A Method for Estimating the Distances of Extragalactic Radio Sources with Jets Using VLBI

B. V. Komberg

Astro Space Center, Lebedev Physical Institute, Moscow, Russia
Received July 25, 2003; in final form, March 15, 2004

Abstract—A method is proposed for estimating the distances to extragalactic radio sources using spectroscopic data on the sizes of the broad-line regions of quasars and radio galaxies (R_{BL}) and VLBI data on the transverse angular sizes of the radio jets emerging from their nuclei ($\theta_{\perp}^{j}$). This approach is based on the fact that observational data on the transverse linear sizes of radio jets ($l_{\perp}^{j}$) associated with the nearest radio sources and theoretical concepts about the formation of radio jets in the central regions of a magnetized accretion disk suggest that $l_{\perp}^{j}$ and R_{BL} should be approximately equal. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

(1) Several methods have been proposed for the estimation of the distances to extragalactic objects whose spectra do not display strong lines that can be used to derive their redshifts, or for the estimation of the Hubble expansion of the Metagalaxy. These are all based on the assumed existence of some standard quantities (luminosity, surface brightness, velocity, dimensions, etc.)[1] or on the time evolution of these quantities. For example, Faber and Jackson [2] constructed a relationship between the luminosities of elliptical galaxies and the velocity dispersions for the stars in them, σ_v, which were estimated from the widths of absorption lines arising in the stellar atmospheres: $L_{EG} \sim \sigma_v^4$. Tully and Fisher [3] obtained an analogous relationship for spiral galaxies based on the width of the 21-cm neutral-hydrogen line, which characterizes the rotational speed of the gas in the galaxy and the galaxy's inclination: $L_{SG} \sim (v_{\max} \sin i)^2$. Relationships between the luminosities of Seyfert-galaxy nuclei in the X-ray ($L_{0.5-4.5\ \mathrm{keV}}$) and radio ($L_{400\ \mathrm{MHz}}$) and the widths of the Balmer lines in their spectra were obtained in [4]:

$$L_{\mathrm{XR}} \sim (\Delta V_{\mathrm{H}(\beta,\alpha)}^0)^2 \quad \text{and} \quad L_{400\ \mathrm{MHz}} \sim (\Delta V_{\mathrm{H}(\beta,\alpha)}^0)^2.$$

It was elucidated that there exists a similar dependence between the luminosities and widths of these lines [5]:

$$L_{\mathrm{H}\beta} \sim (\Delta V_{\mathrm{H}\beta}^0)^2.$$

(2) The quadratic dependence of the line luminosities on the line widths (i.e., the velocity dispersions) provides evidence that the broad lines are formed in disklike circumnuclear structures.[2] This conclusion has recently been supported by data on the masses of the black holes in the nuclei of nearby galaxies and quasars ($M_{\bullet}$). A number of studies (see, for example, [7−10]) have demonstrated the relations for supermassive black holes

$$M_{\bullet} = 10^8 M_{\odot}(\sigma_*^{\mathrm{bul}}/200\ \mathrm{km/s})^4, \tag{1}$$

where σ_*^{bul} is the velocity dispersion for stars in the bulge of the host galaxy and $M_{\odot}$ is the mass of the Sun, and

$$M_{\bullet} = 10^8 M_{\odot}(\Delta V_{1/2}(\mathrm{H}\beta)/3000\ \mathrm{km/s})^2 \tag{2}$$
$$\times (L_{5100\ \text{Å}}/10^{44}\ \mathrm{erg/s})^{0.5},$$

where $\Delta V_{1/2}(\mathrm{H}\beta)$ is the half-width of the Hβ line and $L_{5100\ \text{Å}}$ is the optical luminosity. Relations (1) and (2) for galaxies with various luminosities (masses) yield values for $M_{\bullet}$ from 10^5 to $3 \times 10^9 M_{\odot}$.

Since we have from the definition of the Keplerian velocity $V^2 = GM_{\bullet}/R$, it follows that $M_{\bullet} \sim V^2 R$; comparing this relation with (2), it is easy to see that $R \sim L^{0.5}$. That is, the size of the broad-line region (BLR) in the circumnuclear disk, R_{BL}, is proportional

---

[1] In [1], it is proposed to estimate the distances to ultraluminous IR galaxies based on observations at 450 and 850 μm, since the spectra of these galaxies are "standard," having their maxima at $\lambda_0 \approx 100\ \mu$m. The accuracy of this method is ~30%.

[2] Although there are also other points of view; see, for example, [6].

to the square of the luminosity of the ionizing radiation from the nucleus.

(3) Various dependences that can, in principle, be used to estimate R_{BL} can be found in the literature. For example, according to [11−13],

$$R_{\mathrm{BL}}(\text{light days}) \tag{3a}$$
$$= (26.1 \pm 3.6)(\lambda L_{3000\,\text{Å}}/10^{37}\ \text{W})^{0.5}$$

or

$$R_{\mathrm{BL}}(\text{light days}) \tag{3b}$$
$$= (35.5 \pm 4.9)(\lambda L_{5100\,\text{Å}}/10^{37}\ \text{W})^{0.5}.$$

The relation

$$M_{\bullet} \tag{4}$$
$$= 10^5 R_{\mathrm{BL}}(\text{light days})(\Delta V_{1/2}(\mathrm{H}\beta)/10^3\ \text{km/s})^2 M_{\odot}$$

is given in [14], which, together with (1), can be used to obtain the approximate expression

$$R_{\mathrm{BL}}(\text{light days}) \tag{5}$$
$$= 10^3 (\sigma_*^{\mathrm{bul}}/200\ \text{km/s})^4/(\Delta V_{1/2}(\mathrm{H}\beta)/10^3\ \text{km/s})^2.$$

Assuming $\Delta V_{1/2}(\mathrm{H}\beta) = 10^4$ km/s [15] and $\sigma_*^{\mathrm{bul}} = 200$ km/s, we obtain

$$R_{\mathrm{BL}} \approx \text{tens of light days.}$$

This is in good agreement with estimates of R_{BL} derived from the time delays (Δt) between variations in the ionizing flux and the onset of variations in the line profiles: $R_{\mathrm{BL}} \approx \Delta tc$ (see, for example, [16−18]).

For nearby galaxies or quasars, the size $R_{\mathrm{BL}} \approx 10$ light days ≈ 0.01 pc can already be resolved by current optical interferometers, such as the VLTI. For example, it should be possible to resolve R_{BL} for active galactic nuclei (AGN) with $z < 0.4$ when $\Delta\theta = 0.1$ milliarcsecond (mas) and for AGN with $z = 2$ when $\Delta\theta = 0.01$ mas [19, 20]. This raises the possibility of estimating the distances to AGN using the values of R_{BL} derived from observations [formula (5)] and data on $\Delta\theta_{\mathrm{BL}}$ obtained using optical interferometers.

(4) Even more promising possibilities of this sort are opened if we consider radio observations, since ground-based VLBI, and all the moreso space VLBI, can attain angular resolutions of $10^{-4}-10^{-5}$ arcsec at centimeter wavelengths (for example, in the Russian project "RadioAstron" [21]). With such angular resolutions, it becomes possible to obtain information about the angular diameters of the bases of the radio jets ejected by AGN. Indeed, the transverse angular sizes of the radio jets ($\theta_{\perp}^{j}$) of the most nearby radio galaxies (Cygnus A, Perseus A, and Virgo A) and some Seyfert galaxies (NGC 4151, for example) have been measured using VLBI and have proven to be comparable to $R_{\mathrm{BL}} =$ tens of light days [22−25]. These small radio sizes provide hope that evolution will not play an appreciable role in the cosmological dependence $\theta_{\perp}^{j}(z)$, in contrast to the situation with the angular sizes of the extended radio structures, which are subject to strong evolutionary effects over the course of their long lifetime ($\sim 10^8-10^9$ yrs). This fact has already been considered in a number of studies, such as [26, 27], where the dependence $\theta(z)$ is constructed using the smallest components of the radio jets resolved with VLBI.

In addition, the transverse size of the radio jet near its base ($l_{\perp}^{j}$) should not depend strongly on the angle between the line of sight and the direction of the jet. This is not true of the sizes of radio components along the jet, which must be corrected for the effect of projection.

2. RELATIONSHIP BETWEEN THE CORE RADIO PROPERTIES AND THE SPECTRAL CHARACTERISTICS OF AGN

(1) The existence of certain relationships between the radio properties of the VLBI core components and the spectral characteristics of AGN has already been noted in the literature. In particular, there have been attempts to identify various types of AGN on diagrams analogous to the Hertzsprung−Russell diagram for stars (see, for example, [28−30]). Indeed, certain trends are visible when AGN of various types are plotted in a "Principal Component 2 (PC2)− Principal Component 1 (PC1)" plane, where PC1 reflects the width of the $\mathrm{H}\beta$ line (which, according to (2), is related to the mass of the central black hole) and PC2 reflects the ratio of the equivalent widths (W) of the Fe 4570 Å and $\mathrm{H}\beta$ lines (which is related to the accretion rate $\dot{m}$). The nonthermal radio emission of the core grows relative to the quasi-thermal optical emission of the disk as $M_{\bullet}$ increases and as $\dot{m}/\dot{m}_{\mathrm{Edd}}$ decreases ($\dot{m}_{\mathrm{Edd}} = 1.38 \times 10^{18} M_{\bullet}/M_{\odot}$ g/s and corresponds for an accretion efficiency $\eta = 0.1$ to the luminosity $L_{\mathrm{Edd}} = 1.38 \times 10^{38} M_{\bullet}/M_{\odot}$ erg/s). Thus, AGN that are stronger in the radio (for a given L_{opt}) should have broader optical lines. This, in turn, leads to smaller values of R_{BL}, in accordance with (5).

This raises the question of whether the sizes of the BLRs in AGN are related to the transverse dimensions of the bases of their radio jets. If these two quantities are related, we can estimate the distances to strong radio sources by estimating R_{BL} from spectroscopic data and deriving the transverse angular size of the radio jet from VLBI data.

What do we find in the literature in connection with this question?
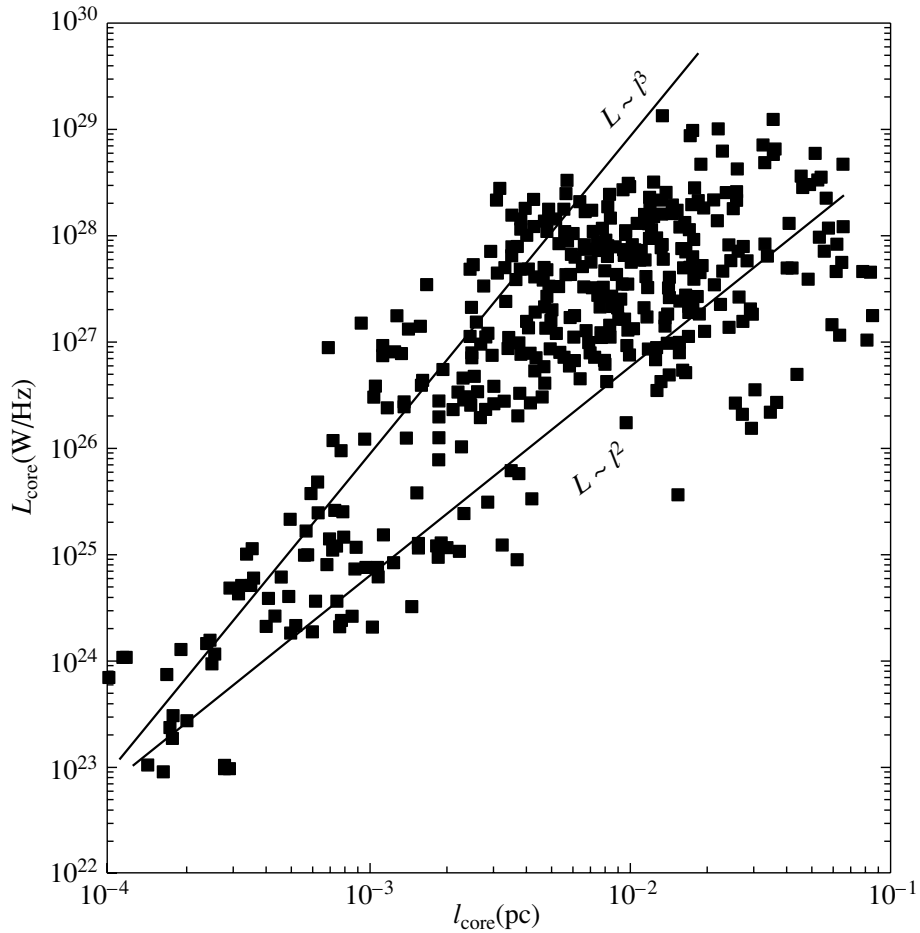
Dependence of the core radio luminosity at 15 GHz (L_{core} in W/Hz) on the size of the radio core (l_{core} in pc), constructed using VLBA data for 160 extragalactic radio sources [33].

