale magnetic fields are obviously in disagreement (especially for t_{mA}), and two values are given in the table. In two cases, the epoch of restructuring could not be unambiguously determined based on the CGL brightness.

Note that, when the idea of reference points was suggested [26], it was assumed that the reference-point epochs based on different solar-activity indices would coincide. This is usually the case. However, it is fairly difficult to accurately determine the reference points. Therefore, as much information as possible should be employed, and the use of coronal data is very promising in this regard.

Figure 4 presents nine yearly latitude–time synoptic maps for cycle 21. Two "rivers" and their evolution are clearly visible in these maps. The brightness distribution during the ascending branch of the activity cycle definitely exhibits 27-day and 13-day periods; accordingly, two antipodal (separated by 180°) longitudes arise, which are present almost synchronously in both hemispheres of the Sun. During 1982, at the beginning of the decline phase, the 27-day period can easily be identified from the outer enveloping contour. The 13-day period is still observed near the equator. As 1984 is approached, the bright regions begin to penetrate into the opposite hemisphere. This effect is especially pronounced in 1984. This phenomenon could be called "alternating" active longitudes, since the active longitudes in the northern and southern hemispheres are activated alternately. This is manifest in the formation of a four-sector structure of the corona (the 13-day period in the brightness variations). Finally, only the 27-day period is observed in the equatorial zone during 1985. This means that only one active longitude was present on the Sun at that time.

4. CONCLUSION

We have analyzed cyclic variations in the brightness distribution of the coronal green line using our movie illustrating these variations in a database comprising data for 1943 to 2001. Our analysis of the spatial brightness distribution of the CGL for five cycles demonstrates the following.

(1) The spatial brightness distribution of the CGL experiences substantial changes within relatively

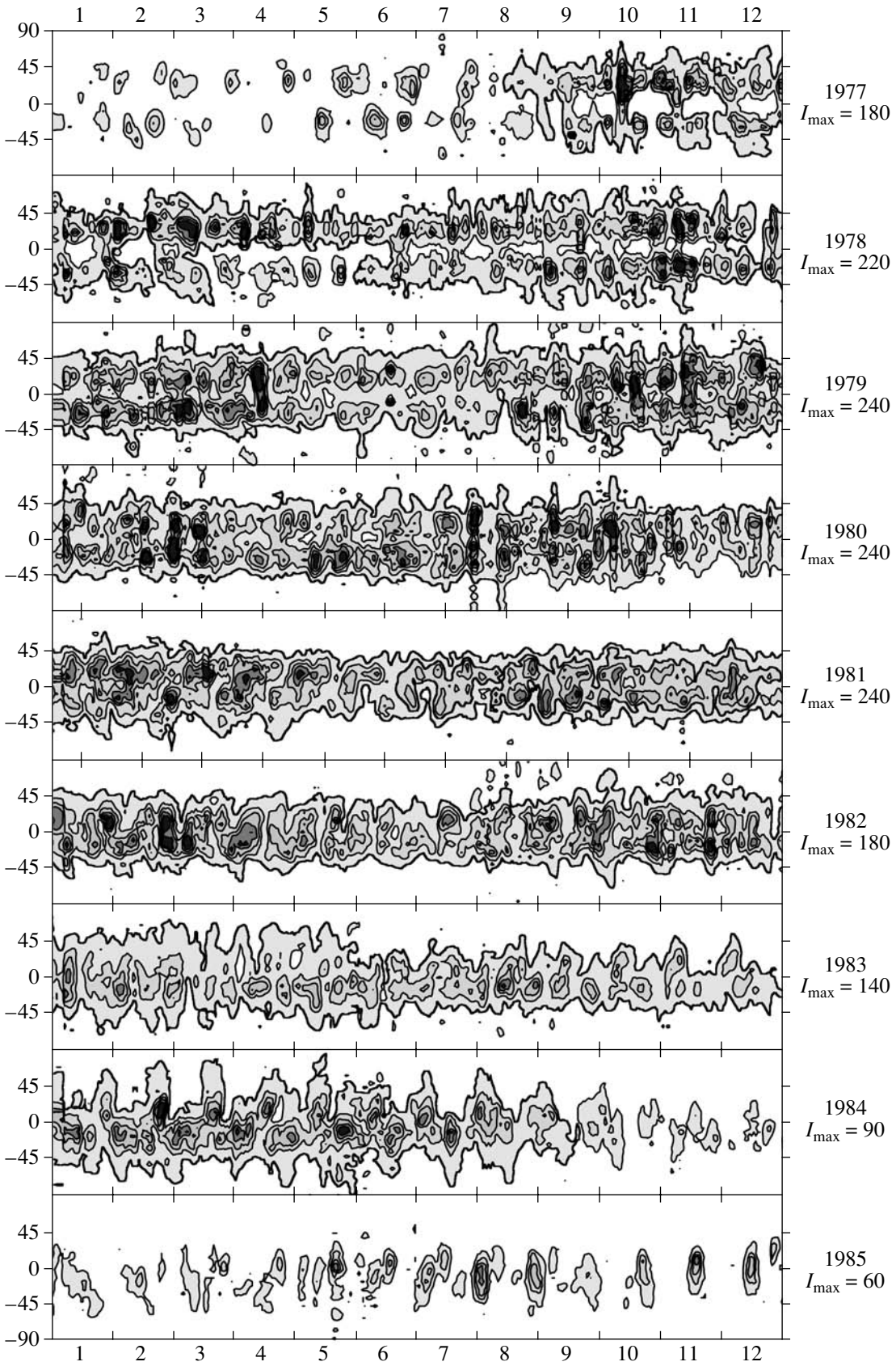


Fig. 4. Sequence of yearly latitude–time maps for cycle 21. The brightness range for each map is divided into six gradations. The year is indicated on the right of the map, and the month numbers in the given year are indicated on the horizontal axis.

short time intervals near the reference points of the cycle determined using other solar-activity indices.

(2) Active longitudes are manifest in the brightness of the corona over time scales of about 1.5–3 years. In many cases, as one of these longitudes becomes fainter, another that is shifted by 180° becomes activated.

(3) During periods of high activity, two active longitudes separated by 180° are present, resulting in variations with a period of 13–14 days. When the activity is lower, only one active longitude is observed in each hemisphere, and these two longitudes are shifted relative to each other by half a Carrington rotation.

Our analysis demonstrates that CGL brightness data are highly informative. The brightness of the CGL is among the few solar-activity indices that can be used to study solar activity over all heliographic latitudes using the same observational data. The CGL intensity clearly reflects the effect of variations in solar-activity mechanisms, of which the magnetic field is probably the most important. Our analysis demonstrates how the green corona fits into the general scenario for cyclic variations in the solar activity. Long series of CGL brightness observations can be used to reconstruct series of magnetic-field data or to test series obtained using other indirect techniques. These series can also be employed to identify coronal holes—an important element of solar activity in terms of their effects at the Earth—during periods when they are not directly observed from spacecraft.

Our movie, which illustrates the spatial and temporal variations in the green corona over a long time interval, can be used to compare the CGL brightness with other characteristics of solar activity, in particular, with magnetic fields.

ACKNOWLEDGMENTS

Thus work was supported by the Russian Foundation for Basic Research (project no. 02-02-16199), INTAS (grant no. 2000-840), and VEGA (grant no. 2/4013/24) of the Slovak Academy of Sciences.

REFERENCES

1. M. Arnaud and R. Rothenflug, Astron. Astrophys., Suppl. Ser. **60**, 425 (1985).
2. O. G. Badalyan, V. N. Obridko, and J. Sýkora, Astron. Astrophys. Trans. (2004, in press).
3. O. G. Badalyan and V. N. Obridko, Astron. Zh. **81**, 746 (2004) [Astron. Rep. **48**, 678 (2004)].
4. Yu. I. Vitinskii, M. Kopetskii, and G. V. Kuklin, *Statistics of Spot-Forming Activity of the Sun* (Nauka, Moscow, 1986) [in Russian].
5. G. A. Stewart and S. Bravo, Adv. Space Res. **17** (4–5), 217 (1996).
6. V. Bumba, A. Garcia, and M. Klvaňa, Sol. Phys. **196**, 403 (2000).
7. M. Neugebauer, E. J. Smith, A. Ruzmaikin, J. Feynman, and A. H. Vaughan, J. Geophys. Res. **105**, 2315 (2000).
8. S. V. Berdyugina and I. G. Usoskin, Astron. Astrophys. **405**, 1121 (2003).
9. E. E. Benevolenskaya, A. G. Kosovichev, and P. H. Scherrer, in Solar Physics Division Meeting, No. 31, Abstract No. 02.26; Bull. Am. Astron. Soc. **32**, 815 (2000).
10. J. Xanthakis, B. Petropoulos, V. P. Tritakis, H. Mavromichalaki, and L. Marmatsuri, Adv. Space Res. **11** (1), 169 (1991).
11. J. Xanthakis, H. Mavromichalaki, B. Petropoulos, *et al.*, Adv. Space Res. **17** (4–5), 277 (1996).
12. M. Waldmeier, Z. Astrophys. **38**, 219 (1956).
13. M. Waldmeier, Sol. Phys. **70**, 251 (1981).
14. R. H. Munro and G. L. Withbroe, Astrophys. J. **176**, 511 (1972).
15. A. S. Krieger, A. F. Timothy, and E. C. Roelof, Sol. Phys. **29**, 505 (1973).
16. G. S. Vaiana, A. S. Krieger, and A. F. Timothy, Sol. Phys. **32**, 81 (1973).
17. V. Letfus and J. Sýkora, *Atlas of the Green Corona Synoptic Charts for Period 1947–1976* (Veda, Bratislava, 1982).
18. J. Sýkora, Sol. Phys. **140**, 379 (1992).
19. J. Sýkora, Bull. Astron. Inst. Czech. **22**, 12 (1971).
20. J. Sýkora, Contrib. Astron. Obs. Skalnaté Pleso **22**, 55 (1992).
21. O. G. Badalyan, V. N. Obridko, and J. Sýkora, Sol. Phys. **199**, 421 (2001).
22. J. Sýkora and J. Rybák, Adv. Space Res. (2004) (in press).
23. J. Sýkora, *Solar and Interplanetary Dynamics*, Ed. by M. Dryer and E. Tandberg-Hanssen (Reidel, Dordrecht, 1980), p. 87.
24. J. Sýkora, Adv. Space Res. **14** (4), 73 (1994).
25. J. Sýkora, O. G. Badalyan, and M. Storini, Adv. Space. Res. **29**, 1975 (2002).
26. Yu. I. Vitinskii, G. V. Kuklin, and V. N. Obridko, Soln. Dannye, No. 3, 53 (1986).
27. G. V. Kuklin, V. N. Obridko, and Yu. I. Vitinsky, *Solar–Terrestrial Predictions—1*, Ed. by R. J. Thompson, D. G. Cole, P. J. Wilkinson, *et al.* (Boulder, CO, 1990), p. 474.
28. V. N. Obridko and G. V. Kuklin, in *Solar–Terrestrial Predictions—IV*, Ed. by J. Hruska, M. A. Shea, D. F. Smart, and G. Heckman (Boulder, CO, 1994), p. 273.
29. G. V. Kuklin and V. N. Obridko, Izv. Ross. Akad. Nauk, Ser. Fiz. **59**, 12 (1995).
30. V. N. Obridko and B. D. Shel'ting, Astron. Zh. **80**, 1034 (2003) [Astron. Rep. **47**, 953 (2003)].