(i) It would be of interest to obtain the relation $l_\perp^j - L_r^j$ observationally, to try to confirm that the transverse size of the base of the radio jet and the size of the BLR are approximately equal. Based on this equality, we could draw certain conclusions about mechanisms for the nonthermal radio emission of the jets. Such a relation was constructed for the smallest radio components of radio-loud quasars resolved using VLBI in [31]:

$$l_r = l_r^0 (L_r/L_r^0)^\beta (1+z)^n,$$

where

$$\beta = 0.26, \quad n = 0.3 \quad \text{and}$$
$$l_r^0 = (4.8 \pm 1) \text{ pc} \quad (\text{see also [32]}).$$

Using data for a large number ($N = 160$) of extragalactic radio sources resolved by the VLBA at 15 GHz [33], Yu.Yu. Kovalev (private communication) constructed the relation shown in the figure:

$$L_r \sim l_r^{2-3},$$

which is consistent with the conclusions of [31]. The relation $l_r^c - M_\bullet^{0.64}$ was obtained in [34] using data on 60 extragalactic radio sources and the model of Königl for an inhomogeneous (in B and n_e) jet; using the relation $M_\bullet \sim L_r^{0.5} \Delta V_{1/2}^2$, this can be rewritten in the form $l_r^c \sim L_r^{0.32}$ for objects with equal Hβ line widths, consistent with the data of [32, 33].

Thus, we can obtain the relation $L_r \sim (l_r^c)^3$ from observations of the compact radio sources in the cores of AGN. However, it cannot be ruled out that the dependence $L_r^j - l_\perp^j$ will be found to have another form when data on the transverse sizes of radio jets are obtained. Based on certain theoretical arguments concerning the magnetohydrodynamical nature of the processes accelerating the radiating particles in the jets, we may find an inverse relation of the form $L_r^j - (l_\perp^j)^{-2}$ (see, for example, [35]). This is associated with the fact that narrower jets will form at the inner edge of the magnetized accretion disk, so that the jet magnetic field will be tightly wound by the

Radio sources with jets (sample with $z \leq 0, 2$; $P_c^{5 \text{ GHz}} > 10^{24}$ W/Hz) from [42]

Coordinates	Name	Type	z
$0415 + 379$	3C 111	Sy I	0.0485
$0430 + 052$	3C 120	Sy I	0.0334
$0518 - 458$	Pic A	RG	0.0342
$0838 + 325$	–	RG	0.0694
$0917 + 458$	3C 219	Sy I	0.1744
$0958 + 290$	3C 234	Sy I	0.185
$1003 + 351$	3C 236	Sy I	0.0989
$1028 + 313$	B2 (OL 347)	QSO	0.177
$1226 + 023$	3C 273	QSO	0.158
$1345 + 125$	4C 12.50	Sy I	0.121
$1441 + 522$	3C 303	Sy I	0.141
$1638 + 32$	B2	RG	0.140
$1641 + 173$	3C 346	Sy I	0.160
$1712 + 638$	–	RG	0.0829
$1845 + 797$	3C 390.3	Sy I	0.057
$2043 + 749$	4C 74.26	Q	0.104
$2116 + 818$	S5	Sy I	0.086

rotation of the disk. This creates favorable conditions for reconnection of the magnetic field and enhances the reprocessing of the gravitational energy into the energy of the radiating particles (via the magnetic field).

(ii) Equality of the various types of pressures (due to the relativistic particles, magnetic field, and accreting gas) is supposed in [36]. In this case, it is possible to estimate the synchrotron power of the relativistic electrons (W_j) in the "equilibrium" magnetic field ($B \sim 10^2$ G). This leads to $B^2 = W_j/c(R_\perp^j)^2$ and $R_\perp^j \approx (W_j/cB^2)^{1/2}$. Adopting $W_j \approx 3 \times 10^{46}$ erg/s and $B = 100$ G, we obtain $R_\perp^j \approx 10^{16}$ cm, which is close to the value $R_{\text{BL}} \approx 10$ light days. A similar result can be obtained using the completely different reasoning of [37], which focuses attention on the fact that the stabilization of a narrow directed jet of relativistic particles requires a strong magnetic field, which can be generated in the magnetized accretion disk. This field is supported by currents flowing in the jet itself (see, for example, [35, 38]). The poloidal magnetic field that becomes part of the jet structure is supported by the disk and should be perpendicular to the disk. However, this condition can be satisfied only in the corotation region (r_c), where the Keplerian angular rotational velocity of the disk

plasma (Ω_K) is comparable to the rotational velocity of the magnetic-force lines (Ω_F) due to the rotation of the black hole, which has angular velocity Ω_{BH} and moment of inertia I (the so-called Blandford–Znajek effect [39]). It follows that $\Omega_K \approx \Omega_{\text{BH}}$, where $\Omega_K = v_k/r_c = r_c^{-1}(GM_\bullet/r_c)^{1/2}$, $\Omega_{\text{BH}} = a(c/R_g) \approx a(c^3/GM_\bullet)$, and $a = I/I_{\text{Kerr}} = Ic/GM_\bullet^2$ is the ratio of the moment of inertia of the black hole and the limiting moment of inertia of a Kerr black hole: $I_{\text{Kerr}} \approx M_\bullet cR_g$. These relations can be used to obtain the approximate expression $r_c = (GM_\bullet/c^2)a^{-2/3} = R_g a^{-2/3}$.

The quantity a can be estimated using the approximate empirical formula of [40], which is applicable for astronomical bodies with a wide range of masses and momenta:

$$a = 300(M_\bullet/\Delta m_{\text{accr}})^{-3/2} = 10^{-4},$$

where we have adopted $M_\bullet = 10^8 M_\odot$ and a mass for the giant molecular clouds that are accreting onto the massive black hole of $\Delta m_{\text{accr}} = 10^3 M_\odot$ [37].

Thus, $r_c = 3 \times 10^{13} \times 10^{8/3} \approx 2 \times 10^{16}$ cm $\approx R_{\text{BL}}$.

Finally, a model in which rotational energy is extracted from the disk plasma by "torsional Alfvén waves" in the course of disk accretion onto a rotating black hole is constructed in [41]. The initially weak magnetic field in the dissipatively dominant disk plasma is amplified by the dynamo effect and begins to influence the dynamics of the accretion flows at a distance of r_{tr} from the black hole ($E_{\text{mag}}/E_{\text{therm}} \geq 1$). A poloidal magnetic field is generated, $B_p \sim 1/r^2$, which brings about the reprocessing of rotational energy into the energy of the synchrotron-radiating charged particles (via the reconnection of magnetic-force lines). The value of r_{tr} grows when the accretion rate onto the black hole is low. This leads to an increase in the surface of the disk from which there is an electromagnetic extraction of angular momentum—the ejection of relativistic particles is more energetic and collimated. Hujerrat and Blandford [41] estimate that, for $M_\bullet = 3 \times 10^8 M_\odot$ and $m = 10^{-3}m_{\text{Edd}}$, the size of the transition layer (r_{tr}), which also determines the size of the jet, is $\sim 30R_g \approx 3 \times 10^{15}$ cm. Again, this is not very different from R_{BL}.

3. CONCLUSIONS

Thus, we can see that both theoretical arguments and the (still relatively few) available observational data support the hypothesis that the size of the BLRs in AGN and the transverse sizes of the radio jets in these objects are approximately equal. This hypothesis can be tested with more certainty by selecting a

dozen or so relatively nearby ($z < 0.2$) radio sources (radio galaxies, Seyfert galaxies, or quasars) for which there are good spectroscopic data and conducting high-resolution measurements of the transverse sizes of their radio jets via space-VLBI observations. Using the linear quantity $l_{\perp}^{j} \approx R_{BL}$ derived from spectroscopic data and the angular quantity $\Theta_{\perp}^{j}$ derived from the radio observations, it is then straightforward to estimate the distances for a sample of these objects, such as the one given in tha table. These objects were selected using the data collected in [42] for 661 radio sources with jets. We selected radio sources with $z \leq 0.2$ and $P_c^{5\ GHz} > 10^{24}$ W/Hz. Since all these radio sources have redshifts, it is possible to verify the proposed method directly and obtain an independent estimate of the Hubble constant.

Unfortunately, it is not possible to estimate the distances of even nearby, well-studied radio sources with the currently available data. For example, in the case of Virgo A (3C 274), the stellar velocity dispersion measured at an angular distance from the nucleus of $\sim 0.3''$ by the Hubble Space Telescope reaches ~ 550 km/s [43]. This corresponds to $M_{\bullet} \approx 3 \times 10^9 M_{\odot}$. The transverse angular size of the radio jet at $\nu = 22$ GHz is ~ 2 mas [44], which corresponds to 0.2 pc on a linear scale, since the distance to 3C 274 is ~ 16 Mpc. However, we can see from expression (5) that, in order to have $l_{\perp}^{jet} \approx R_{BL}$, the widths of the optical lines must be $\sim 30\,000$ km/s. That is, we are restricted not by the angular resolution of the radio image but by the angular resolution of the spectrum, since we do not have data on the widths of lines that are radiated closer than $\sim 0.3''$ from the nucleus (which is a factor of 10^2 worse resolution than obtained with the VLBI data). There are indirect data suggesting the presence of very high speeds of gaseous clouds in circumnuclear disks, based on the splitting of broad lines in some AGN (see, for example, [15]). The rotation curve of the central regions of the circumnuclear disk of the nearby Seyfert galaxy NGC 4258 (0.1 pc) obtained using VLBI observations of 1.35-cm water masers in the disk also provides evidence for high rotational speeds [45].

ACKNOWLEDGMENTS

REFERENCES

1. T. Wiklind, Preprint No. 1588 (Space Telesc. Sci. Inst., 2003).
2. S. N. Faber and R. E. Jackson, Astrophys. J. **204**, 668 (1976).
3. R. B. Tully and J. K. Fisher, Astron. Astrophys. **56**, 661 (1977).
4. B. V. Komberg and E. Yu. Shafer, Pis'ma Astron. Zh. **9**, 478 (1983) [Sov. Astron. Lett. **9**, 277 (1983)].
5. É. A. Dibai, Astron. Zh. **61**, 417 (1984) [Sov. Astron. **28**, 245 (1984)].
6. J. M. Pittrad, J. E. Dyson, S. A. Falle, and T. W. Hartquist, astro-ph/0309087.
7. S. Tremaine, K. Gelhardt, R. Bender, et al., Astrophys. J. **574**, 740 (2002).
8. G. A. Shields, K. Gelhardt, S. Salviander, et al., Astrophys. J. **583**, 124 (2003).
9. S. Kaspi, P. S. Smith, H. Netzer, et al., Astrophys. J. **533**, 631 (2000).
10. A. J. Barth, L. C. Ho, and W. L. W. Sargent, astro-ph/0310435.
11. R. J. McLure and M. J. Jeavis, astro-ph/0204473.
12. E. A. Corbett, S. M. Croom, B. J. Boyle, et al., astro-ph/0304541.
13. A. Wandel, astro-ph/0303562.
14. J.-H. Woo and C. M. Urry, astro-ph/0207249.
15. M. Eracleous and J. P. Halpern, astro-ph/0309149.
16. V. M. Lyutyi and A. M. Cherepashchuk, Astrophys. Lett. **13**, 165 (1973).
17. M. J. Fromerth and F. Melia, astro-ph/0207194.
18. W. Kollatschny, Astron. Astrophys. **407**, 461 (2003).
19. M. Karovska, M. Elvis, and M. Mazengo, astro-ph/0303445.
20. M. Elvis and M. Karovska, astro-ph/0211385.
21. V. V. Andreyanov, Experim. Astron. **9**, 103 (1999).
22. W. Junor and J. A. Biretta, Astron. J. **109**, 500 (1995).
23. T. P. Krichbaum, A. W. Witzel, J. A. Zensus, et al., Astron. Astrophys. **329**, 873 (1998).
24. V. Dhawan, K. I. Kellermann, and J. D. Romney, Astrophys. J. **498** 111 (1998).
25. C. G. Mundell, J. M. Wrobel, and A. Pedlar, Astrophys. J. **583**, 192 (2003).
26. L. I. Gurvits, K. I. Kellermann, and S. Frey, Preprint No. 3, JIVE (1998).
27. J. A. S. Lima and J. S. Alcaniz, astro-ph/0109047.
28. T. A. Boronson, Astrophys. J. **565**, 78 (2002).
29. P. Marziani, J. W. Sulentick, T. Zwitter, et al., Astrophys. J. **558**, 553 (2001).
30. Z. Shany, B. J. Wills, E. L. Robinson, and D. Wills, astro-ph/0211641.
31. L. I. Gurvits, Astrophys. J. **425**, 442 (1994).
32. V. I. Zhuravlev and B. V. Komberg, Astron. Zh. **76**, 163 (1999) [Astron. Rep. **43**, 135 (1999)].
33. Yu. Yu. Kovalev and N. S. Kardashev, Preprint No. 21, FIL RAN (Lebedev Phys. Inst., Russ. Acad. Sci., Moscow, 2000).
34. X. Cao and D. R. Jiany, astro-ph/0110541.
35. Ya. N. Istomin and V. I. Pariev, Mon. Not. R. Astron. Soc. **28**, 1 (1996).