Translated by A. Getling

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 485–494. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 544–554.

Turbulence in the Inner Solar Wind Determined from Frequency Fluctuations of the Downlink Signals from the ULYSSES and GALILEO Spacecraft

A. I. Efimov[1], I. V. Chasheĭ[2], M. K. Bird[3], L. N. Samoznaev[1], and D. Plettemeier[4]

[1] *Institute of Radio Engineering and Electronics, Russian Academy of Sciences, ul. Mokhovaya 11-7, Moscow, 125009 Russia*

[2] *Lebedev Institute of Physics, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 117924 Russia*

[3] *Radioastronomisches Institut, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*

[4] *Institut für Hochfrequenztechnik, Ruhr-Universität Bochum, Gebäude IC 6/133, Universitätsstrasse 150, D-44780 Bochum, Germany*

Received August 15, 2004; in final form, December 3, 2004

Abstract—The results of several sets of measurements of the frequency of radio signals during coronal-sounding experiments carried out from 1991 to 2000 using the ULYSSES and GALILEO spacecraft are presented and analyzed. The S-band signals (carrier frequency $f = 2295$ MHz) were received at the three 70-m widely spaced ground stations of the NASA Deep Space Network. As a rule, the frequency-fluctuation spectra at frequencies above 1 mHz are power-laws. At small heliocentric distances, $R < 10R_\odot$ ($R_\odot$ is the solar radius), the spectral index is close to zero; this corresponds to a spectral index for the one-dimensional turbulence spectrum $p_1 = 1$. The index of the frequency-fluctuation spectra in the region of the supersonic solar wind at distances $R > 30\,R_\odot$ is between 0.5 and 0.7 ($p_1 = 1.5$–1.7). The results demonstrate a substantial difference between the turbulence regimes in these regions: in the region of the established solar wind, the power-law spectra are determined by nonlinear cascade processes that pump energy from the outer turbulence scale to the small-scale part of the spectrum, whereas such cascade processes are absent in the solar wind acceleration region. Near the solar minimum, the change in the turbulence regime of the fast, high-latitude solar wind occurs at greater distances than for the slow, low-latitude solar wind. Spectra with a sharp cutoff at high frequencies have been detected for the first time. Such spectra are observed only at $R < 10\,R_\odot$ and at sufficiently low levels of the electron density fluctuations. The measured cutoff frequencies are between 10 and 30 mHz; the cutoff frequency tends to increase with heliocentric distance. The variance of the plasma-density fluctuations has been estimated for the slow, low-latitude solar wind. These estimates suggest that the relative fluctuation level at distances $7\,R_\odot < R < 30\,R_\odot$ does not depend on heliocentric distance. The cross correlation of the frequency fluctuations recorded at widely spaced ground stations increases with the index of the frequency-fluctuation spectrum. At distances $R \approx 10\,R_\odot$, the rate of temporal changes in irregularities on the scale of several thousand kilometers is less than or comparable to the solar wind velocity. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Long-term studies using local methods and sounding of circumsolar plasma using radio signals from both natural and man-made sources demonstrate that turbulence is a constant property of the solar wind: all plasma parameters at all heliocentric distances and heliolatitudes experience random fluctuations, which are characterized by a broad spectrum of spatial and temporal scales. The evolution of the turbulence and its interaction with the moving plasma flow is of primary interest for the physics of the solar wind; this is, in particular, relevant to the inner region of the flow, where it is accelerated to supersonic and super-Alfvénic velocities. However, precisely these regions, at both low and high heliolatitudes, are currently accessible only to radio-propagation methods. Interplanetary scintillations (intensity fluctuations) of radio waves are due to relatively small-scale (of the order of 100 km) irregularities in the electron density of the solar wind plasma [1]. Sounding using monochromatic spacecraft signals allows us to measure the phase and frequency fluctuations of radio signals [2–4]. Under the experimental conditions that are realized, as a rule, such measurements yield information about larger-scale ($\geq$1000 km) irregularities of the electron density. Radio-sounding experiments using monochromatic signals have been carried out since 1991 using the ULYSSES spacecraft, and

1063-7729/05/4906-0485$26.00

since 1995 using GALILEO. A considerable amount of data on the frequency fluctuations of radio signals at various heliocentric distances and heliolatitudes and at various phases of the solar cycle has been accumulated in these experiments. Even the very first results demonstrated that the available observational material contains valuable information on solar wind irregularities. In particular, due to the extended durations of individual measurement sets and the high stability of the spacecraft radio transmitters, it was possible to obtain the first estimates of the outer scale of the turbulence, at heliocentric distances from 7 $R_\odot$ to 80 $R_\odot$ [5, 6].

We report here new results obtained from radio-sounding data using the ULYSSES and GALILEO spacecraft, related to the level and spatial spectrum of density fluctuations in the inner regions of the solar wind, as well as the dynamics of the density irregularities during their motion relative to the path of the radio signal.

2. BASIC THEORETICAL RELATIONSHIPS

In this section, we present basic theoretical expressions relating the observed statistical characteristics of the frequency fluctuations of radio-sounding signals to the spatial spectrum of the turbulence and the velocity of the solar wind. We assume that the three-dimensional spatial spectrum of the electron density fluctuations in the solar wind $\Phi_N(\mathbf{q})$ can be represented

$$\Phi_N(\mathbf{q}) = C(r)(q^2 + q_0^2)^{-p/2}\exp(-q^2/q_m^2), \quad (1)$$

where r is the heliocentric distance, q is the wave number, p is a three-dimensional exponent, and $L_0 = 2\pi/q_0$ and $L_m = 2\pi/q_m$ are the outer and inner scales of the turbulence, respectively ($L_0 \gg L_m$). If the spectrum (1) is normalized to the variance of the density fluctuations $\sigma_N^2(r)$, the structural constant $C(r)$ is

$$C(r) = A(p, q_0, q_m)\sigma_N^2(r), \quad (2)$$

where

$$A(p, q_0, q_m) \quad (3)$$
$$= \begin{cases} \dfrac{(p-3)\Gamma(p/2)q_0^{p-3}(2\pi)^{-3/2}}{\Gamma[(p-1)/2]} & \text{for} \quad 3 < p < 4, \\ \left[4\pi \ln\left(\dfrac{2q_0}{q_m}\right)\right]^{-1} & \text{for} \quad p = 3. \end{cases}$$

If we assume that the spatial fluctuations are "frozen" into the plasma and are convectively transported at velocity V, then, as is shown in [7], the temporal power spectrum of the frequency fluctuations of the radio-sounding signal $G'_f(\nu)$ for the electron density fluctuation spectrum (1) is

$$G'_f(\nu, V) = D(r, V, \lambda)\nu^2(\nu_0^2 + \nu^2)^{-(\alpha+2)/2} \times \exp(-\nu^2/\nu_0^2), \quad (4)$$

where

$$D(r, V, \lambda) = B\pi^{-1}(\lambda r_e)^2\sigma_N^2(R)V\nu_0^2\Lambda_{eff}, \quad (5)$$

$\alpha = \alpha_f = p - 3$ is the index of the frequency-fluctuation spectrum, λ is the radio wavelength, $r_e = 2.82 \times 10^{-13}$ cm is the classical radius of the electron, $\Lambda_{eff} \approx R$ is the effective thickness of the layer with modulating irregularities that is adjacent to the point of closest approach, $\nu_0 = V/L_0$ and $\nu_m = V/L_m$ are temporal frequencies corresponding to the outer and inner turbulence scales, and the constant B is

$$B = \begin{cases} \alpha & \text{for} \quad 0 < \alpha < 1, \\ \left[\ln\left(\dfrac{2q_m}{q_0}\right)\right]^{-1} & \text{for} \quad \alpha = 0. \end{cases} \quad (6)$$

However, strictly speaking, the spatial pattern of the fluctuations is not frozen in; this may be due either to an intrinsic change in the irregularities during their transport by the solar wind (this takes place for irregularities associated with waves propagating relative to the moving plasma) or to the presence in the line of sight of an instantaneous spread of outflow velocities [8]. The deviations from freezing can be described by some velocity distribution function $\varphi(V)$, which we assume to be one-dimensional. Strong variability of the fluctuation pattern will correspond to a considerable spread of velocities $\langle \Delta V^2 \rangle \geq \langle V \rangle^2$; in the opposite case, when $\langle \Delta V^2 \rangle \ll \langle V \rangle^2$, the effects of deviation from freezing will be negligible. With the variability of the moving pattern taken into account, relationship (4) for the frequency-fluctuation spectrum becomes

$$G_f(\nu) = \int G'_f(\nu, V)\varphi(V)dV. \quad (7)$$

Let us suppose that, as is the case for the solar wind, the spread of possible velocities in the distribution $\varphi(V)$ is limited by some values $V_{\min}$ and $V_{\max}$ (so that $V_{\min} \leq V \leq V_{\max}$), $V_{\min}$ and $V_{\max}$ are of the same order, and the power-law part of the density-fluctuation spectrum (1) is fairly broad ($q_0 \ll q_m$). It then follows from (7) that, even in the presence of considerable variability of the pattern, the shape and parameters of the frequency-fluctuation spectrum $G_f(\nu)$ will be determined by (4), with the velocity of the irregularities V replaced by its effective value V_{eff}:

$$V_{eff} = (\langle V^{p-2} \rangle)^{1/(p-2)}, \quad (8)$$
$$\langle V^{p-2} \rangle = \int V^{p-2}\varphi(V)dV.$$