36. S. Hein and R. Sunyaev, astro-ph/0305252.

37. Ya. N. Istomin, New Astron. (in press).

38. Ya. N. Istomin and B. V. Komberg, Astron. Zh. **78**, 871 (2001) [Astron. Rep. **45**, 755 (2001)].

39. R. D. Blandford and R. L. Znajek, Mon. Not. R. Astron. Soc. **179**, 433 (1977).

40. P. S. Wesson, Phys. Rev. D **23**, 1730 (1981); Space Sci. Rev. **98**, 329 (2002).

41. A. Hujerrat and R. Blandford, astro-ph/0307317.

42. F. K. Liu and Y. H. Zhang, Astron. Astrophys. **381**, 757 (2002).

43. R. P. Vander Marel, Preprint No. 15, IASSNS-AST (1994).

44. W. Junior and J. A. Biretta, Astron. J. **109**, 500 (1995).

45. M. Miyoshi, J. Moran, J. Herrnstein, *et al.*, Nature **373**, 127 (1995).

Translated by D. Gabuzda

The Effect of the Ionization Rate on the Chemical Composition of Dense Cores of Dark Molecular Clouds

M. S. Kirsanova[1] and D. S. Wiebe[2]

[1]Ural State University, pr. Lenina 51, Yekaterinburg, 620083 Russia
[2]Institute of Astronomy, Russian Academy of Sciences, Pyatnitskaya ul. 48, Moscow, 119017 Russia
Received December 10, 2003; in final form, March 15, 2004

Abstract—The effect of variations in the rate of ionization of neutral chemical species by cosmic rays, ζ, on the abundances of some observed molecules in the dense cores of dark molecular clouds is studied. Changes in molecular abundances accompanying an increased (decreased) ionization rate have a single origin: the acceleration (deceleration) of processes that are affected directly or indirectly by chemical reactions with charged species. In addition to affecting the gas-phase chemistry, an increased cosmic-ray flux leads to the more efficient destruction of dust-grain mantles and also accelerates the freezing of some components onto dust. In particular, in a model with an increased ζ, the destruction of the volatile N_2 molecule by ionized helium leads to the rapid accumulation of nitrogen atoms in dust-phase ammonia, which has a higher desorption energy than N_2. As a result, the gas-phase abundance of NH_3 and N_2H^+ decreases significantly. This mechanism can explain the unusual chemical structures of some dense globules, such as B68, where surprisingly low abundances of nitrogen-bearing molecules are observed together with a central drop in the NH_3 and N_2H^+ column densities. Observations of clouds in HCN and HNC lines can discriminate between the two possible origins of the reduced NH_3 and N_2H^+ abundances: an increased cosmic-ray flux or N_2 freezing due to the higher desorption energy of this molecule. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The dense cores of dark molecular clouds are believed to be regions of formation of low-mass stars in the Galaxy. Studies of these objects enable us to estimate the initial conditions for star formation and parameters of the earliest stages of this process. Many cores host compact infrared sources; however, there are also cores in which infrared sources are absent. These cores are frequently called prestellar, though, as a rule, it is not possible to be sure whether or not a star will be formed in a particular core. One of the main sources of information about the physical characteristics of prestellar cores are spectral lines of a number of chemical species. Owing to developments in various observational techniques, it is now possible not only to determine the average abundances of individual species in a cloud but also to map its distribution, thereby obtaining the spatial distributions of various molecular abundances.

Such observations demonstrate that the chemical structures of most observed prestellar cores are nearly identical [1]: the column densities of CO, HCO^+, CS, and some other molecules decrease from the core periphery toward its center. The depletion of these molecules in the densest part of a cloud is usually explained by proposing that they have frozen onto dust particles [2, 3]. On the other hand, in most objects, the NH_3 and N_2H^+ molecules are concentrated in the central part of the core; this is explained, first, by their being associated with so-called "late-type chemistry" and, second, by the low desorption energy of molecular nitrogen, which precludes the early concentration of nitrogen in dust. In general, the characteristic "onion-skin" structure of prestellar cores can be successfully explained using modern concepts about interstellar chemistry.

This makes the exceptions to this rule all the more interesting: cores of molecular clouds and isolated globules in which the NH_3 and N_2H^+ abundances also decrease toward the densest part of the core. Examples include the well-known globules B68 and L183. It is shown in [4, 5] that the NH_3 and N_2H^+ abundances in B68 are appreciably lower than their abundances in other dense cores; the ring structure typical of CO and CS is also observed in the N_2H^+ distribution. There are indications of a similar structure in the distribution of NH_3 [6]. A drop in the abundances of the N_2D^+ and NH_2D isotopomers is found in the densest part of the core of L183 [7]. The absence of a central concentration of N_2H^+ or NH_3 is also observed in L1536, L1512, and L1521 [8–10]. These observations testify that the region of the depletion of nitrogen-bearing molecules is fairly compact and may not be distinguished in observations with low angular

resolution. Interferometric observations will probably reveal other similar objects.

The central depression in the N_2H^+ distribution could have a number of origins. The study of the globule B68 [4] (see also [11]) attributes this depression to the freezing out of N_2 (parent to both N_2H^+ and ammonia). However, since the desorption energy of molecular nitrogen is low (~ 750 K) [2], accumulating a significant amount of solid N_2 requires a time that can exceed the mean lifetime of a prestellar core [12]. On the other hand, Shematovich et al. [13] noted that a ring-shaped (devoid of a central peak) distribution of the column densities of NH_3 and N_2H^+ arises in models for the collapse of a prestellar core if the cosmic-ray ionization rate ζ appreciably exceeds the "canonical" value $\zeta = 1.3 \times 10^{-17}$ s^{-1}.

The effect of the rate of ionization (by cosmic rays and X rays) on dark-cloud chemistry has been considered many times. For example, Farquhar et al. [14] showed based on a gas-phase chemistry model (neglecting the freezing of molecules onto dust and reactions on dust particle surfaces) that the time necessary to reach chemical equilibrium decreases with increasing ionization rate. Lepp and Dalgarno [15] studied the equilibrium abundances of a number of observed molecules as functions of ζ and showed that, when only gas-phase reactions are taken in account, an increase of the ionization rate by two to three orders of magnitude over the standard value does not lead to considerable changes in the equilibrium abundances of CO and HCO^+. Both these papers dealt only with effects due to a considerable increase in ζ that were able to explain the peculiarities of the molecular structure of active galactic nuclei.

In our Galaxy, as well, the scatter in the estimates of ζ reaches an order of magnitude. In particular, estimates of ζ for L183 range from 4×10^{-18} s^{-1} [14] to 5×10^{-17} s^{-1} [16]. In the latter paper, to study the fractional ionization in a sample of dark cores, the fixed $\zeta = 5 \times 10^{-17}$ s^{-1} is used, which is slightly higher than the standard value; it is noted, however, that the agreement between the results of chemical simulations and observations is also satisfactory for $\zeta = 10^{-16}$ s^{-1}. Using a model for the chemical evolution of a collapsing cloud, El-Nawawy et al. [17] showed that the abundances of most molecules in TMC-1 can be reproduced with the standard value, $\zeta \sim 10^{-17}$ s^{-1}; however, for some species (in particular, NO), agreement with observations can be achieved only with $\zeta \sim 10^{-16}$ s^{-1}. In their analysis of the degree of ionization x_e in the cores of 24 dark clouds, Caselli et al. [18] concluded that the cosmic-ray ionization rate in the Galaxy varies from 10^{-18} s^{-1} to 10^{-16} s^{-1}. Van der Tak and van Dishoeck [19] came to a similar conclusion. Of course, the uncertainty

in ζ is partly due to the use of different approaches in the chemical modeling of the objects studied and uncertainties in the constants of some key reactions, most importantly the dissociative recombination reaction for the H_3^+ ion. However, a comparison of the available information on chemical simulations with data on the propagation of gamma rays tells us that there are also real variations of ζ in the Galaxy, with this parameter exceeding the standard value [19].

We consider here the effect of a changing ionization rate ζ on the chemical structure of prestellar cores within the limits set by the uncertainties in the observational estimates of this parameter. Our main purpose is to determine the reasons behind the tendency for the abundances of NH_3 and N_2H^+ to decrease at the centers of molecular-cloud cores as ζ increases. To find out whether the appearance of a central depression in the distributions of NH_3 and N_2H^+ as ζ increases is accompanied by changes in the abundances of other molecules contained in the cloud, we analyzed the abundances of CO, HCO^+, HCN, and HNC, which are frequently studied in observations. We also investigated the possibility of identifying the most probable origin of the decreased N_2H^+ and NH_3 abundances using these same molecules: the rapid freezing of N_2 (because its desorption energy exceeds the commonly accepted value) or the effect of an enhanced ionization rate on the chemistry of N_2H^+ and NH_3. In addition, we analyzed the effect of ζ on the abundance of water. Our choice of this molecule is due to the basic similarity between the chains of reactions that form water and ammonia; this has enabled us to trace regularities in the variations of abundances due to the chemical similarity of H_2O and NH_3.

In Section 2 of this paper, we describe the model for the chemical evolution of the interstellar medium used in our calculations. Section 3 presents our analysis of the effect of the ionization rate on the main mechanisms for the formation of NH_3, N_2H^+, HCN, CO, HCO^+, and H_2O. These results are discussed in Section 4, and our main conclusions are formulated in Section 5.

2. THE CHEMICAL MODEL AND INITIAL CONDITIONS

We used the model described in detail in [20, 21] to study the chemical evolution of a molecular cloud. This model includes chemical reactions in the gas and dust phases, as well as the accretion of molecules onto dust and their desorption from dust. The rates of gas-phase reactions were taken from the UMIST 95 database [22]. Reactions on surfaces of dust particles were modeled using the technique proposed in [23]. In UMIST 95, the interaction of cosmic rays with

molecules in the gas phase is described by the ionization reactions for H, H_2, He, C, O, and N

$$X + CRP \longrightarrow X^+ + CRP' + e^-,$$
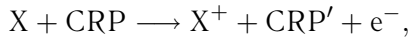
where CRP is a cosmic-ray proton and CRP' is a cosmic-ray proton that has lost part of its energy as a result of a reaction. Fast electrons formed during the cosmic-ray ionization of neutral particles collisionally excite H and H_2. The excited H atoms and H_2 molecules radiate photons, resulting in a weak UV radiation field in the cores of molecular clouds for any value of the extinction [24]. The photoreactions induced by this field are also included in the UMIST 95 calculations and represent a secondary channel for the effect of ζ on the gas-phase chemical compositions of prestellar cores.

Colliding with dust grains, cosmic-ray particles locally heat the molecular mantle and evaporate some of the molecules contained in it. We calculated the rate of this process in accordance with the results of [21, 25]. In addition to desorption by cosmic rays, we also took into account thermal desorption and photodesorption. The desorption energies of chemical species were taken from [25], except for several updated values from [2]. The mass fraction of the dust particles is 0.01 of the gas density. The adopted gas and dust temperatures were both 10 K. The initial abundances listed in Table 1 were taken from [3].

Calculations were done for atomic hydrogen number densities of $n_H = 10^3$, 10^4, 10^5, and 10^6 cm^{-3}. These different values of n_H can be used to compare the chemical compositions of different regions of the core and/or to estimate the role of the increase in the core density during its gravitational contraction. Neglecting variations of n_H considerably simplified the analysis by enabling us to separate the effects of chemical and dynamic factors on changes in the molecular abundances. We chose the calculations with $n_H = 10^6$ cm^{-3} as the primary model.

We present two sets of results as typical examples: in model SI (standard ionization), we use the canonical ionization rate $\zeta = \zeta_0 = 1.3 \times 10^{-17}$ s^{-1}, and in model HI (high ionization), $\zeta = 10^{-16}$ s^{-1}, corresponding to an increase in the flux of cosmic rays incident on the cloud by a factor of $\zeta/\zeta_0 \approx 8$, due, for example, to the presence of nearby supernova remnants. This could be happening, in particular, with the object B68, which is of interest in the context of the problem studied here [26]. We also calculated several models with higher and lower values of ζ, which will be discussed in Section 4.

To estimate the effect of UV radiation, we also calculated several models with a nonzero UV flux and $A_V \geq 10$, which agrees with estimates of the extinction in the cores of dark clouds. The modeling

Table 1. Initial abundances of the components of the chemical model

Component	Abundance
He	0.0975
Na	2.1×10^{-6}
Mg	1.5×10^{-7}
Fe	1.0×10^{-7}
C	5.0×10^{-5}
N	2.2×10^{-5}
O	1.7×10^{-4}
S	8.0×10^{-8}
Si	8.0×10^{-9}

demonstrates that, with such strong absorption, even an increase in the UV flux incident on the cloud by a factor of 10^3 over the mean interstellar background does not result in any appreciable changes in the chemical structure in the cloud interior. The role of the incident UV radiation becomes dominant only at the periphery of the cloud, where A_V is much lower. Since we are interested in the inner regions of a cloud, we will neglect the effect of the incident UV flux below.

3. RESULTS

One of the main characteristics of dark molecular clouds that depends directly on ζ is their degree of ionization, $x(e^-)$. The upper panel in Fig. 1 shows the evolution of the electron abundance in models with standard and enhanced cosmic-ray ionization rates. We can see that the curves essentially repeat each other, but the degree of ionization in model HI is always approximately an order of magnitude higher. We will show below that an enhanced electron abundance is an important but not the sole factor determining the differences between these two models.

The first peak in both curves is due to the evolution of the sodium-ion abundance (Fig. 1c). As sodium atoms stick to dust, the role of the dominant ion is transferred to H_3^+: the evolution of this ion is connected with the increase of $x(e^-)$ at $t > 10^6$ years and with the equilibration of the electron abundance at later times (Fig. 1d). Figure 1b displays the time dependence of the N_2 abundance, which is important in determining the chemistry of nitrogen-bearing molecules.

Figure 2 shows the abundances of CO, HCO$^+$, H_2O, N_2H^+, NH_3, HCN, and HNC as functions of time for $n_H = 10^6$ cm^{-3} in models SI and HI. Though the values of ζ in two these models differ
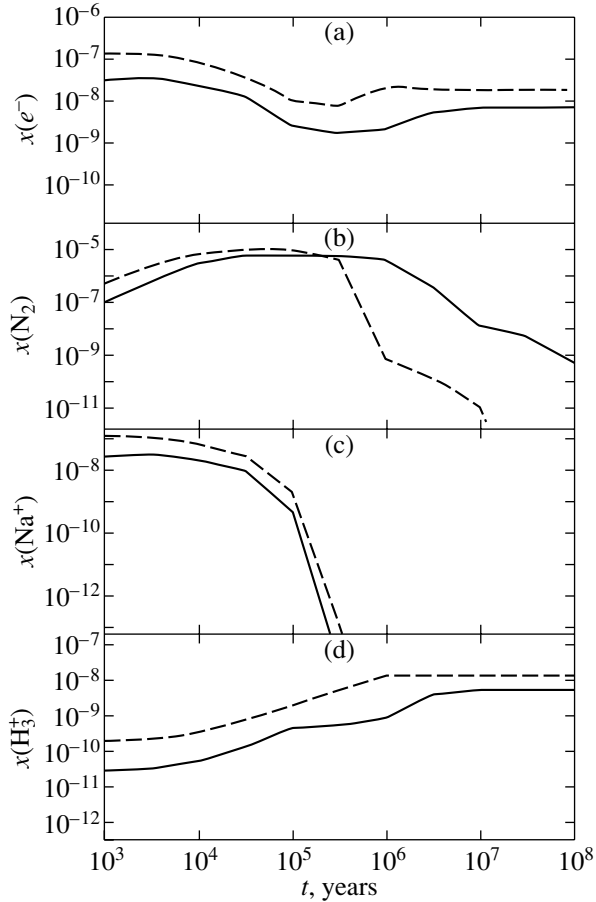
Fig. 1. Evolution of the abundances of electrons, N_2, Na^+, and H_3^+ for the standard (SI; solid curve) and enhanced (HI; dashed curve) cosmic-ray ionization rates.

by a factor of only a few, the differences between the abundances are obvious, especially for the nitrogen-bearing molecules. To find out the origins of these differences, we analyzed in detail the rates of chemical reactions taking place in the medium: which ones change appreciably in the transition from model SI to model HI and, conversely, which remain almost unchanged, and whether there are changes in the priorities of any reactions for each species considered.

The graphs in Fig. 2 showing the evolution of the abundances of molecules we have chosen can be divided into three segments. The first corresponds to an abundance increase on time scales from 10^4 to 5×10^5 years. The molecular abundances reach their maxima earlier in model HI than in model SI; furthermore, the maxima of the abundances are higher in HI (CO is an exception). A sharp drop in the abundances begins in the second segment, which continues to approximately $t = 5 \times 10^6$ years; this is due to the freezing of gas-phase molecules onto dust. The third segment, in which the molecular abundances do not vary appreciably, corresponds to an equilibrium be-

tween the processes of accretion and desorption. With a kinetic temperature of 10 K and $A_V \geq 10$, molecules return to the gas phase almost solely via desorption induced by cosmic rays (this is especially efficient in model HI). We can see that the drop in the molecular abundances in the model with an enhanced ionization rate begins and ends earlier than in the model with the standard ζ. The time dependence of the NH_3 and N_2H^+ abundances (Figs. 2a, 2b) on time scales of the order of 10^8 years includes only the first two segments.

We describe the results for each of the above molecules in detail below, and Table 2 compares the theoretical and observed abundances of some of these molecules in B68 and L183. The theoretical abundances are given for $t = 10^6$ years. The data for B68 are taken from [4] (N_2H^+, CO), [5] (HCO^+, NH_3), and [27] (H_2O). The CO abundance is recalculated from the $C^{18}O$ abundance assuming an isotope ratio of $^{16}O/^{18}O = 500$. We used the data of [28] for L183 (except for water). The upper limit for the water abundance in L183 is taken from [29].

The theoretical abundances are in satisifactory agreement with the observational data for B68, except for the abundances of CO and CS, for which the calculations yield considerably lower values due to the efficient freezing of these components onto dust particles. Note that we did not aim here to reproduce the abundances for any particular object in detail. A detailed comparison with a number of observed cores will be carried out in a forthcoming paper.

3.1. N_2H^+ and NH_3

After reaching their maxima, the abundances of N_2H^+ and NH_3 fall off on time scales of $5 \times 10^5 - 10^6$ years, with this decrease occurring much more rapidly for the enhanced than the standard ionization rate. For instance, with the standard ζ, the N_2H^+ and ammonia abundances at $t = 10^6$ years are 7.4×10^{-10} cm^{-3} and 3.1×10^{-8} cm^{-3}, respectively; with the enhanced ionization rate, the abundances of these molecules drop to 1.1×10^{-12} cm^{-3} and 8.0×10^{-10} cm^{-3}, respectively, by this same time.

In both models, the main channel for the formation of N_2H^+ is the addition of a proton to an N_2 molecule in reactions with ions; the most efficient of these is

$$N_2 + H_3^+ \longrightarrow N_2H^+ + H_2. \qquad (1)$$

The recombination reaction

$$N_2H^+ + e^- \longrightarrow N_2 + H, \qquad (2)$$
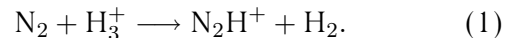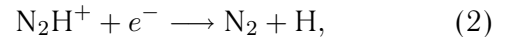
is responsible for the destruction of N_2H^+ molecules. The equilibrium between these two reactions leads to the relatively flat tops of the graphs in Fig. 2a. Since

the abundances of H_3^+ and electrons both increase with the enhanced ionization rate, the N_2H^+ abundance in this short-lived equilibrium is virtually the same in both models.

The presence of N_2 molecules in the gas phase helps to support the N_2H^+ abundance at a level of about 10^{-9}. This equilibrium is disrupted by the destruction of molecular nitrogen by He^+ ions:

$$N_2 + He^+ \longrightarrow N^+ + N + He. \qquad (3)$$
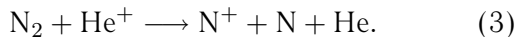
Reactions (1) and (3) compete with each other; however, compared to model SI, the ionized-helium abundance in model HI increases faster than the H_3^+ abundance; therefore, reaction (3) is more efficient when the ionization rate is enhanced. As a result, the decrease in the N_2H^+ abundance begins earlier and proceeds faster in model HI than in model SI.

The formation of ammonia mainly follows the chain

$$N, N^+ \longrightarrow NH^+ \longrightarrow NH_2^+ \longrightarrow NH_3^+ \qquad (4)$$
$$\longrightarrow NH_4^+ \longrightarrow NH_3.$$
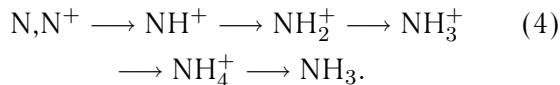
Since there are neutral and ionized nitrogen atoms at the beginning of the chain, it is more efficient in model HI at the expense of reaction (3), and also due to the accelerated ionization of N atoms by cosmic-ray particles.

This initially results in a higher gas-phase abundance of ammonia. However, at later stages, nearly all the nitrogen is bound in ammonia molecules that are frozen in ice mantles (Fig. 3a shows the NH_3 abundance on dust as a function of time), while at the same time, the gas-phase abundance of ammonia sharply decreases. Ammonia ice forms on dust particles via the condensation of gas-phase NH_3, as well as the formation of ammonia from NH and NH_2—products of the recombination of intermediate ions in the ammonia-formation chain—on the dust particles themselves. Virtually none of the ammonia that is formed on dust or accreted from the gas returns, since the NH_3 desorption energy in our model corresponds to a temperature of 3080 K [2].

The evolution of NH_3 and N_2H^+ can be summarized as follows. In the early stages, nearly all the nitrogen atoms are bound in N_2 molecules and only a small fraction are contained in NH_3 and N_2H^+. The desorption energy of N_2 is low, so that these molecules remain predominantly in the gas phase. However, in model HI, N_2 is efficiently destroyed by helium ions and the released nitrogen atoms fall onto the surfaces of dust particles (in atomic form or in the form of NH_x hydrides), where they accumulate as ammonia ice. An increase in the cosmic-ray flux *accelerates* the freezing of nitrogen: the abundance of ammonia on dust, which is approximately equal
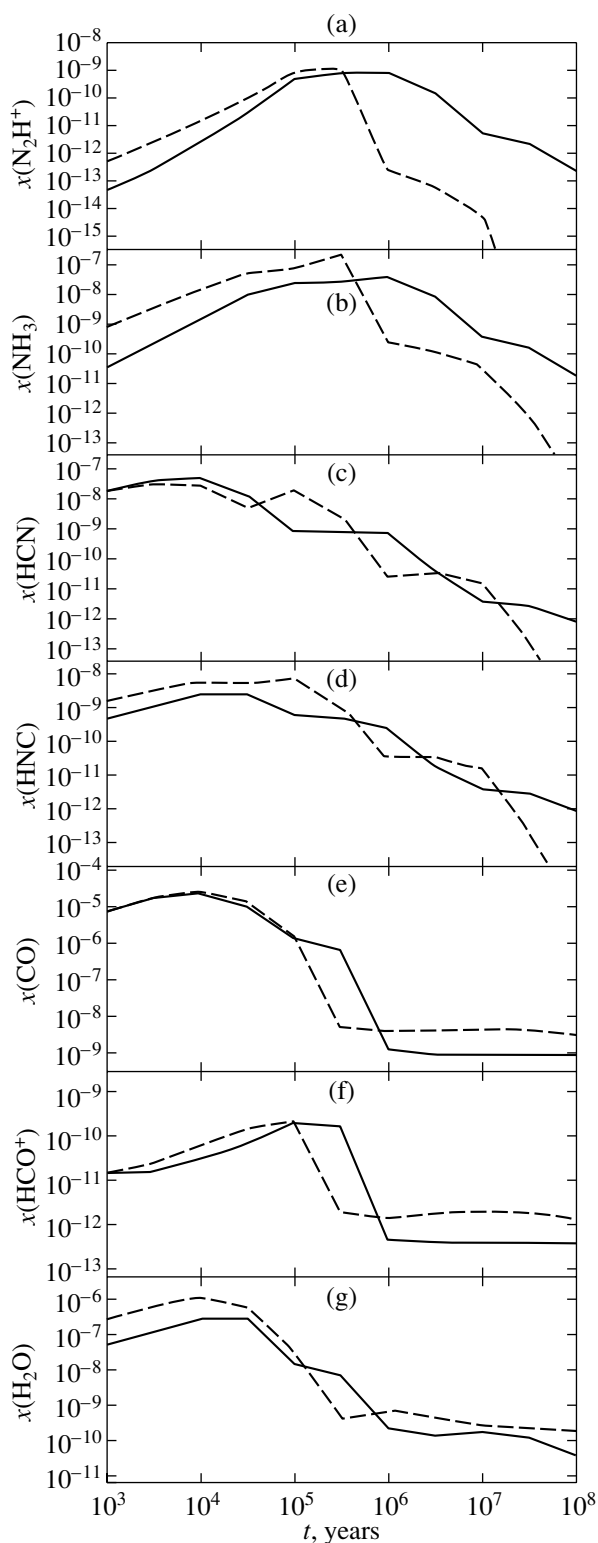

Fig. 2. Evolution of the abundances of (a) N_2H^+, (b) NH_3, (c) HCN, (d) HNC, (e) CO, (f) HCO^+, and (g) H_2O for standard (model SI, solid) and enhanced (model HI, dashed) cosmic-ray ionization rates.