In the particular case $p = 3$, the velocity V_{eff} [formula (8)] coincides with the average velocity of the irregularities $\langle V \rangle$. The frequency-fluctuation spectrum (4), (7) has a maximum at the fluctuation frequency

$$\nu_{\max} = \begin{cases} \left(\frac{2}{\alpha}\right)^{1/2} \nu_0 & \text{for} \quad 0 < \alpha < 1, \\ (\nu_0 \nu_m)^{1/2} & \text{for} \quad \alpha = 0. \end{cases} \tag{9}$$

It can readily be verified that there is no such maximum in the wave-phase fluctuation spectrum $G_{\rm ph}(\nu) \sim \nu^{-2} G_f(\nu)$. According to [5, 6], the expected characteristic frequency $\nu_{\max}$ [formula (9)] in the region of the developed solar wind, $R > 20\,R_\odot$ with $\alpha \approx 0.6$–0.7 and $V = \text{const}$, is $\nu_{\max} \sim 0.1$ mHz. In the region of the solar wind acceleration at $R < 10\,R_\odot$, where spectra with $\alpha \approx 0$ are observed [2], the values of $\nu_{\max}$ can be much higher: $\nu_{\max} \geq 1$ mHz. However, in this case, the spectral maximum (9) will be pronounced weakly and will be of little use for determining the outer (or inner) turbulence scale. Substituting $\nu_{\max}$ from (9) into (4), we find the spectral density of the fluctuations at the peak frequency:

$$G_f(\nu_{\max}) = \lambda^{-1} B_1 (\lambda r_e)^2 \sigma_N^2(r) R, \tag{10}$$

where

$$B_1 = \begin{cases} 2\left[\frac{\alpha}{\alpha+2}\right]^{(\alpha+2)/2} & \text{for} \quad 0 < \alpha \leq 1, \\ \left[\ln\left(\frac{2q_m}{q_0}\right)\right]^{-1} & \text{for} \quad \alpha = 0. \end{cases} \tag{11}$$

If we know the values of α, V_{eff}, and $G_f(\nu_{\max})$ from measurements, (10) and (11) can be used to estimate the variance of the density fluctuations in the solar wind plasma, $\sigma_N^2(R)$.

The motional velocity of the irregularities can be estimated from simultaneous measurements of the frequency fluctuations at widely spaced ground stations. In this case, the temporal cross-correlation function for the frequency fluctuations $K(\boldsymbol{\rho}, \tau)$ is related to the corresponding cross-correlation function for the phase fluctuations $K_{\rm ph}(\boldsymbol{\rho}, \tau)$ as

$$K(\boldsymbol{\rho}, \tau) = -\frac{\partial^2 K_{\rm ph}(\boldsymbol{\rho}, \tau)}{\partial \tau^2}, \tag{12}$$

where $\boldsymbol{\rho}$ is the projection of the baseline vector onto the region of efficient phase modulation adjacent to the point of closest solar approach in the line of sight and τ is the time shift. The function $K(\boldsymbol{\rho}, \tau)$ peaks at some shift $\tau = \tau_0$; this value can be used to find the apparent convective velocity of the spatial pattern of the fluctuations:

$$V_{app} = x/\tau_0, \tag{13}$$

where x is the projection of the vector $\boldsymbol{\rho}$ onto the radial direction. The velocity V_{app} does not depend on the anisotropy of the irregularities [9]; for widely spaced measurements of the frequency fluctuations, it is determined by the third and fourth moments of the distribution $\varphi(V)$:

$$V_{app} = \langle V^4 \rangle / \langle V^3 \rangle. \tag{14}$$

Under certain assumptions about the velocity distribution function $\varphi(V)$, the measured value of V_{app} (14) can be used to find V_{eff} (8). In particular, if the velocity dispersion is insignificant, then $V_{eff} \approx V_{app}$. Even in the case of two flows with different velocities and equal weights, when $\varphi(V) = \frac{1}{2}[\delta(V - V_1) + \delta(V - V_2)]$, $\Delta V = V_2 - V_1$, $V_{eff} \approx \frac{1}{2}(V_1 + V_2)$, the relative velocity difference is $\frac{V_{app} - V_{eff}}{V_{eff}} \approx \frac{3\Delta V^2}{(4V_{eff}^2 + 3\Delta V^2)}$, and for $\Delta V \leq V_{eff}$ does not exceed 0.4.

3. OBSERVATIONS AND DATA REDUCTION

We used as our input data the results of radio sounding of the circumsolar plasma carried out from 1991 to 2000 with the GALILEO and ULYSSES spacecraft. During the ingress and egress phases of the solar conjunction, the three 70-m radio telescopes of the Deep Space Network in Goldstone (USA), Canberra (Australia), and Madrid (Spain) recorded the frequency of the S-band downlink carrier signal emitted by an onboard transmitter ($\lambda = 13$ cm, $f = 2295$ MHz) at a sampling rate of 1 Hz. The observations were carried out for radio ray path closest approach distances of $R = 5$–$80\,R_\odot$.

The processing of the raw data included the creation of long continuous series of frequency measurements containing thousands and even tens of thousands of one-second samples. The series were formed from consecutive individual sessions carried out at different stations. The spacecraft was always visible by at least one ground station, ensuring continuity of the frequency record. The frequency fluctuations produced by turbulence in the circumsolar plasma were found by subtracting the slowly varying component (trend), which was approximated by a polynomial fit to the observational segment to be processed.

The subsequent data processing consisted of finding the temporal power spectra of the frequency fluctuations of the downlink sounding signals via a Fast Fourier Transform (FFT). In all time intervals when the spacecraft was visible from two stations, the frequency measurements were carried out at both stations. The data of these joint observations were cross

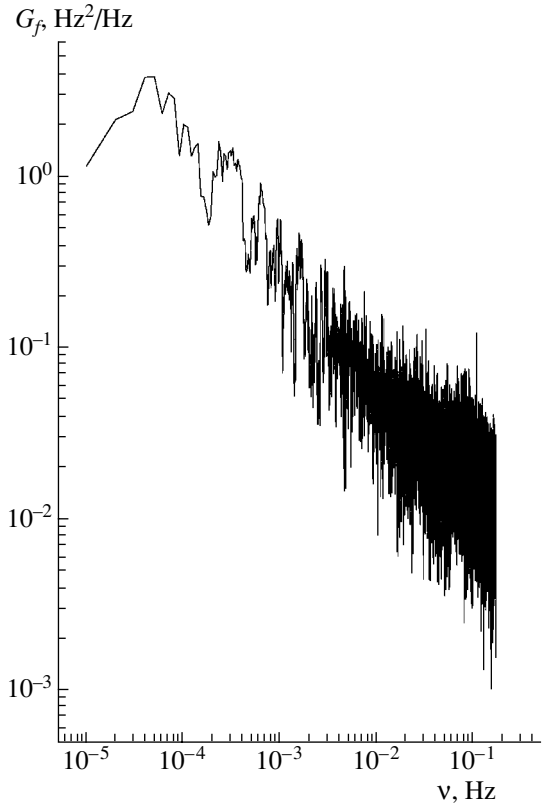


Fig. 1. Temporal power spectra of the frequency fluctuations based on GALILEO data.

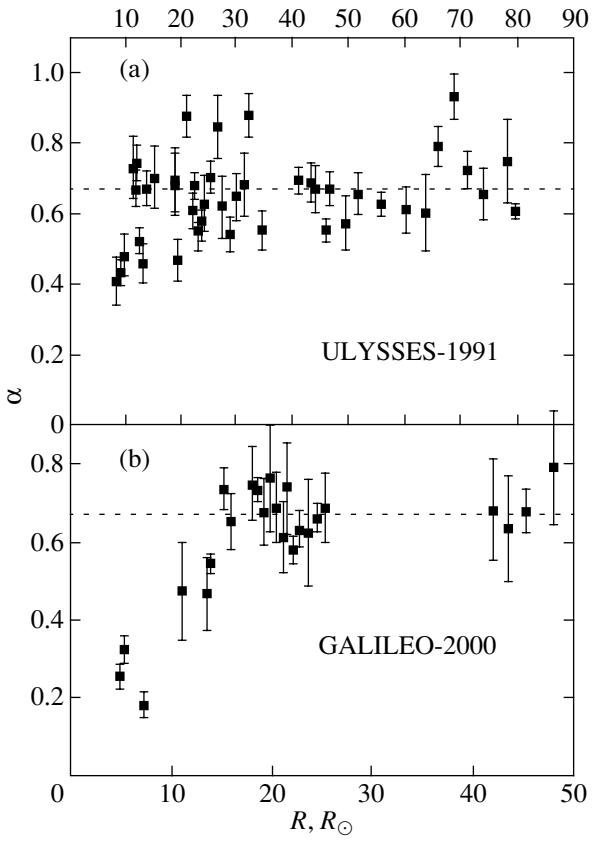


Fig. 2. Spectral index α of the temporal spectra of the frequency fluctuations of the radio signals from the (a) ULYSSES and (b) GALILEO spacecraft at various heliocentric distances R.

correlated to find the convective velocity of the irregularities across the downlink radio path. This is approximately equal to the solar wind speed, since the change of the closest-approach distance with time, $v_{gr} \approx 24$ km/s, is much lower than this speed.

4. SPECTRAL PARAMETERS OF THE SOLAR WIND TURBULENCE

Figure 1 shows the frequency-fluctuation spectrum obtained from continuous measurements over more than a day at the Canberra, Madrid, Goldstone, and again Canberra stations on January 9–10, 1997. The duration of the observations was about 27 h ($\sim 10^5$ measurements), and the average distance of closest approach was $R \approx 27.3\,R_\odot$. These data were averaged over three measurements to reduce the effect of instrumental noise. We obtained the temporal spectrum using an FFT with $2^{15} = 32\,768$ points.