Table 2. Comparison of theoretical and observed abundances

Component	SI	HI	B68	L183
CO	1.3(−09)	4.0(−09)	2.5(−7)	4(−5)
HCO$^+$	4.4(−13)	1.4(−12)	<7(−11)	4(−9)
N$_2$H$^+$	7.8(−10)	2.6(−13)	1.5(−11)	2.5(−10)
NH$_3$	4.2(−08)	2.8(−10)	3.5(−10)	1(−7)
H$_2$O	8.9(−10)	5.3(−10)	<1.5(−8)	<1.5(−7)
CS	4.7(−16)	2.5(−13)	2(−10)	5(−10)
H$_2$CO	4.3(−10)	9.6(−11)	2(−10)	1(−8)
C$_3$H$_2$	3.1(−12)	2.6(−12)	6(−12)	1(−9)

Note: For powers of ten, we use the notation $a(b) = a \times 10^b$.

to 2.0×10^{-5} (only slightly lower than the total nitrogen abundance), is reached considerably earlier in model HI than in model SI. This acceleration could be crucial observationally. With the standard ionization rate in a medium with a density of 10^6 cm^{-3}, freezing of nitrogen-bearing species begins only after several millions of years, which is comparable to the maximum estimated life times of dense cores and globules. Our calculations show that an increase of ζ by a factor of a few results in a shorter freezing time scale for the bulk of the NH$_3$ and N$_2$H$^+$, making it possible to see the effects of nitrogen freezing in much younger objects. This scenario is qualitatively consistent with the calculations of Charnley and Rodgers [30], who studied the evolution of the abundances of nitrogen-bearing molecules in a model with $n_{\rm H} = 10^7$ cm^{-3} and $\zeta = 5 \times 10^{-17}$ s^{-1} taking into account accretion and desorption.

3.2. HCN and HNC

The HCN and HNC abundances in molecular clouds are determined by a vast complex of neutral−neutral and ion−molecular reactions involving nitrogen-bearing molecules and numerous hydrocarbons [31]. This considerably complicates the analysis; nevertheless, general tendencies can be revealed in this case, as well.

The HCN molecule differs from other nitrogen-bearing molecules in that it has a secondary abundance maximum in the model with enhanced ζ; this considerably extends the time interval in which the gas phase contains the maximum possible number of HCN molecules. Whereas x(HCN) already falls below 10^{-9} at $t \approx 10^5$ years in model SI, this takes place at $t \approx 10^6$ years in model HI. The behavior of x(HCN) is determined by the two main channels for
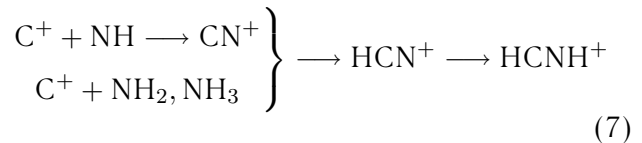
the formation of this molecule: the neutral−neutral reaction

$$N + CH_2 \longrightarrow HCN + H \qquad (5)$$

and the recombination reaction

$$HCNH^+ + e^- \longrightarrow HCN + H. \qquad (6)$$

In model HI, reaction (5) is responsible for the first peak of the HCN abundance, and reaction (6), for the second peak (Fig. 2c). The relatively slow ion−molecular chain for the formation of HCNH$^+$

$$\left. \begin{array}{l} C^+ + NH \longrightarrow CN^+ \\ C^+ + NH_2, NH_3 \end{array} \right\} \longrightarrow HCN^+ \longrightarrow HCNH^+$$

$$(7)$$

is accelerated in model HI due to the higher number densities of the participating ions and the accelerated formation of NH$_x$ species via the ammonia chain. As a result, the HCN abundance falls appreciably at $t \sim 3 \times 10^4$ years and then increases again by $t \sim 10^5$ years. In the time interval from 3×10^4 to $\sim 3 \times 10^5$ years, the HCN abundance in model HI exceeds the HCN abundance in model SI by more than an order of magnitude. The early drop in the HCN abundance in model SI is due to a decreased number density of atomic nitrogen, which is bound up in the more complex molecules N$_2$, NH$_x$, and others. Reaction (5) loses its efficiency due to the low density of N atoms, and reaction (6) is unable to compensate for this loss with the low ion density in model SI.

In the case of the HNC molecule, the situation is simpler. In both models, the main channel for the formation of this molecule over nearly the entire evolutionary time scale is the recombination reaction

$$H_2NC^+ + e^- \longrightarrow HNC + H. \qquad (8)$$

The rate of the recombination reaction for HCNH$^+$, which is analogous to reaction (6) but leads to the formation of HNC, is no greater than 10% of the rate of reaction (8). In turn, the H$_2$NC$^+$ ion is formed in the reactions

$$C^+ + NH_3 \longrightarrow H_2NC^+ + H \qquad (9)$$

$$CH_3^+ + N \longrightarrow H_2NC^+ + H. \qquad (10)$$

Since the ammonia abundance is enhanced until $t \sim 3 \times 10^5$ years in model HI compared to model SI, the HNC abundance in this model is also higher. As for the other nitrogen-bearing molecules, the sharper drop in the HCN and HNC abundances in model HI compared to model SI is due to the accumulation of ammonia in ice mantles and to the corresponding decrease in the content of gas-phase nitrogen atoms.

**Table 3.** Abundances of the main observed molecules in models SI and HI and their ratios at $t = 10^6$ years

Species	$x(\text{SI})$	$x(\text{HI})$	$x(\text{SI})/x(\text{HI})$	Species	$x(\text{SI})$	$x(\text{HI})$	$x(\text{SI})/x(\text{HI})$
N_2H^+	7.8(−10)	2.6(−13)	3000	C_3O	6.9(−15)	2.8(−15)	2.46
HC_9N	9.1(−20)	1.5(−22)	607	H_2CO	2.2(−10)	1.2(−10)	1.83
NH_3	4.2(−08)	2.8(−10)	150	C_3H	2.0(−11)	1.3(−11)	1.54
HC_5N	1.4(−14)	1.0(−16)	140	C_4H	1.7(−12)	2.2(−12)	7.73(−1)
CH_2CO	1.7(−12)	1.3(−14)	131	C_3H_2	7.8(−12)	1.4(−11)	5.57(−1)
C_4H_2	3.8(−13)	8.1(−15)	46.9	C_2	1.7(−12)	3.5(−12)	4.86(−1)
HCN	7.2(−10)	3.1(−11)	23.2	CO	1.3(−09)	4.0(−09)	3.25(−1)
C_5H	3.7(−14)	2.0(−15)	18.5	HCO^+	4.4(−13)	1.4(−12)	3.14(−1)
CH_2CN	9.2(−13)	5.0(−14)	18.4	CN	7.4(−11)	2.9(−10)	2.55(−1)
HC_7N	7.7(−17)	4.2(−18)	18.3	H_2S	1.5(−14)	6.8(−14)	2.21(−1)
CCO	3.7(−14)	2.6(−15)	14.2	CH	6.9(−11)	3.2(−10)	2.16(−1)
C_3N	1.5(−12)	1.2(−13)	12.5	CH_3OH	6.9(−15)	7.4(−14)	9.32(−2)
C_6H	5.4(−15)	5.5(−16)	9.82	OCS	1.7(−19)	2.1(−18)	8.10(−2)
C_6H_2	1.3(−15)	1.5(−16)	8.67	SO	1.4(−15)	2.0(−14)	7.00(−2)
C_3H_4	1.0(−13)	1.3(−14)	7.69	OH	2.2(−10)	4.7(−09)	4.68(−2)
CH_3CN	1.7(−13)	2.3(−14)	7.39	SO_2	8.1(−18)	1.8(−16)	4.50(−2)
$HCNH^+$	3.0(−12)	4.4(−13)	6.82	H_2CS	1.5(−16)	1.6(−14)	9.37(−3)
HNC	3.0(−10)	4.6(−11)	6.52	C_2S	6.2(−17)	9.5(−15)	6.53(−3)
C_2H	3.2(−11)	5.0(−12)	6.40	C_3S	1.5(−18)	2.5(−16)	6.00(−3)
HC_3N	1.4(−11)	3.4(−12)	4.12	CS	7.6(−16)	4.8(−13)	1.58(−3)
NO	9.4(−10)	2.3(−10)	4.09	HCS^+	2.3(−18)	2.5(−15)	9.20(−4)
CH_3CHO	9.8(−15)	2.9(−15)	3.38				

Note: For powers of ten, we use the notation $a(b) = a \times 10^b$.

3.3. CO, HCO^+, and H_2O

The abundances of CO, HCO^+, and H_2O change with increasing ζ less strongly than do the NH_3 or N_2H^+ abundances. The enhanced ionization rate likewise accelerates certain chemical processes involving these molecules; however, this acceleration is manifest differently than in the case of nitrogen-bearing molecules.

The peaks of the CO abundance in models SI and HI coincide (Fig. 2e). This is due to the efficiency of the main channels for the formation of CO at early times: neutral−neutral chains of reactions that do not depend on the ionization rate, in which the final reactions are

$$O + CH \longrightarrow CO + H \qquad (11)$$
$$O + CH_2 \longrightarrow CO + 2H. \qquad (12)$$
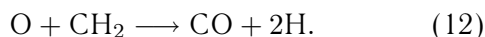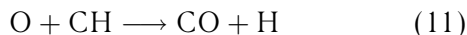
The molecular ion HCO^+ (Fig. 2f) is formed in reactions of CO, H_2O, and H_2CO with C^+, H_3^+, N_2H^+, CH_5^+, and some other ions. Owing to the increase in the number of ions in a cloud as ζ increases, the maximum HCO^+ abundance is reached in a shorter time. However, the drop in the HCO^+ abundance in the second segment also begins earlier in model HI than in model SI.

Water (Fig. 2g) is formed in a chain similar to the ammonia-formation chain:

$$O \longrightarrow OH^+ \longrightarrow H_2O^+ \longrightarrow H_3O^+ \longrightarrow H_2O. \qquad (13)$$
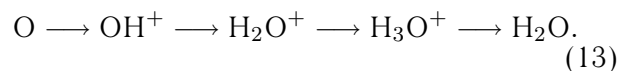
The hydroxyl formed from the recombination of intermediate ions of this chain can be accreted onto dust, where it will participate in the formation of water ice. As for ammonia and HCO^+, the beginning of the decrease in the water abundance is slightly earlier
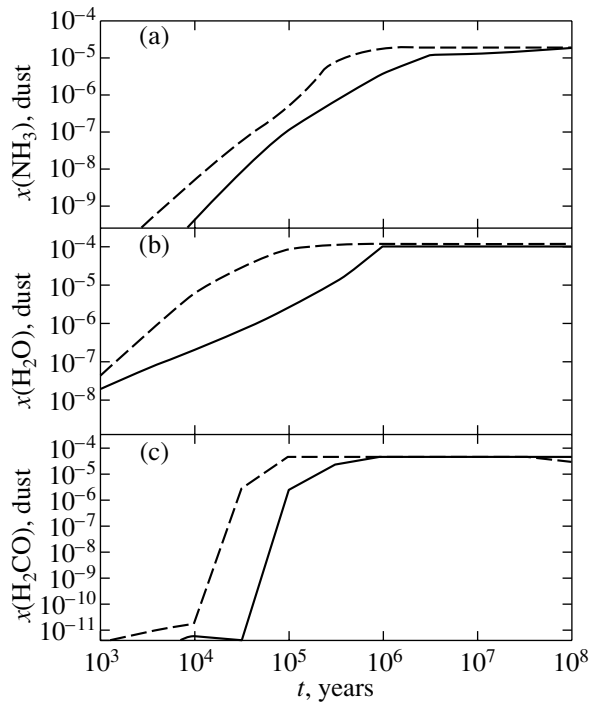
Fig. 3. Evolution of the abundances of NH_3, H_2O, and H_2CO in ice mantles for the standard (model SI; solid) and enhanced (model HI; dashed) cosmic-ray ionization rates.

in model HI than in the model with the standard ionization rate.