The shape of the temporal spectrum presented in Fig. 1 is typical of most of the observational series processed in this way: there is a local, smoothly varying maximum at a low fluctuation frequency and a power-law interval $\nu > \nu_{\max}$, which extends to frequencies where the measured fluctuations become lower than the noise level. The main parameters of the frequency-fluctuation spectrum [from which we can find the outer turbulence scale L_0 and the rms fluctuations of the electron density σ_N using (9) and (10)] are the frequency of the maximum $\nu_{\max}$ (equal, in this case, to 4×10^{-5} Hz), the spectral density at this maximum $G_f(\nu_{\max}) \approx 2$ Hz, and the index α in the inertial interval $\nu > \nu_{\max}$ ($\alpha = 0.73$). The average velocity V in the time interval when the spectrum was determined (derived from cross-correlation processing of the simultaneous frequency measurements at two stations) was $V \approx 2 \times 10^2$ km; hence, we obtain $L_0 \approx 8.28 \times 10^6$ km, $\sigma_N = 2 \times 10^2$ cm^{-3}, and the index of the spatial spectrum of the electron density turbulence $p = 3.73$.

Figure 2 shows the index α of the frequency-fluctuation spectra as a function of heliocentric distance; α is equal to the index of the spatial spectrum of the plasma turbulence minus three. Figure 2a presents the values of α obtained from the ULYSSES data (1991). The horizontal axis for Fig. 2a is labeled above. Figure 2b (horizontal axis labeled below) presents the results for GALILEO frequency measurements carried out from April 24 to May 24, 2000.

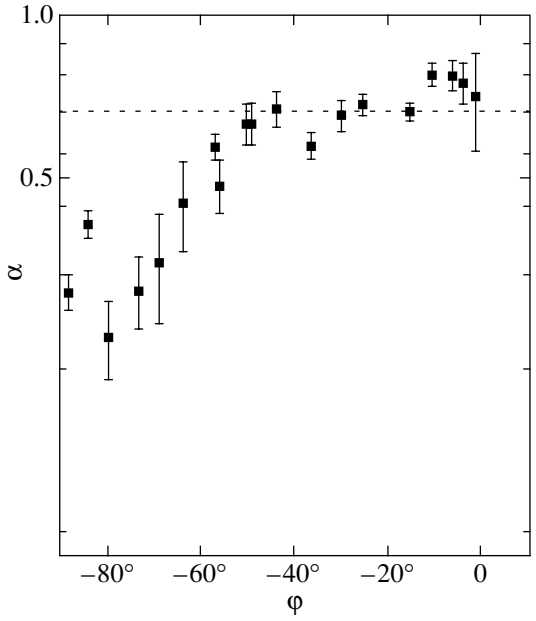


Fig. 3. Spectral index α of the frequency-fluctuation spectrum as a function of the heliolatitude of the sounded region (ULYSSES data).

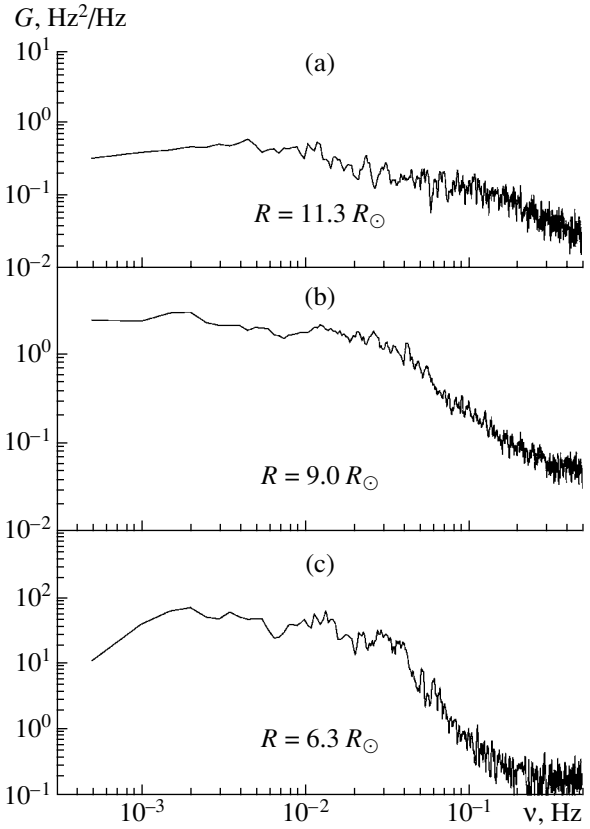


Fig. 4. Frequency-fluctuation spectra of the GALILEO radio downlink signals (experiments in 1999–2000) at small heliocentric distances R: (a) $R = 11.3\,R_\odot$, (b) $R = 9.0\,R_\odot$, (c) $R = 6.3\,R_\odot$.

The dashed lines in Figs. 2a and 2b show the spectral index for the Kolmogorov–Obukhov turbulence spectrum (2/3). On the whole, the data of Fig. 2 characterize turbulence modes of the circumsolar plasma for mid-latitude (Fig. 2a) and low-latitude (Fig. 2b) regions during epochs of high solar activity. These results testify that, near the solar-activity maximum, the turbulence is well developed at heliocentric distances $R > 12\,R_\odot$ (its index is close to 2/3), whereas the turbulence remains undeveloped at smaller distances, $R < 10\,R_\odot$.

Similar changes in the index for low-latitude regions of the solar wind were found earlier [2] from measurements of the phase-fluctuation spectra for radio signals from the VIKING spacecraft. However, the phase-fluctuation spectra [2] were obtained in a narrower spectral interval that was entirely within the power-law segment of the power spectrum. The indices of the frequency-fluctuation spectra for various heliolatitudes are given in Fig. 3. These were obtained from ULYSSES data acquired in 1995, when the spacecraft was moving from the Southern hemisphere to the equatorial region. For all the data of Fig. 3, the point of closest approach in the line of sight was within a fairly narrow interval of heliocentric distances, from $22\,R_\odot$ to $32\,R_\odot$; therefore, Fig. 3 characterizes the actual heliolatitude dependence of α. This is confirmed by the fact that there is no radial dependence of α for the low-latitude and mid-latitude data of Fig. 2 in the given interval of heliocentric distances. As we can see from Fig. 3, $\alpha \approx 0.6$–0.7 at low latitudes ($\varphi < 50°$), whereas the spectral index shows a tendency to systematically decrease toward the poles. The turbulence remains undeveloped for latitudes above 50°, and only in low-latitude regions ($\varphi < 40°$) is the turbulence close to Kolmogorov.

Examples of flat spectra with low spectral index are shown in Figs. 4a–4c. All three spectra were obtained via an FFT of 2048 frequency-fluctuation measurements for GALILEO radio-sounding data acquired in 1999 and 2000. Figure 4a shows the spectrum derived from measurements in a session on May 11, 2000, when the sounded region was at a heliocentric distance of $R = 11.3\,R_\odot$. The spectrum in Fig. 4b was obtained for April 3, 1999, with $R = 9\,R_\odot$, and the spectrum in Fig. 4c for May 11, 2000, with $R = 6.3\,R_\odot$. We can see that the spectral density depends on the fluctuation frequency only weakly in the frequency interval 10^{-3}–4×10^{-2} Hz. Another important feature of the spectra at small heliocentric distances is the presence of a sharp decrease in the spectral density at high fluctuation frequencies. This feature is most pronounced at the smallest distances of closest approach (Fig. 4c, $R =$

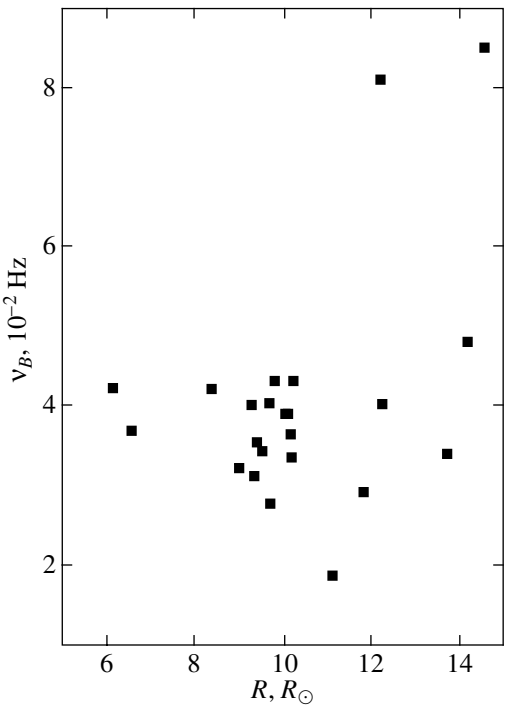


Fig. 5. Break frequency ν_B of the frequency-fluctuation spectra of the GALILEO downlink radio signal at various heliocentric distances.

$6.3\,R_\odot$). As R increases, the break steepness decreases (Figs. 4b, 4c) and the spectrum becomes purely power-law. The break frequency ν_B tends to increase with increasing R: $\nu_B \approx 4 \times 10^{-2}$ Hz at $R = 6.3\,R_\odot$, $\nu_B \approx 4.9 \times 10^{-2}$ Hz at $R = 9\,R_\odot$ and $\nu_B > 5 \times 10^{-2}$ Hz at $R = 11\,R_\odot$.

Figure 5 presents the break frequency as a function of heliocentric distance derived from GALILEO data acquired from 1997 to 2000. The value of ν_B, which is in the range 0.01–0.05 Hz, indeed, tends to increase with R. As we noted above (see Fig. 3), nearly flat spectra with low spectral indices α were also obtained for the 1995 ULYSSES radio-sounding data for high-latitude regions of the solar supercorona. However, in those experiments, the sounded regions were rather far from the Sun ($R > 20\,R_\odot$), and no sharp break was observed in the frequency-fluctuation spectrum.

5. CROSS-CORRELATION PARAMETERS FOR THE FREQUENCY FLUCTUATIONS AND DYNAMICS OF THE DENSITY IRREGULARITIES

Let us spend some time considering some of the dynamic properties of the turbulence that follow from an analysis of the cross-correlation functions for the frequency fluctuations measured simultaneously at widely spaced receiving stations. We chose a set of simultaneous ULYSSES frequency-fluctuation measurements received at Madrid and Goldstone on August 16, 1991 with a total duration of about 2 h for a detailed study of the cross-correlation functions. The session began at 16^h45^m UT and ended at 18^h45^m UT. The heliocentric distance of the point of closest solar approach in the line of sight was about $12\,R_\odot$, the projection of the baseline onto the closest-approach region was $\rho \approx 6500$ km, and the baseline was directed almost radially relative to the Sun. We divided the observational dataset into six 20-min intervals in order to study the variability of the irregularities.