The acceleration of the drop in the abundances of CO, HCO^+, and H_2O is due again to the more efficient destruction of N_2 in model HI: the release of N atoms speeds up the reaction

$$N + OH \longrightarrow NO + H. \qquad (14)$$
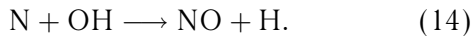
As a result, a considerable fraction of the oxygen atoms is bound in NO molecules. Note that the maximum abundances of nitrogen-bearing molecules in model HI fall right at the beginning of the drop in the CO and HCO^+ abundances. NO molecules settle onto dust and react to form HNO, which, in turn, reacts with oxygen located on the dust,

$$O + HNO \longrightarrow NO + OH, \qquad (15)$$
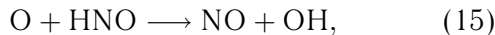
thereby converting O atoms into OH molecules. The desorption energy of hydroxyl is almost a factor of 1.5 higher than that of atomic oxygen (1260 and 800 K, respectively), considerably hindering the return of oxygen to the gas phase. In the end, the accelerated destruction of nitrogen molecules promotes the retention of oxygen atoms on the surfaces of dust particles. An OH molecule settled on a dust particle will quickly join with a hydrogen atom and be transformed into water ($T_D = 4820$ K). Oxygen atoms become depleted in the gas phase; in model HI,

this results in a more rapid drop in the abundances of CO, HCO^+, and H_2O.

CO molecules on dust undergo the sequence of chemical transformations $CO \longrightarrow HCO \longrightarrow H_2CO$. After 10^6 years, most of the carbon atoms are bound into formaldehyde molecules on the surfaces of dust particles. The bulk of the oxygen atoms will ultimately be bound into water molecules, likewise stuck to dust (Figs. 3b, 3c).

The flatter shape of the third segment of the graphs (Figs. 2e, 2f) is due to an approximate equilibrium between the accretion and desorption of carbon-bearing molecules. Cosmic rays evaporate H_2CO and HCO molecules from dust mantles. In the gas phase, the following sequence of reactions involving ions (H^+, H_3^+, He^+) and electrons is realized:

$$H_2CO, HCO \longrightarrow H_3CO^+, H_2CO^+ \qquad (16)$$
$$\longrightarrow HCO^+, CO.$$
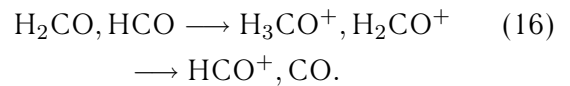
Since the desorption of H_2CO and HCO by cosmic rays is more efficient in model HI, the equilibrium abundances of CO and HCO^+ are reached more rapidly than in the case of the standard ionization rate.

At $t > 3 \times 10^5$ years for the enhanced ionization rate and at $t > 3 \times 10^6$ years for the standard ionization rate, the sharp drop in the gas-phase water abundance also ceases due to the desorption of formaldehyde, with the subsequent reaction

$$H_2CO + He^+ \longrightarrow O + CH_2^+ + He, \qquad (17)$$
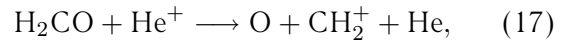
which returns oxygen to the gas phase.

Thus, the only general regularity in the variations of the abundances of NH_3 and H_2O due to their chemical similarity is the accelerated formation of molecules in model HI at times before 10^4 and 10^5 years, respectively, as a consequence of the increased efficiency of reaction chains (4) and (13).

4. DISCUSSION

For most dense cores of molecular clouds, the NH_3 and N_2H^+ abundance peaks usually coincide with the regions of the maximum gas and dust densities indicated in [1]; however, there are several exceptions. The observation of low abundances of nitrogen-bearing molecules in regions of dense gas is usually attributed to the chemical youth of the objects (the maximum abundance of NH_3 and/or N_2H^+ has not yet been reached [10]) or to their advanced age (freezing of NH_3 and N_2 on dust has begun [4]). We have shown here that regions of depleted NH_3 and N_2H^+ can also arise in prestellar cores for another reason: as a result of a local enhancement in the rate of ionization of neutral components by cosmic rays
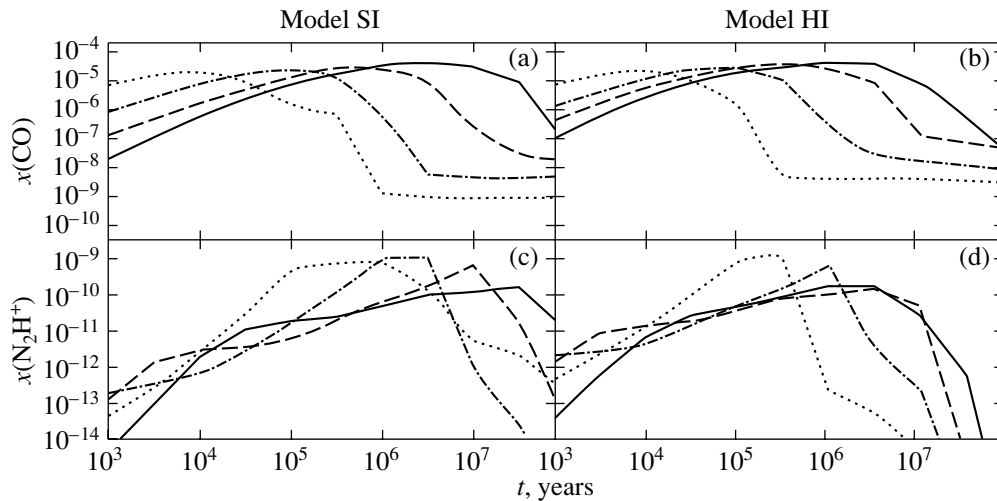
Fig. 4. Evolution of the abundances of N_2H^+ and CO for the standard (model SI) and enhanced (model HI) cosmic-ray ionization rates and for $n_H = 10^3$ cm^{-3} (solid), 10^4 cm^{-3} (dashed), 10^5 cm^{-3} (dot−dashed), and 10^6 cm^{-3} (dotted).

(or X rays). An increase of ζ within the limits of the observational data considerably accelerates the desorption of molecules from dust mantles (manifest as a relative increase in the abundances of CO, HCO^+, and H_2O on long time scales) and also promotes the condensation of nitrogen-bearing molecules onto the surfaces of dust particles. This latter effect comes about because an increased cosmic-ray flux accelerates the conversion of the volatile N_2 molecule to ammonia, whose desorption energy is much higher.

The reactions involving this transformation are two-particle, so that their efficiency increases with increasing density. Figure 4 shows the calculated evolution of the abundances of CO and N_2H^+ for gas number densities of 10^3, 10^4, 10^5, and 10^6 cm^{-3}, which are representative of different parts of a cloud, from its outer envelope to the core. We can see that, as the density increases, the effects of freezing of molecules begin to appear increasingly earlier. In observations of a specific object, this would be manifest as a decrease in the molecular abundances at the cloud center compared to its outer parts, i.e., in the spherically-symmetric case, as a ring-shaped column-density distribution. The radius of the N_2H^+ ring will always be smaller than the radius of the CO ring in the same object.

The difference between the behavior of the abundances of CO and N_2H^+ is associated with the times of the onset of freezing. The central depression in the CO distribution appears at nearly the same time in models SI and HI, whereas the central drop of the N_2H^+ abundance emerges much earlier in model HI than in model SI. For $n_H = 10^6$ cm^{-3} and the standard ionization rate, the abundance of this ion is decreased by an order of magnitude relative to

its peak value at $t \sim 3 \times 10^6$ years. In model HI, a similar decrease in this abundance is reached at $t \sim 3 \times 10^5$ years.

Of the species that are usually observed in regions where low-mass stars are formed, such behavior is characteristic only of NH_3 and N_2H^+. Our calculations show that the difference between the abundances of HCO^+, CO, and H_2O in models SI and HI is much less significant and will be much more difficult to detect in observations, especially in the relatively early stages of a cloud's evolution.

Table 3 lists the computed abundances and their ratios for molecules observed in the molecular clouds TMC-1 and L183 (the list is taken from [28]). A depletion of cyanopolyines and some other carbon-bearing molecules can be noted in model HI at $t = 10^6$ years but is later replaced by a higher abundance compared to model SI, similarly to the case for CO and HCO^+. For species at the bottom of Table 3, the abundances are already determined by the balance between accretion and desorption by $t = 10^6$ years; consequently, they are relatively enhanced in model HI due to the more efficient cosmic-ray desorption. For NH_3 and N_2H^+, the balance between accretion and desorption is achieved much later, at $t > 10^9$ years, due to the large desorption energy of ammonia, and the decrease in their abundances in model HI is therefore quite prolonged. On the whole, from the observational point of view, this means that a slight enhancement of the cosmic-ray flux over the standard flux is manifest in "unusual" abundances of NH_3 and N_2H^+ against the background of "usual" abundances for other species, as, e.g., in ρ Oph A [32].
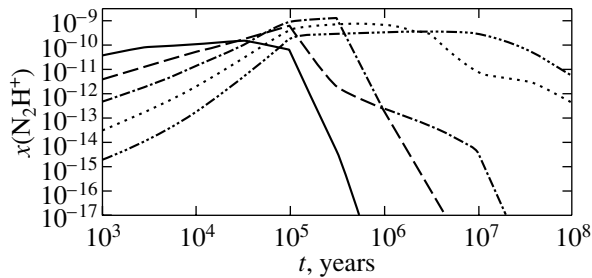
Fig. 5. Evolution of the N_2H^+ abundance for cosmic-ray ionization rates of 10^{-14} s^{-1} (solid), 10^{-15} s^{-1} (dashed), 10^{-16} s^{-1} (dot−dashed), 10^{-17} s^{-1} (dotted), 10^{-18} s^{-1} (dot−dashed with three dots).

Fig. 6. Evolution of the HCN and HNC abundances in models SI (solid), HI (dot−dashed) and SI(N_2) (dashed).

Further increase of ζ strengthens this effect only if the gas temperature remains unchanged. If the model includes heating of the gas by cosmic rays, the situation becomes somewhat more complicated. Figure 5 shows the calculated evolution of the N_2H^+ abundance for ionization rates $10^{-19} < \zeta < 10^{-14}$ s^{-1} and a gas number density of 10^6 cm^{-1}. We computed the gas temperature for this case based on the calculations of [14] [equation (8) of that paper]; according to [14], the gas temperature is about 50 K for $\zeta < 10^{-14}$ s^{-1}. An enhanced cosmic-ray flux results in the more efficient destruction of N_2; however, due to the increased thermal desorption, nitrogen atoms return to the gas phase more actively. As a result, the time scale for the decrease in the N_2H^+ number density with increasing ζ changes in these models much less than in models that assume a constant temperature.

When $\zeta = 10^{-19}$–10^{-18} s^{-3}, the small drop in the N_2H^+ abundance begins no earlier than at 10^7 years. If the actual ionization rates in dark clouds approached these low values, we would detect no objects with depleted N_2H^+ (and NH_3) in their cores. Such low fluxes of cosmic rays are probably unlikely in Galactic molecular clouds but could be characteristic, e.g., of the Magellanic Clouds and other dwarf galaxies.

The above relation between the abundances of nitrogen-bearing molecules and the ionization rate is closely connected to the parameters of the interaction between the gas and dust, in particular, to the desorption energies of molecules involved either directly or indirectly in the synthesis of ammonia and N_2H^+. Bergin *et al.* [4] suggest that the depletion of N_2H^+ in dark-cloud cores could be due to the freezing of molecular nitrogen onto dust. With the low N_2 desorption energy adopted in [4] ($T_D = 750$ K), this process is inefficient and cannot compete with the destruction of N_2 by helium ions.

The uncertain value of the N_2 desorption energy hinders us from unambiguously identifying the most
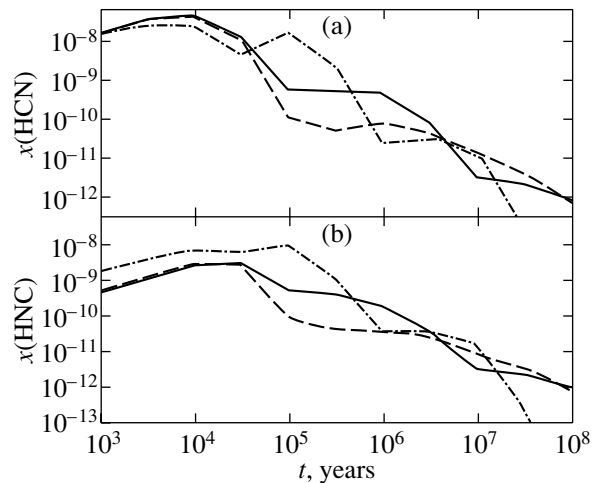
probable origin for the central depressions in the N_2H^+ and NH_3 distributions: whether they are a consequence of N_2 freezing onto dust or of an enhanced flux of cosmic rays. Therefore, we investigated the possibility of estimating the relative roles of these factors based on observations of other molecules. For this purpose, we performed calculations for model SI(N_2), which has the standard ionization rate and a molecular-nitrogen desorption energy equal to that of C_2 and O_2 ($T_D = 1210$ K [25]). The abundances of nitrogen-bearing molecules change appreciably in this case, primarily due to the fact that most of the nitrogen atoms are bound in N_2 molecules on the surfaces of dust particles. As a result, the gas-phase number densities of NH_3 and N_2H^+ remain low compared to the standard SI model during most of the calculation time. The behavior of the abundances of HCN and HNC (Fig. 6) also changes: after reaching their maxima, x(HCN) and x(HNC) decrease much more rapidly than in model SI. These molecules turned out to be most sensitive to the N_2 desorption energy. The difference between the abundances of HCN and HNC in models HI and SI(N_2) exceeds two orders of magnitude at times of the order of 5×10^5 years. Thus, observations of cloud cores in HCN and HNC lines provide the possibility of discriminating between the two possible origins of the depletion of NH_3 and N_2H^+ in the early stages of the core evolution.

The NH_3 abundance in ice mantles in model SI(N_2) also remains low—0.01% of the water-ice abundance at $t = 10^6$ years, as opposed to 4% in model SI with the standard N_2 desorption energy. Since observational estimates yield NH_3(ice)/H_2O(ice) ratios of about 1% [33], the efficient freezing of N_2 seems unlikely. From this point

of view, explanation of the observed distributions of nitrogen-bearing molecules using a somewhat higher value of ζ is preferable.

5. CONCLUSIONS

In this paper, we have studied the effect of the cosmic-ray ionization rate on molecular abundances in a model for the chemical evolution of a dark cloud that includes reactions in the gas phase and on dust particles. All changes in the abundances of molecule with increasing ζ are due to the direct or indirect acceleration of chemical reactions involving charged species. The increase or decrease in the model molecular abundances with enhanced ζ is determined by the time scale for the establishment of equilibrium between accretion and desorption of the parent species of the given molecule. In particular, an increase in the ionization rate over its standard value within the accuracy of the observational data results in a considerable decrease of the NH_3 and N_2H^+ abundances in the core of the dark cloud. We propose this mechanism as the probable origin of the anomalously low abundances of nitrogen-bearing molecules observed in the cores of some molecular clouds.

ACKNOWLEDGMENTS

REFERENCES

1. M. Tafalla, P. C. Myers, P. Caselli, *et al.*, Astrophys. J. **569**, 815 (2002).
2. Y. Aikawa, N. Ohashi, S. Inutsuka, *et al.*, Astrophys. J. **552**, 639 (2001).
3. Z.-Y. Li, V. I. Shematovich, D. S. Wiebe, and B. M. Shustov, Astrophys. J. **569**, 792 (2002).
4. E. A. Bergin, J. Alves, T. Huard, and C. J. Lada, Astrophys. J. **570**, L101 (2002).
5. J. Di Francesco, M. R. Hogerheijde, W. J. Welch, and E. A. Bergin, Astron. J. **124**, 2749 (2002).
6. S.-P. Lai, T. Velusamy, W. D. Langer, and T. B. H. Kuiper, Astron. J. **126**, 311 (2003).
7. L. Pagani, A. J. Apponi, A. Bacmann, *et al.*, *SF2A-2003: Semaine de l'Astrophysique Francaise, Bordeaux, France, 2003*, Ed. by F. Combes, D. Barret, and T. Contini (EdP-Sciences, Conf. Ser., 2003), p. 92.
8. P. Caselli, P. J. Benson, P. C. Myers, and M. Tafalla, Astrophys. J. **572**, 238 (2002).
9. J.-E. Lee, N. I. Evans, Y. L. Shirley, and K. Tatematsu, Astrophys. J. **583**, 789 (2003).
10. T. Hirota, T. Ito, and S. Yamamoto, Astrophys. J. **565**, 359 (2002).
11. E. A. Bergin, in *Proceedings of the 4th Zermatt-Cologne Meeting* (in press); astro-ph/0311216.
12. L. Hartmann, J. Ballesteros-Paredes, and E. A. Bergin, Astrophys. J. **562**, 852 (2001).
13. V. I. Shematovich, D. S. Wiebe, B. M. Shustov, and Z.-Y. Li, Astrophys. J. **588**, 894 (2003).
14. P. R. A. Farquhar, T. J. Millar, and E. Herbst, Mon. Not. R. Astron. Soc. **269**, 641 (1994).
15. S. Lepp and A. Dalgarno, Astron. Astrophys. **306**, 21 (1996).
16. J. P. Williams, E. A. Bergin, P. Caselli, *et al.*, Astrophys. J. **503**, 689 (1998).
17. M. S. El-Nawawy, D. A. Howe, and T. J. Millar, Mon. Not. R. Astron. Soc. **292**, 481 (1997).
18. P. Caselli, C. M. Walmsley, R. Terzieva, and E. Herbst, Astrophys. J. **499**, 234 (1998).
19. F. F. S. van der Tak and E. F. van Dishoeck, Astron. Astrophys. **358**, L79 (2000).
20. D. Wiebe, D. Semenov, and Th. Henning, Astron. Astrophys. **399**, 197 (2003).
21. D. Semenov, D. Wiebe, and Th. Henning, Astron. Astrophys. **417**, 93 (2004).
22. T. J. Millar, P. R. A. Farquhar, and K. Willacy, Astron. Astrophys., Suppl. Ser. **121**, 139 (1997).
23. T. I. Hasegawa, E. Herbst, and C. M. Leung, Astrophys. J., Suppl. Ser. **82**, 167 (1992).
24. S. S. Prasad and S. P. Tarafdar, Astrophys. J. **267**, 603 (1983).
25. T. I. Hasegawa and E. Herbst, Mon. Not. R. Astron. Soc. **261**, 83 (1993).
26. J. F. Alves, C. J. Lada, and E. A. Lada, Nature **409**, 159 (2001).
27. E. A. Bergin and R. L. Snell, Astrophys. J. **581**, L105 (2002).
28. M. Ohishi, W. M. Irvine and N. Kaifu, *Astrochemistry of Cosmic Phenomena*, Ed. by P. D. Singh (Kluwer, Dordrecht, 1992), p. 171.
29. R. L. Snell, J. E. Howe, M. L. N. Ashby, *et al.*, Astrophys. J. **539**, L101 (2000).
30. S. B. Charnley and S. D. Rodgers, Astrophys. J. **569**, 133 (2002).
31. B. E. Turner, L. Pirogov, and Y. C. Minh, Astrophys. J. **483**, 235 (1997).
32. R. Liseau, B. Larsson, A. Brandeker, *et al.*, Astron. Astrophys. **402**, L73 (2003).
33. E. L. Gibb, D. C. B. Whittet, and J. E. Chiar, Astrophys. J. **558**, 702 (2001).

Translated by G. Rudnitskiĭ

Relative Variations of the Physical Parameters of Variable Extragalactic Radio Sources

S. A. Tyul'bashev and P. A. Chernikov

Pushchino Radio Astronomy Observatory, Astro Space Center, Lebedev Physical Institute, Moscow, Russia
Received September 10, 2003; in final form, January 9, 2004

Abstract—A method for studying the physical conditions in compact components of extragalactic radio sources displaying variability on time scales of hundreds of days is proposed. The method can be used to estimate the relative variations of the magnetic-field strength and number density of relativistic electrons in superluminal jets from the cores of quasars and radio galaxies. Results are presented for the jets of the quasars 3C 120, 3C 273, 3C 279, and 3C 345. The energies of the magnetic field and relativistic particles in these objects are not in equipartition. As a rule, the magnetic-field strength decreases appreciably during the evolution of an expanding jet, while the number of relativistic electrons grows. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The radio emission of extragalactic sources is variable on time scales from fractions of a second to several years. This variability can have both intrinsic (occurring in the source itself) and extrinsic (not associated with the source) origins. Many intensity variations on time scales from fractions of a second to several minutes were rapidly recognized as scintillation of the source radiation on fluctuations in the electron density of the solar plasma or the Earth's ionosphere. However, various problems arose in connection with attempts to explain the variability on longer time scales, some of which remain unresolved to this day.

There is some debate about the origin of variability on time scales from several hours to days. There are arguments in support of both intrinsic and extrinsic origins. Currently, the hypothesis that such variability is due to scintillation in the interstellar plasma appears to be the most convincing. The origin of variability on time scales from several months to years is likewise not fully understood.

Variability with a characteristic time scale of several years was first detected in the source CTA 102 in 1965 by Sholomitskiĭ, whose observations were at 900 MHz [1]. Several months later, Dent [2] reported variability of 3C 273 at 8 GHz.

In 1986, Witzel *et al.* [3] showed that variability of extragalactic radio sources on time scales of days or even fractions of a day is a widespread phenomenon. The most rapid variability (on time scales of several hours) that is not associated with ionospheric or interplanetary scintillations has been observed in J1819+3845 [4].

A number of reviews have been devoted to long-timescale and short-timescale variability, of which [5, 6] (long-timescale) and [7] (short-timescale) are most noteworthy. It is currently believed that long-timescale variability at frequencies above 1 GHz is associated with processes occurring within the sources. In this case, the linear size of the variable region in the source should be related to the characteristic time scale for the variability. For example, if the emission is variable on a time scale of one year, the linear dimensions of the variable region should not exceed about one light year.

Two regions can be distinguished in a variable source, both located near the galactic nucleus. The first, contained in the core component in VLBI images, has an observed flux density that is essentially constant in time. The second region, which is also physically close to the core, is the location of flares associated with the ejection of material from the core. It is this second region that is the site of the radio variability. If we know the flux density from the source and we have an upper limit for its angular size, we can obtain a lower limit for the brightness temperature of the flare region. It turns out that, in sources with variability time scales shorter than a year, this temperature is close to the inverse-Compton limit; the brightness temperature can exceed the inverse-Compton limit by several orders of magnitude in sources with variability time scales of hours. These high brightness temperatures are explained by supposing that the variable region is located in a jet that moves toward the observer with relativistic speed, so that the emission is Doppler boosted.

The dimensions of the flare region are fairly small. Observations that have resolved the features associated with the long-timescale variability have recently been obtained for several sources (for example, [8]). The sizes of these features are of the order of a light year.

In 1966, van der Laan [9] proposed a model for a variable extragalactic source in which a compact component (shock, flare) is formed as a result of some series of processes (similar to a supernova explosion), then adiabatically expands and gradually decays. This is precisely the component that we have called the flare component, which is visible in the variable source's light curves at various radio frequencies. The van der Laan model describes the shape of the light curve and variations in the amplitude, flux density, and epoch of the maximum brightness as a function of frequency. This theory was further developed in [10−14].

Thus, we have a theory that can explain the evolution of a flare, including its effect on the light curve and the time variations of the source spectrum. At the same time, there are no unambiguous answers to a number of questions related to the physics of active galactic nuclei. In particular, it is not clear how energy is transported from the black hole to the radio-core and flare regions, what physical conditions exist in these regions, or how they vary with time.

We attempt here to estimate variations in the physical conditions occurring in the core region during a flare based on an analysis of the corresponding light curves.

2. METHOD AND RESULTS

As a rule, the physical conditions in radio sources are usually estimated assuming equipartition between the energies of the magnetic field and relativistic particles. However, if a turnover due to synchrotron self-absorption is observed in the radio spectrum, it is possible to estimate the physical conditions without assuming energy equipartition. Formulas showing that it is not essential to assume energy equipartition when estimating the physical conditions in a source are presented in [15−17]. The method described in [15] was applied to a study of the physical conditions (parameters) in giant radio galaxies in [18]. We will now present the main assumptions underlying this method.

According to Slysh [19], the magnetic-field strength in the source can be estimated if the synchrotron self-absorption turnover is observed in the spectrum:

$$H_\perp = 2 \times 10^{-11} K(\gamma) \nu^5 \theta^4 S^{-2} (1+z)^{-1}, \quad (1)$$

where ν is the turnover frequency (where the optical depth becomes $\tau > 1$) in MHz, S is the flux density in Jy at this frequency, θ is the angular size of the source in arc seconds, z is the source's redshift, and $K(\gamma)$ (where $\gamma = 2\alpha + 1, S \sim \nu^{-\alpha}$) is a coefficient tabulated by Pacholczyk [20].

Thus, the magnetic-field strength can be estimated from the low-frequency radio spectrum. If the frequency ν is taken near the turnover, not far from the peak of the spectrum, the resulting value of $H_\perp$ will characterize the average magnetic field of the source perpendicular to the line of sight.