Our analysis shows that the dispersion and shape of the frequency-fluctuation spectra calculated for the short intervals are approximately identical at the two ground stations. At the same time, these parameters vary randomly from one short interval to another: α varies between 0.25 and 1.05, while the rms frequency fluctuations σ_f vary much less, from 0.9 to 1.2 Hz. The velocity of the fluctuation pattern V_{app}, derived from the shift of the cross-correlation maximum (a delay of about 15 s), changed only insignificantly, and was close to 400 km/s, both in the entire observational interval and in the 20-min subintervals for which the maximum level of cross-correlation $K_{\max}$ exceeded the statistical errors.

Of all the correlation and spectral parameters, the correlation at the maximum, $K_{\max}$, was subject to the strongest changes: it varied from statistically insignificant values to 0.4. A comparison of the power spectra and cross-correlation functions obtained in the same time intervals for the joint measurements shows that $K_{\max}$ increases almost linearly with increasing spectral index α. When α is about unity, $K_{\max}$ can reach 0.5, even without filtration of the raw data. When $\alpha < 0.4$, no correlation is observed. Therefore, the efficiency of using joint measurements of the signal frequency at widely spaced stations to find the velocity of motion of the solar wind irregularities decreases at small heliocentric distances ($R < 10\,R_\odot$), where the spectra become flat and the index of the spatial plasma-turbulence spectrum is close to three. We should point out an important origin of the observed variations of the frequency-fluctuation correlation coefficient. As was shown in [10], the correlation depends on the angle ψ between the projection of the baseline onto the plane of the sky and the radial direction. We have $\psi \approx 0$ for simultaneous reception at Madrid and Goldstone, which are at approximately the same latitude. For joint sessions at these stations, a correlation is always observed (at $R > 10\,R_\odot$), even if no high-frequency filtration of the fluctuations is applied.

Figure 6 shows the cross-correlation function for the ULYSSES frequency fluctuations for all the data of August 16, 1991 received at the Madrid and Goldstone stations. The upper curve (Fig. 6a) was obtained for the data averaged over $T_a = 13$ s, and the lower curve for the raw data ($T_a = 1$ s). The correlation coefficient increases strongly with T_a, and the $K_{\max}(\alpha)$ dependence persists even after smoothing. The time shift for which the cross-correlation function reaches its maximum is $\tau_{\max} \approx 15$ s; this corresponds to a velocity of $V_{app} = 430$ km/s. When using the two other combinations of stations, Madrid–Canberra and Goldstone–Canberra, for which ψ is considerably greater than zero, no correlation at all is found in a number of sessions, while, in others, there is a correlation only in some time intervals. As was shown in Section 3, we can determine the rms electron density fluctuations in the circumsolar plasma σ_N using (10)–(11), the measured temporal spectra, and simultaneous widely spaced observations of these fluctuations, which enable us to find the velocity of motion of the fluctuations V_{app} via a cross-correlation analysis. In this method, we determine α, $G_f(\nu_{\max})$, σ_f, and $\nu_{\max}$ from the temporal spectra, and V_{app} from the correlation function, which we assume to be close to the velocity V_{eff} from (8). We obtained estimates of σ_N from GALILEO data near the solar-activity minimum, on January 10 to 30, 1997, for heliocentric distances from 7 $R_\odot$ to 30 $R_\odot$. The value of σ_N varies from 50 to 150 electrons/cm^3. We estimated the relative level of the plasma-density fluctuations using data on the mean density, which was also measured near the solar minimum at low heliolatitudes based on the group delay of the VIKING downlink signals [11]. The value of σ_N/N is approximately constant for heliocentric distances from 7 $R_\odot$ to 30 $R_\odot$, and is 0.12–0.23 for the conditions under which the experiments were carried out: low solar activity, equatorial latitudes, slow solar wind, and a characteristic size of the irregularities close to the outer turbulence scale ($\sim 10^6$ km).

Another radio-physics method was implemented in an experiment with the ULYSSES spacecraft: the solar wind electron density, its variations, and the degree of plasma inhomogeneity were determined from measurements of the differential group delay at two coherent frequencies [12]. The results of this experiment demonstrate that the degree of inhomogeneity σ_N/N depends strongly on the position of the sounded region relative to the neutral line of the heliospheric current system. Our data, which were obtained for the slow solar wind ($V \approx 250$–350 km/s), are consistent with the corresponding results of [12]. The dependence of the degree of inhomogeneity on heliocentric distance is different for the fast solar wind ($V > 500$ km/s): in the region of formation and acceleration of the solar wind, it is quite small (a fraction of a percent), while it increases to 7–10% with increasing R in the region of steady flow.

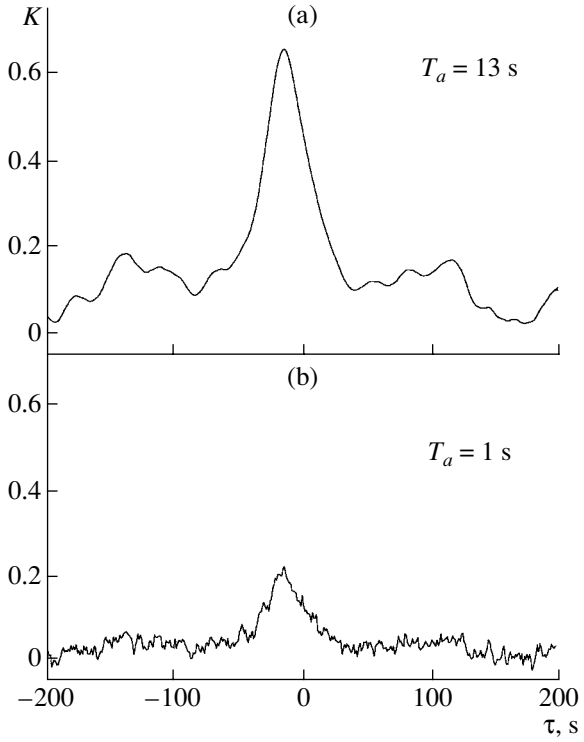


Fig. 6. Cross correlation of the frequency fluctuations recorded at widely spaced stations in Madrid and Goldstone (ULYSSES 1991 data), with fluctuation averaging times of (a) 13 s and (b) 1 s.

6. DISCUSSION

Our results indicate that the power spectrum of the solar wind density fluctuations is a power-law at frequencies exceeding several millihertz. The exponent recalculated for a three-dimensional spectrum at heliocentric distances less than 20 $R_\odot$ is $p \approx 3.0$–3.2, whereas the spectrum becomes steeper at distances exceeding 30 $R_\odot$, with $p \approx 3.6$–3.7. There are indications that the power spectrum of the magnetic-field fluctuations behaves in a similar way. Indeed, the temporal spectra of Faraday-rotation fluctuations at heliocentric distances of 5–10 $R_\odot$ correspond to magnetic-field fluctuation power spectra with $p \approx 3.0$ [13]. Local measurements of the magnetic-field fluctuations on the HELIOS spacecraft at distances of 0.3–1 AU correspond to steeper power spectra with $p \approx 3.6$–3.7 [14].

The as yet unclarified question of the physical processes leading to the abrupt change in the turbulence regime in the transition from the plasma-acceleration

region to the steady-flow region requires a separate quantitative analysis. We will restrict our consideration of this observational fact below to a general qualitative scheme that is consistent with the model proposed in [15] to explain the spectra of smaller-scale fluctuations measured using the scintillation method.

We will assume the solar wind turbulence to be an ensemble of Alfvén and magnetoacoustic waves transported by a plasma flow; the largest energy of the turbulent perturbations is associated with weakly damped Alfvén waves, which, at sufficiently low frequencies, are in the direct-propagation regime. The magnetoacoustic waves to which the plasma-density fluctuations are attributed are generated locally by nonlinear stimulated interactions involving Alfvén waves. In the solar wind acceleration region, where the magnetic field is strong and the Alfvén velocity considerably exceeds the sound speed, the turbulence is weak, and the fastest probable nonlinear process is the interaction of oppositely directed Alfvén and slow magnetoacoustic waves [15]. Here, the energy level of slow exceeds that of fast magnetoacoustic waves, and the quasi-stationary power spectra of the magnetic-field fluctuations (Alfvén waves) and plasma-density fluctuations (slow magnetoacoustic waves) have identical one-dimensional exponents $p_1 = p - 2 = 1$ [15], in full consistency with the frequency-fluctuation and Faraday-rotation measurements. Note that spectra with $p_1 = 1$ result from the drift of Alfvén waves toward low frequencies with conservation of their total number and without cascade processes that transfer turbulent energy to high frequencies. The relative level of the Alfvén waves increases with heliocentric distance due to the smooth plasma inhomogeneity related to the expansion. In addition, the ratio of the sound speed to the Alfvén speed also increases. The combination of these effects results in an increase of the relative contribution of fast magnetoacoustic waves and, beginning from heliocentric distances of 10–$20\,R_\odot$ (according to the estimates of [15]), to the activation of cascade processes that pump energy to high frequencies, with a corresponding change in the turbulence spectrum from a flat spectrum with $p_1 = 1$ to a typical inertial spectrum with a constant spectral-energy flux with $p_1 = 1.5$–1.7 ($p_1 = 3/2$ for an Iroshnikov–Kraichnan spectrum, and $p_1 = 5/3$ for a Kolmogorov spectrum).

As is shown in Section 4, near a solar minimum, the region where the turbulence makes a transition to a mode with energy cascading to higher frequencies is located farther from the Sun for the high-latitude, fast solar wind than for the low-latitude, slow wind. This can be explained if there are stronger magnetic fields in the regions above polar coronal holes and, as a consequence, lower ratios of the sound speed to the Alfvén velocity and lower relative levels of turbulence energy, so that the cascade processes are activated somewhat farther from the Sun than in the slow solar wind.

Turbulence is weak in the acceleration region, possibly due to an increase in the relative amplitude of the Alfvén waves that generate the magnetoacoustic waves. The frequency-fluctuation spectra with a sharp cutoff observed in near-solar regions (Section 4)—in particular, the increase in the cutoff frequency with heliocentric distance—testify to an absence of cascade processes. Indeed, frequencies of the order of 10^{-2} Hz are considerably lower than the frequencies corresponding to the characteristic scales of the magnetized plasma (the gyroradius or the inertial scale of the ions), on which we could expect strong linear absorption or a change in the behavior of the wave dispersion. If we assume that the cutoff frequency results from the pumping of energy to smaller scales, it would have displayed the opposite radial dependence. We can suppose that the cutoff in the turbulence spectrum is initially formed in the middle corona, near the coronal temperature maximum, due to the linear absorption of Alfvén waves propagating outward. If the turbulence level is anomalously low in this case, then nonlinear processes outside the corona have no time to smooth the cutoff in the power spectra of the waves leaving the corona.

According to the estimates of [16], the characteristic nonlinear increments for the evolution of the angular distribution of the wave vectors are small compared to the reciprocal timescale for the establishment of the power spectrum with $p_1 = 1$. Therefore, the strong anisotropy of the angular distribution of the waves, with the prevalence of transverse wave vectors, which is initially formed in the primary region where the inhomogeneities are generated at the base of the corona, will persist to a distance of about $10\,R_\odot$. Therefore, we expect that the change in the turbulence regime will also be accompanied by a change in the anisotropy of the irregularities: from strongly elongated in the radial direction at small heliocentric distances to weakly anisotropic in the region of the steady solar wind.

Note that, in principle, there is an alternative explanation of the collected spectral data, based on the idea that nonlinear interactions between the irregularities are completely suppressed in the solar wind acceleration region, and that the flat power spectra of the magnetic-field and plasma-density fluctuations with $p_1 = 1$ either reflect the initial conditions in the zone in which the irregularities form in the lower corona [17, 18] or are related to some universal mechanisms for the formation of known spectra, such as flicker noise. Without considering this possibility in detail, we note that, in this case, as well, the change

in the turbulence regime can only be due to the activation of cascade processes that are absent in regions that are closer to the Sun. Here, it is interesting that spectra with $p_1 = 1$ can also be formed by the local generation of waves with a linear dispersion law due to instabilities with increments that are proportional to the frequency, as a result of the pumping of energy to lower frequencies during the stimulated scattering of waves on ions [19].

Finally, let us consider the data of Section 5 relating to the dynamics of the spectral parameters and the changes in the spatial pattern of the fluctuations during its motion between the radio paths. The fast variations of the shape of the frequency-fluctuation spectrum indicate that the variation timescale, which was determined by the duration of the short samples and was about 20 min, is shorter than or comparable to the timescale for the transport of irregularities with scales comparable to the outer turbulence scale. This is consistent with the results of [6], which indicate that the convection timescale for the outer turbulence scale is $\geq$1 h.

The observed decorrelation between the fluctuations measured at different stations could be due to (1) motions of the irregularities transverse to the baseline between ground stations, (2) a spread in the instantaneous velocity in the line of sight due to the presence of flows with different velocities, and (3) intrinsic changes in the irregularities during their motion between the two radio-signal paths. In all three cases, the decorrelation of the frequency fluctuations is amplified with an increase in the ratio of the longitudinal baseline to the irregularity scale. As the index of the fluctuation spectrum increases, the relative contribution of large-scale irregularities also increases; this explains the corresponding increase in the cross-correlation. Filtration of high-frequency fluctuations should produce a similar result, as is observed in the actual experiments.

The first of the origins of the decorrelation listed above does not play an important role, since we have analyzed a specially selected set of variations, for which the angle between the baseline and the radial direction was small. To describe the other two possible origins, we can introduce a certain rate of random variability w. Since the chosen filtration timescale, 13 s, is similar to the delay of the fluctuations, the main contribution to the cross-correlation comes from irregularities on scales close to the radial baseline projection. The maximum cross-correlation for the filtered data is fairly high ($K_{\max} \geq 0.6$), but is, at the same time, appreciably different from unity; thus, we conclude that w is comparable to, but does not exceed, the solar wind speed V ($w \lesssim V$).

The variability of the fluctuation pattern at the considered heliocentric distance, about $12\,R_\odot$, is probably related to intrinsic changes in the irregularities, since, due to the fairly small thickness of the efficiently modulating layer of the medium, the instantaneous velocity spread in the line of sight cannot be comparable to the average solar wind speed. A similar conclusion for smaller scale inhomogeneities (100 km) in a comparable interval of heliocentric distances was drawn in [20]. The physical explanation that the variations of the irregularities during their motion (the "seething" pattern in the terms of [20]) play the main role, and not the instantaneous velocity spread in the line of sight (the wind pattern), is that, at distances $\sim 10\,R_\odot$ from the Sun, the propagation velocities of MHD waves in the solar wind plasma are comparable to the average solar wind speed.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 03-02-16262 and 04-02-17332), the Program of the Division of Physical Sciences of the Russian Academy of Sciences on the Solar Wind (Generation and Interaction with the Earth and other Planets), and the Federal Program in Astronomy of the Ministry of Education and Science of the Russian Federation.

REFERENCES

1. V. I. Vlasov, I. V. Chashei, V. I. Shishov, and T. D. Shishova, Geomagn. Aéron. **19**, 401 (1979) [Geomagn. Aeron. **19**, 269 (1979)].
2. R. Woo and J. W. Armstrong, J. Geophys. Res. **84** (A12), 7288 (1979).
3. A. I. Efimov and O. I. Yakovlev, *Electromagnetic Waves in the Atmosphere and Space* (Nauka, Moscow, 1986) [in Russian].
4. M. K. Bird and P. Edenhofer, *Remote Sensing Observations of the Solar Corona,* in *Physics of the Inner Heliosphere*, Ed. by R. Schwenn and E. Marsch (Springer-Verlag, Heidelberg, 1990), p. 13.
5. R. Wohlmuth, D. Plettemeier, P. Edenhofer, *et al.*, Space Sci. Rev. **97**, 9 (2001).
6. A. I. Efimov, I. V. Chashei, L. N. Samoznaev, *et al.*, Astron. Zh. **79**, 640 (2002) [Astron. Rep. **46**, 579 (2002)].
7. N. A. Armand, A. I. Efimov, and O. I. Yakovlev, Astron. Astrophys. **183**, 135 (1987).
8. I. V. Chashei, V. I. Shishov, M. Kojima, and H. Misawa, J. Geophys. Res. **105** (A12), 27409 (2000).
9. I. V. Chashei, A. I. Efimov, V. K. Rudash, and M. K. Bird, Astron. Zh. **77**, 713 (2000) [Astron. Rep. **44**, 634 (2000)].
10. N. A. Armand, A. I. Efimov, L. N. Samoznaev, *et al.*, Radiotekh. Elektron. (Moscow) **48**, 1058 (2003).
11. D. O. Muhleman and J. D. Anderson, Astrophys. J. **247**, 1093 (1981).

12. R. Woo, J. V. Armstrong, M. K. Bird, and M. Patzold, Geophys. Res. Lett. **22**, 329 (1995).
13. I. V. Chashei, A. I. Efimov, L. N. Samoznaev, M. K. Bird, and M. Patzold, Adv. Space Res. **25**, 1973 (2000).
14. K. U. Denskat, H. J. Beinroth, and F. M. Neubauer, J. Geophys. **54**, 60 (1983).
15. I. V. Chashei and V. I. Shishov, Astron. Zh. **60** (4), 594 (1983) [Sov. Astron. **27** (4), 346 (1983)].
16. I. V. Chashei, Astron. Zh. **78**, 761 (2001) [Astron. Rep. **45**, 659 (2001)].
17. W. H. Matthaeus and M. L. Goldstein, Phys. Rev. Lett. **57**, 495 (1986).
18. A. A. Ruzmaikin, B. E. Goldstein, E. J. Smith, *et al.*, in *Solar Wind Eight*, Ed. by D. Winterhalter *et al.* (AIP Press, Woodbury, NY, USA, 1996); AIP Conf. Proc. **382**, 225 (1996).
19. B. B. Kadomtsev, *Collective Phenomena in Plasmas* (Nauka, Moscow, 1976) [in Russian].
20. R. D. Ekers and L. T. Little, Astron. Astrophys. **10**, 310 (1971).

Translated by G. Rudnitskiĭ

Astronomy Reports, Vol. 49, No. 6, 2005, pp. 495–499. Translated from Astronomicheskiĭ Zhurnal, Vol. 82, No. 6, 2005, pp. 555–560.

Modern Progress in Solar Paleoastrophysics and Long-Range Solar-Activity Forecasts

M. G. Ogurtsov

Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia

Received March 25, 2004; in final form, December 3, 2004

Abstract—Decade-averaged Wolf numbers are reconstructed for the time interval 8005 B.C.–1945 A.D. using radiocarbon data derived from tree rings. Comparisons of other paleoastrophysical reconstructions of solar activity with this temporal series verify its validity and reliability. A prediction of the mean solar activity for the next forty years is made using these reconstructed Wolf numbers. It is likely that the mean solar activity during 2005–2045 will be lower than the activity of recent decades. This conclusion is compared with the long-term predictions proposed by other researchers. The prospects for paleoastrophysical predictions for the long-term variations of solar activity in the future are discussed. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The forecasting of solar activity is both an interesting scientific problem and very important for the life of present-day humankind. The Sun's activity governs a number of processes in the Earth's atmosphere and environment, thus affecting numerous aspects of modern civilization. For example, the powerful corpuscular radiation emitted during solar flares and coronal ejections disrupts radio communications and disables spacecraft equipment, as well as affecting meteorological processes (the transparency of the Earth's atmosphere, atmospheric circulation, and the temperature distribution in the atmosphere) [1]. Plausible effects exerted on the Earth's climate by Galactic cosmic rays modulated by solar activity have been intensively discussed [2–4]. Possible relations of long-term climatic changes to corresponding variations in the solar luminosity are discussed in [5, 6]. Numerous indications of the reality of solar–climatic relations confirm that solar activity is one of the most important factors determining the state of the lower terrestrial atmosphere.

The forecasting of changes in solar activity is becoming more and more important for meteorology and climatology. Such predictions are especially important from the standpoint of the global warming that has been observed over the past century and has growing implications for the entire existence of humankind. Although global warming is usually associated with the greenhouse effect, there are indications that increases in solar activity, along with increases in the carbonic acid content, are also factors responsible for the growth in temperature over the 20th century (see, for example, [7–9]). Thus, long-term (over tens of years and more) forecasting of solar activity becomes useful for making predictions about the climate in the future.

However, such forecasting is not simple, since it requires information on solar activity over hundreds of years or more; i.e., on the history of our star over numerous centuries. It is obvious that targeted observations of the Sun using various instruments cannot supply us with this information. The longest series of solar data, namely data on numbers of sunspot groups, covers only the four last centuries, with the trustworthy part of the data series being even shorter. Analysis of indirect indicators of solar activity studied by solar paleoastrophysicists cover numerous centuries and millennia, and can thereby clarify prospects for long-range forecasting.

Paleoastrophysics is concerned with astrophysical phenomena whose signals reached the Earth before the time of instrumental astronomy. The basic ideas of paleoastrophysics were formulated by B.P. Konstantinov and G.E. Kocharov in the mid-1960s. It was shown that we can obtain reliable information on a number of astrophysical phenomena of the distant past, including long-term variations in the intensity of solar and Galactic cosmic rays, magnetic and flaring activity of the Sun, supernova outbursts, and gamma-ray bursts, using various natural archives. This paper presents a brief review of long-term reconstructions of solar activity obtained recently. The mean solar activity for the first half of the 21st century is predicted using the Wolf numbers reconstructed from radiocarbon data for 8005 B.C.–1945 A.D.

1063-7729/05/4906-0495$26.00

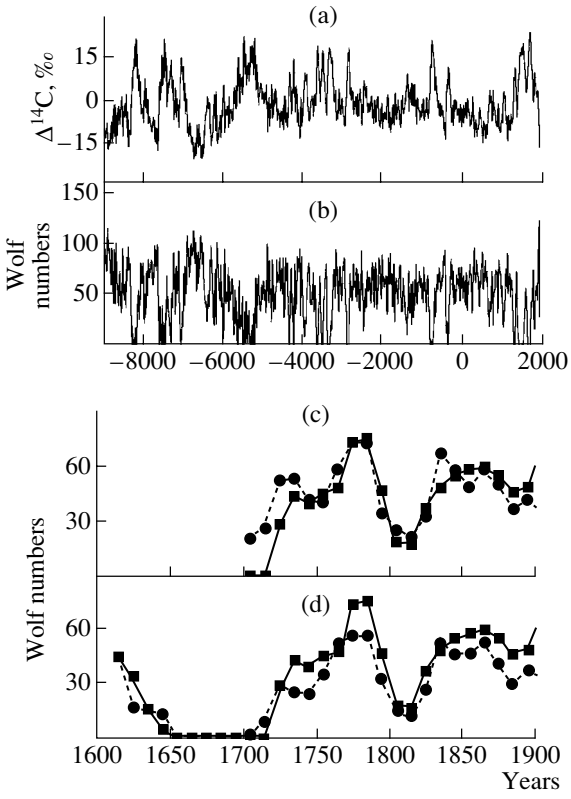


Fig. 1. (a) The ^{14}C concentration in tree rings from the northwest US [14] after removing the long-term trend. (b) Decade-averaged Wolf numbers reconstructed from the radiocarbon data. (c) Decade-averaged Wolf numbers reconstructed from the radiocarbon data for 1700–1900 (solid curve) and detected instrumentally (dashed curve). (d) Decade-averaged Wolf numbers reconstructed from the radiocarbon data for 1600–1900 (solid curve) and numbers of sunspot groups detected instrumentally (dashed curve).

2. MODERN PALEO-RECONSTRUCTIONS OF THE SOLAR ACTIVITY

The study of concentrations of cosmogeneous isotopes in natural archives is a one of the basic methods of experimental paleoastrophysics. Cosmogeneous radiocarbon ^{14}C and radioberyllium ^{10}Be originate in the Earth's stratosphere and upper troposphere due to the effect of energetic Galactic cosmic rays (GCR) modulated by the solar activity. Molecules containing ^{14}C and ^{10}Be that are formed are rapidly oxidized to $^{14}CO_2$ and ^{10}BeO. The beryllium oxide is then captured by aerosols, washed out by precipitation, and preserved in polar ices and sea-bottom deposits. $^{14}CO_2$ is involved in a chain of geophysical and geochemical processes forming the global carbon cycle, and is finally fixed in tree rings [10–12]. Thus, the concentrations of ^{10}Be in ices and of radiocarbon in tree rings depend on solar activity. Studies of various historical chronicles (sunspot observations with the unaided eye or auroral observations) provide another method for obtaining information on the solar activity in the past.

The study of past solar activity using radiocarbon data has been carried out since the end of the 1960s, and has already revealed numerous important features (long maxima and minima, long-term periodicity, etc.). Note, however, that ^{14}C describes the solar activity in a nonlinear fashion; namely, the carbon cycle is similar to a high-frequency filter that attenuates the short-term variations in the atmospheric radiocarbon production much more strongly than the long-term variations. Therefore, the statistics of the solar activity recorded in the radiocarbon record may be seriously garbled. We can avoid such distortion by employing the Wolf numbers reconstructed from the radiocarbon series using some carbon-cycle model, rather than the original ^{14}C data.

We predict long-term solar activity based on a reconstruction of the decade-averaged Wolf numbers for 8005 B.C.–1945 A.D. performed by Ogurtsov [13] using data on $\Delta^{14}C$ for tree rings for the northwest US [14] and a five-reservoir model of the carbon cycle. This model was thoroughly studied by Soviet researchers, who showed that the model adequately describes the ^{14}C variations for time periods exceeding several decades (see, for example, [15, 16]). The Wolf numbers were calculated from the radiocarbon data by solving a set of integro-differential equations, presented in [13]. The model parameters (time periods for exchanges between reservoirs and the initial radiocarbon concentrations) were adjusted to obtain the best agreement with the Wolf numbers detected instrumentally.

The reconstructed Wolf (RW) numbers obtained using this technique are presented in Fig. 1b, and the original radiocarbon series in Fig. 1a (after removal of the long-term trend). Figure 1c compares the RW numbers with those detected instrumentally, averaged over 11-year periods, and interpolated for decades in 1700–1900 (before the start of anthropogenic carbon emission, which strongly distorts the atmospheric ^{14}C concentration). We can see that these series agree well, with the correlation coefficient being 0.82—appreciably higher than the correlation coefficient for the Wolf numbers and the original $\Delta^{14}C$ series, −0.62. The correlation coefficient for the RW series and similarly averaged and interpolated numbers of sunspot groups [17] reaches 0.9 for 1615–1895 (Fig. 1d).

Figure 2 shows the temporal RW series along with other recent paleo-reconstructions of solar activity for the last millennium.

Figure 2b shows the Wolf numbers reconstructed by Stuiver and Quay [18] (scanned and digitized) based on the same radiocarbon series as for our RW

series, but using another model for the carbon cycle, Fig. 2c the reconstruction of the solar activity performed in [19] based on data on the ^{10}Be concentration in Antarctic ices, and Fig. 2d the reconstructed Wolf numbers obtained by Nagovitsyn [20] using the data of Schove [21], who had estimated the periods of maxima and minima of quasi-11-year solar cycles and the approximate Wolf numbers at the maxima starting from 1090 (mainly using historical accounts of auroras). Observations of sunspots with the unaided eye carried out by Eastern astronomers and collected by Wittmann and Xu [22] and averaged over 45-year periods are presented in Fig. 2e. Figure 2 demonstrates that the various paleo-indicators of solar activity obtained from such different sources as the cosmogeneous isotopes stored in terrestrial archives and historical accounts of auroras and sunspots are generally in reasonably good agreement. The global maxima and minima of the solar activity are fairly clear in most of these temporal series. The correlation coefficients between the various reconstructions are 0.52–0.80. This supports the idea that modern paleoastrophysics can supply us with trustworthy information, at least about the qualitative features of the temporal behavior of solar activity.

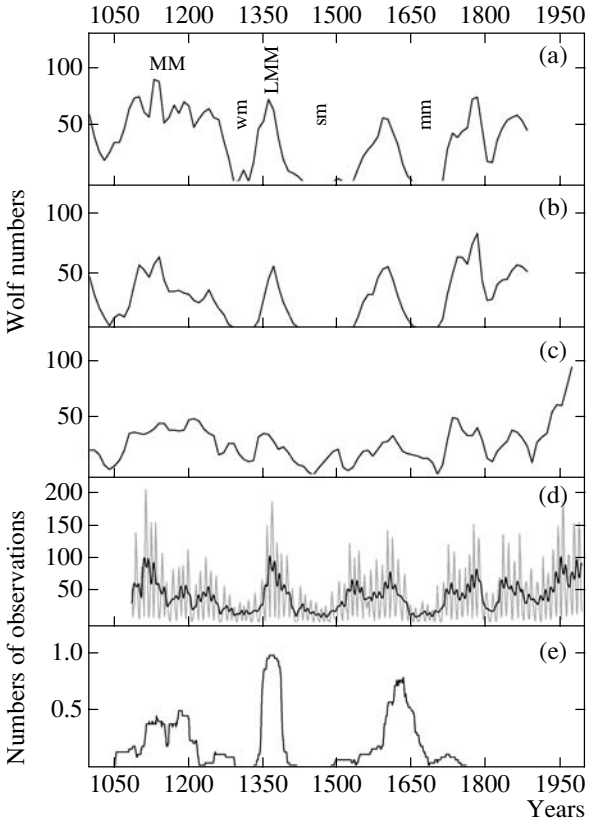


Fig. 2. (a) Wolf numbers reconstructed in [18] from radiocarbon data [14]. (b) Wolf numbers reconstructed in [19] from radiocarbon data [14]. (c) Wolf numbers reconstructed in [20] from the ^{10}Be concentration in Antarctic ices. (d) Wolf numbers reconstructed in [20] from the data of [21]. (e) Observations of sunspots by the unaided eye averaged over 45-year intervals (MM is the Medieval maximum, LMM the last Medieval maximum, wm the Wolf minimum, sm the Sperer minimum, and mm the Maunder minimum) [22].

3. PREDICTIONS OF THE MEAN SOLAR ACTIVITY FOR THE COMING DECADES

We used the RW series, which is the longest reconstructed series currently available, to forecast the mean solar activity for the coming decades. We used the technique of Sugihara and May [23] for nonlinear forecasting. In this method, a sequence of d-dimensional vectors is extracted from the initial one-dimensional temporal series, $x_i = x(t_i)$ $(i = 1, \ldots, n)$, using the procedure of Packard and Takens [24, 25]:

$$\mathbf{X}_i = X(t_i) = \{x(t_i), x(t_i + \tau), \quad (1)$$
$$x(t_i + 2\tau), \ldots, x(t_i + (d-1)\tau)\},$$
$$i = 1, \ldots, (n - d\tau).$$

This d-dimensional series $X(t_i)$ $(i = 1, \ldots, n)$ is divided into two (usually roughly equal) portions, namely $Y(t_i) = X(t_i)$ $(i = 1, \ldots, n/2)$ and $Z(t_i) = X(t_{i+n/2})$ $(i = 1, \ldots, n/2)$. The first portion forms a "library" of trajectories of the dynamic system and is used as a "known past" to forecast the "future" for the second portion. For instance, if we wish to forecast the future $x(t_{k+p})$ (after p time steps) for the point $x_k = x(t_k)$ in the second portion of the series, we first determine the d-dimensional point $Z(t_l)$ containing x_k and select some nearest neighbors of $Z(t_l)$ among the points $Y(t_i)$, then move this set of nearest neighbors by p steps into the future, and determine its geometric center, Y_c. The corresponding coordinate Y_c is taken to be the forecast $x(t_{k+p})$.

The RW series was modified for the forecasting. It remained unchanged before 1900, but the true Wolf numbers averaged over 11-year periods and interpolated over decades were used after 1900. We call this series of 1096 points the RW1 series. To obtain a long-term forecast for the decade-averaged solar activity, the first 996 points of the RW1 series were used as a database of trajectories of the dynamic system. The number of nearest neighbors and the parameters d and τ were selected using an empirical procedure providing the minimum error for the forecast (see [26] for details). Figure 3a presents the forecast of the decade-averaged solar activity for the next forty years (to 2045) obtained using the author's own code, while Fig. 3b shows the forecast obtained using the TISEAN code [27].

The accuracy is relatively low, and the forecast

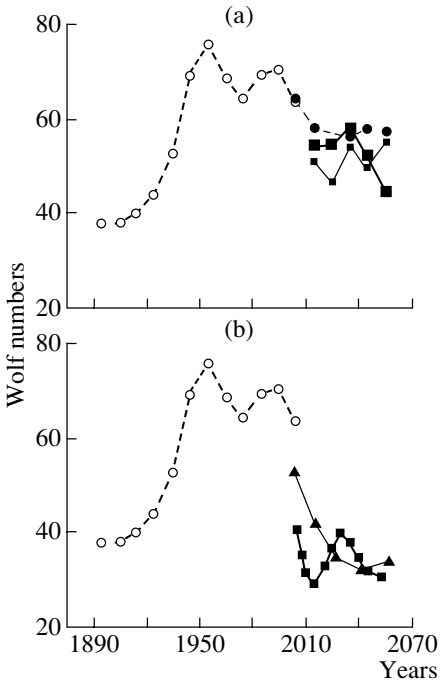


Fig. 3. Predictions of the mean solar activity for the next forty years based on the reconstructed Wolf numbers (8005 B.C.–1945 A.D.). (a) The bold solid curve with large squares and thin solid curve with small squares show forecasts made by the author using his own code, for $d = 3$ and the five nearest neighbors and for $d = 4$ and the four nearest neighbors, respectively. The thin dashed curve with filled circles shows the forecast produced by the TISEAN code, and the bold dashed curve with hollow circles show the observed Wolf numbers (averaged over 11-year periods and interpolated over decades). (b) The bold solid curve with squares shows the forecast presented in [28], the thin solid curve with triangles shows the forecast of [29], and the bold dashed curve with hollow circles is the same as in the upper plot.

uncertainty increases from ≈16 (2015) to almost 25 (2045). In other words, the error is appreciably smaller than the series variance (which is 26) only for predictions over 10–20 yrs (1–2 steps forward). However, despite the low accuracy of the predictions of the mean solar activity for each single point, the net tendency for a decrease in activity in 2005–2045 is quite clear. Thus, this analysis of the RW1 series suggests that the mean solar activity in the first half of this century will probably be lower than it is now. The predictions of solar activity obtained in [28, 29] using more classical linear methods are shown in Fig. 3b. The prediction of Nagovitsyn and Ogurtsov [28] was obtained via multiple-scale cloning using the series [14]. Miletskii [29] used the Wolf-number reconstruction of [20] and employed a linear model. Figure 3b shows that the predictions obtained in [28, 29] also indicate a decrease in solar activity during the first half of the 21st century, though it is more pronounced than is predicted by the present work. Komitov and Kaftan [30] also predicted a significant decrease in solar activity for the 21st century using Schove's series as a paleo-indicator.

4. CONCLUSIONS

We have analyzed a radiocarbon reconstruction of the Wolf numbers reflecting the Sun's behavior over the last 10 000 yrs to predict the mean solar activity over the next several decades. The accuracy of the long-term forecasts is fairly low, and the calculations showed that only predictions over a single step (ten years) forward are reasonably accurate. Nevertheless, the results indicate that, on average, the solar activity in the first half of this century will be lower than it is now. This conclusion agrees with other predictions of [28–30], which likewise used solar paleoastrophysics to deduce that there is likely to be a decrease in solar activity in the first half of the 21st century. In summary, a paleoastrophysical approach to long-term forecasting of solar activity seems promising and deserving of further study.

ACKNOWLEDGMENTS

The author is grateful to N.G. Makarenko, Yu.A. Nagovitsyn, and V.G. Ivanov for numerous helpful discussions and suggestions. This work was supported by the scientific exchange activities of the Russian and Finnish Academies of Sciences (project no. 16), the Program of the Presidium of the Russian Academy of Sciences on Nonstationary Processes in Astronomy, INTAS (project no. 2001-0550), and the Russian Foundation for Basic Research (project nos. 03-04-48769, 03-02-17505, and 04-02-17560).

REFERENCES

1. M. I. Pudovkin, Soros. Obraz. Zh. **10**, 106 (1996).
2. N. Marsh and H. Svensmark, Phys. Rev. Lett. **85**, 5004 (2000).
3. E. Palle and C. J. Butler, J. Atmos. Sol.-Terr. Phys. **64**, 327 (2002).
4. Yu. I. Stozhkov, V. I. Ermakov, and P. E. Pokrevskii, Izv. Ross. Akad. Nauk, Ser. Fiz. **65**, 406 (2001).
5. G. C. Reid, J. Geophys. Res. **96** (D2), 2835 (1991).
6. J. Lean, L. Beer, and R. Bradley, Geophys. Res. Lett. **22**, 3195 (1995).
7. E. Friis-Christensen and K. Lassen, Science **254**, 698 (1991).
8. J. Beer, W. Mende, and R. Stellmacher, Quat. Sci. Rev. **19**, 403 (2000).
9. J. Lean and D. Rind, J. Clim. **11**, 3069 (1998).
10. V. A. Dergachev, Izv. Ross. Akad. Nauk, Ser. Fiz. **59**, 53 (1995).
11. M. Stuiver, T. F. Braziunas, B. Becker, and B. Kromer, Quat. Res. **35**, 124 (1991).

12. G. E. Kocharov, M. G. Ogurtsov, and S. L. Tsereteli, Astron. Zh. **80**, 1144 (2003) [Astron. Rep. **47**, 1054 (2003)].
13. M. G. Ogurtsov, Sol. Phys. **220**, 93 (2004).
14. M. Stuiver, P. J. Reimer, and T. F. Braziunas, Radiocarbon **40**, 1127 (1998).
15. V. A. Dergachev and A. V. Stupneva, Preprint No. 491, Fiz.-Tekh. Inst. im. Ioffe (Ioffe Physicotechnical Inst., USSR Acad. Sci., Moscow, 1975).
16. V. M. Ostryakov, *Abundances of Isotopes in Surrounding the Environment and Astrophysical Phenomena*, Ed. by G. E. Kocharova (Leningrad, 1984), p. 25 [in Russian].
17. D. V. Hoyt and K. H. Schatten, Sol. Phys. **179**, 189 (1998).
18. M. Stuiver and P. D. Quay, Science **207**, 11 (1980).
19. I. G. Usoskin, K. Mursula, S. Solanki, *et al.*, Astron. Astrophys. **413**, 745 (2004).
20. Yu. A. Nagovitsyn, Pis'ma Astron. Zh. **23**, 851 (1997) [Astron. Lett. **23**, 742 (1997)].
21. D. J. Schove, *Sunspot Cycles* (Hutchinson Ross, Stroudsburg, Pennsylvania, 1983).
22. A. D. Wittmann and Z. Xu, Astron. Astrophys., Suppl. Ser. **70**, 83 (1987).
23. G. Sugihara and R. M. May, Nature **344**, 734 (1990).
24. N. H. Packard, J. P. Grutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
25. F. Takens, *Lecture Notes in Mathematics* (Springer, Berlin–Heidelberg–New York, 1981), Vol. 898, p. 366.
26. M. G. Ogurtsov, Sol. Phys. (2005, in press).
27. R. Hegger, H. Kantz, and T. Schreiber, Chaos **9**, 413 (1999).
28. Yu. A. Nagovitsyn and M. G. Ogurtsov, *Climatic and Ecological Aspects of Solar Activity* (Gos. Astron. Obs. Ross. Akad. Nauk, St. Petersburg, 2003), p. 321.
29. E. V. Miletskiĭ, *Climatic and Ecological Aspects of Solar Activity* (Gos. Astron. Obs. Ross. Akad. Nauk, St. Petersburg, 2003), p. 305.
30. B. P. Komitov and V. I. Kaftan, Geomagn. Aéron. **43**, 592 (2003) [Geomagn. Aeron. **43**, 553 (2003)].

Translated by V. Badin