It is also possible to estimate the number density of relativistic electrons using the resulting estimate of the magnetic-field strength and data derived from the high-frequency part of the spectrum [15]:

$$n_e = \int_{E_1}^{E_2} N(E)\, dE = \frac{N_0}{1-\gamma} \left(E_2^{1-\gamma} - E_1^{1-\gamma} \right), \quad (2)$$

$$E_{1,2} = \left(\frac{\nu_{1,2}}{1.9 \times 10^6 H_\perp} \right)^{1/2}, \quad (3)$$

where $N(E)dE = N_0 E^{-\gamma} dE$ is the energy distribution of the electrons and $H_\perp$ is the estimated magnetic-field strength obtained from (1). The integration limits E_1 and E_2 are taken by assuming that all the energy is emitted by the relativistic electrons at frequencies between $\nu_1 = 10$ MHz and $\nu_2 = 100$ GHz. This choice of frequency limits covers the entire radio range. The value of N_0 can be obtained from the equation [15]

$$N_0 \sim H_\perp^{-\frac{\gamma+1}{2}} \nu^{\frac{\gamma-1}{2}} S \theta^{-3}. \quad (4)$$

The energy density of the magnetic field ω_H and the energy density of the relativistic particles ω_e can also be estimated:

$$\omega_H = \frac{H_\perp^2}{8\pi}, \quad (5)$$

$$\omega_e = \int_{E_1}^{E_2} E N(E)\, dE = \frac{N_0}{2-\gamma} \left(E_2^{2-\gamma} - E_1^{2-\gamma} \right). \quad (6)$$

Thus, if the spectrum has a turnover associated with synchrotron self-absorption, and the angular size and redshift of the source are known, it is possible to estimate the magnetic-field strength $H_\perp$, density of relativistic electrons n_e, total number of relativistic electrons $N_e = n_e V$ (where $V \sim \theta^3$ is the volume of the source), and the energies of the relativistic plasma ($E_e = \omega_e V$) and of the magnetic field ($E_H = \omega_H V$).

The accuracy of estimates obtained using this method depends primarily on the accuracy of the estimated angular size of the source. For example, if the size of the source derived from VLBI observations is known with an accuracy of 20−30%, the true value of the magnetic field should differ from the estimated value by no more than a factor of 1.5−2 ($H_\perp \sim \theta^4$).

The use of this method in studies of sources with long-timescale variability is hindered by the lack of good estimates of the angular dimensions of the sources in most cases due to the lack of corresponding VLBI data. In addition, it is necessary to determine a means to extract the spectrum of the studied component from the light curve. However, the main reason the method cannot be applied to studies of variable sources is the presence of superluminal motions in the radio cores, as is indicated by the inference of brightness temperatures exceeding the inverse-Compton limit.

Nevertheless, there exists the possibility of adapting this approach to make it applicable for studies of the physical parameters of variable sources. For example, a modification of the formula of Slysh for the case of a source with superluminal motions was obtained by Marscher [16] in 1983:

$$H_\perp = 2 \times 10^{-11} K\left(\gamma\right) \nu^5 \theta^4 S^{-2} \left(\frac{\delta}{1+z}\right), \quad (7)$$

where $\delta = \frac{\sqrt{1-\beta^2}}{1-\beta\cos\phi}$, $\beta = \frac{v}{c}$ is the ratio of the speed of a clump of relativistic particles (here and below, a "blob") v to the speed of light, and ϕ is the angle between the motion and the direction toward the observer. This formula cannot be used directly to estimate the physical parameters in variable sources because we have no estimates of the speeds of the superluminal components in the vast majority of cases. However, if we know how the spectrum of the variable source (blob) evolves with time, we can trace the relative variations of the physical parameters in this component.

The spectrum of the varying blob can be obtained as follows. A region of enhanced brightness is identified in the light curve at some frequency. The observed flux density at each time is made up of the flux density of the region occupied by the variable blob and the constant flux density of the remainder of the source. The radiation at the minimum brightness is ascribed to the constant component, and all the remaining radiation, to the blob. After subtracting the constant component from the light curves at several frequencies, the resulting data can be used to construct the spectrum of the variable blob at various moments in time.

Such analyses have been carried out in a number of studies. In particular, a statistical analysis of the light curves of 27 variable radio sources was conducted in [21]. This work showed that a so-called "average" flare has a spectrum that resembles that of a uniform synchrotron source, with spectral indices of about 0.2 and −2.5 in the optically thin and optically thick parts of the spectrum (the latter indicates the presence of synchrotron self-absorption in these objects).

The light curves can also be used to derive an upper limit for the linear size of the blob, $l = ct$, where t is the time from the birth to the decay of the flare at the given frequency.

Radio observations show that the evolution of a flare can be traced beginning at a frequency of about 100 GHz down to frequencies of several hundreds of MHz. The maximum of the spectrum is shifted from higher to lower frequencies with time. Let us suppose that the spectrum of the blob has a turnover due to synchrotron self-absorption and that the relativistic motion of the blob is constant as the spectrum's peak shifts towards lower frequencies. These are reasonable assumptions if the frequencies considered are not too far from each other.

Let $H_{\perp 1}$ be the magnetic field in the variable blob when the maximum of its spectrum is at frequency ν_1 (which coincides with the maximum flux density in the light curve at that frequency). Let $H_{\perp 2}$ be the magnetic field in the blob when the spectral maximum has shifted to the lower frequency ν_2. We will now write expressions for $H_{\perp 1}$ and $H_{\perp 2}$ using (7) and divide one by the other. This trivial procedure eliminates the unknown relativistic speed of the blob:

$$\frac{H_{\perp 1}}{H_{\perp 2}} = \left(\frac{\nu_1}{\nu_2}\right)^5 \left(\frac{\theta_1}{\theta_2}\right)^4 \left(\frac{S_1}{S_2}\right)^{-2}. \quad (8)$$

Similarly rewriting the corresponding equations for the number density of the relativistic electrons n_e, the total number of relativistic electrons N_e, and the energies of the magnetic field E_H and the relativistic particles E_e, we obtain

$$\frac{n_{e1}}{n_{e2}} = \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\gamma-1}{2}} \left(\frac{\theta_1}{\theta_2}\right)^{-3} \frac{S_1}{S_2} \left(\frac{H_{\perp 1}}{H_{\perp 2}}\right)^{-1}, \quad (9)$$

$$\frac{N_{e1}}{N_{e2}} = \frac{n_{e1}}{n_{e2}} \left(\frac{\theta_1}{\theta_2}\right)^3 = \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\gamma-1}{2}} \frac{S_1}{S_2} \left(\frac{H_{\perp 1}}{H_{\perp 2}}\right)^{-1}, \quad (10)$$

$$\frac{E_{H1}}{E_{H2}} = \frac{\omega_{H1} V_1}{\omega_{H2} V_2} = \left(\frac{H_{\perp 1}}{H_{\perp 2}}\right)^2 \left(\frac{\theta_1}{\theta_2}\right)^3, \quad (11)$$

$$\frac{E_{e1}}{E_{e2}} = \frac{\omega_{e1}V_1}{\omega_{e2}V_2} = \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\gamma-1}{2}} \frac{S_1}{S_2} \left(\frac{H_{\perp 1}}{H_{\perp 2}}\right)^{-3/2}. \quad (12)$$

These formulas can be used to analyze data for variable sources for which good-quality light curves are available at several frequencies covering a broad range. It is necessary to have two or three points in the optically thick regime (to determine the spectral index at these frequencies), one point near the maximum (to determine the flux density at the maximum), and two or three frequencies in the optically thin regime (to determine the spectral index at these higher frequencies).

In addition, statistical studies of variable sources enable us to apply our method to estimate the relative variations of the physical parameters of the blob.

An analysis for a nearly complete sample of 100 active galactic nuclei with flat spectra and flux densities exceeding 2 Jy at 22 GHz was carried out in [22]. Independent of the redshift and type of source, the variability timescale (which is related to the linear dimensions of the blob) and the flux density at the maximum lie within a fairly narrow range. The typical flare durations were 0.63 light years at 22 GHz and 0.4 light years at 37 GHz. The mean flux density at the flare maximum in the plateau stage was $\Delta S = 1.68$ Jy at 22 GHz and $\Delta S = 1.64$ Jy at 37 GHz. These data are sufficient to apply our method to an "average flare":

$$\frac{H_{\perp 1}}{H_{\perp 2}} = \left(\frac{37}{22}\right)^5 \left(\frac{0.40}{0.63}\right)^4 \left(\frac{1.64}{1.68}\right)^{-2} = 2.3 \pm 0.3, \quad (13)$$

$$\frac{n_{e1}}{n_{e2}} = \left(\frac{37}{22}\right)^{0.2} \left(\frac{0.40}{0.63}\right)^{-3} \frac{1.64}{1.68} (2.3)^{-1} = 1.8 \pm 0.3, \quad (14)$$

$$\frac{N_{e1}}{N_{e2}} = \left(\frac{37}{22}\right)^{0.2} \frac{1.64}{1.68} (2.3)^{-1} = 0.5 \pm 0.2, \quad (15)$$

$$\frac{E_{H1}}{E_{H2}} = (2.3)^2 \left(\frac{0.40}{0.63}\right)^3 = 1.4 \pm 0.3, \quad (16)$$

$$\frac{E_{e1}}{E_{e2}} = \left(\frac{37}{22}\right)^{0.2} \frac{1.64}{1.68} (2.3)^{-3/2} = 0.3 \pm 0.1. \quad (17)$$

The uncertainties in these estimates are calculated assuming that the uncertainties in the flux density at the maximum of the blob's spectrum and in the angular size of the blob are 20%. We expect the properties of individual flares to have fairly large uncertainties, while the uncertainties should become appreciably smaller in statistical analyses such as that carried out in [22].
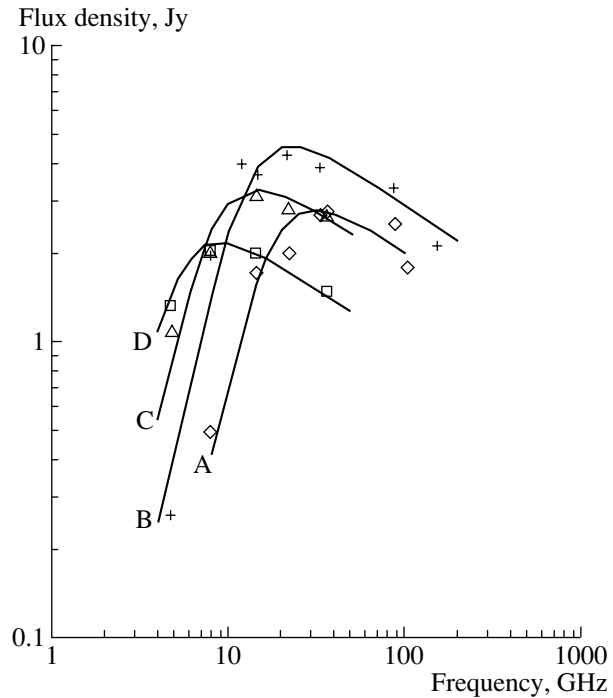
Evolution of the spectrum of a blob in 3C 279 ejected in 1980. The dates and parameters of the spectrum are presented in Table 3.

For example, according to the conservation of magnetic flux, $\frac{H_1}{H_2} = \left(\frac{r_1}{r_2}\right)^{-2}$, where r is the radius of the source. We then find for a blob that expands from 0.4 light years to 0.63 light years in size $H_{\perp 1}/H_{\perp 2} = (0.40/0.63)^2 = 2.48$. This value differs from the ratio of the perpendicular components of the magnetic field obtained above by only 8%, providing indirect evidence for the correctness of this method.

Let us note two especially interesting conclusions. First, the relations we have obtained show that the magnetic-field energy decreases and the relativistic-particle energy grows during the evolution of the blob. This indicates that we should not expect equipartition of the energies of the magnetic field and particles. This is not a new result. Similar conclusions were drawn, for example, in [18] in connection with studies of giant radio galaxies. It was concluded in the study of the physical parameters of compact steep-spectrum sources [23] that most of these compact sources are not in equipartition. However, the linear dimensions of the objects studied earlier ranged from several parsecs to kiloparsecs, and the turnovers in their spectra were observed at frequencies below 1 GHz. Here, we have been able to show an absence of equipartition in compact features fractions of a light year in size with spectral turnovers at frequencies above 10 GHz.

Table 1. Jetlike source

Parameter	$k_H = 10^2$			$k_H = 10^4$			Uniform field
Law for decrease in H	$1/r$	$1/r^2$	$1/r^3$	$1/r$	$1/r^2$	$1/r^3$	const
H_{01}/H_{02}	2.30	2.30	2.30	2.30	2.30	2.30	2.3 ± 0.3
n_{e1}/n_{e2}	1.52	1.54	1.50	1.48	1.51	1.57	1.8 ± 0.3
N_{e1}/N_{e2}	0.37	0.36	0.41	0.39	0.40	0.42	0.5 ± 0.2
E_{H1}/E_{H2}	1.28	1.30	1.29	1.28	1.29	1.29	1.4 ± 0.3
E_{e1}/E_{e2}	0.30	0.29	0.29	0.29	0.30	0.28	0.3 ± 0.1

Table 2. Spherically symmetrical source

Parameter	$k_H = 10^2$			$k_H = 10^4$			Uniform field
Law for decrease in H	$1/r$	$1/r^2$	$1/r^3$	$1/r$	$1/r^2$	$1/r^3$	const
H_{01}/H_{02}	2.30	2.30	2.30	2.30	2.30	2.30	2.3 ± 0.3
n_{e1}/n_{e2}	1.82	1.85	1.85	1.84	1.86	1.83	1.8 ± 0.3
N_{e1}/N_{e2}	0.41	0.37	0.39	0.39	0.39	0.37	0.5 ± 0.2
E_{H1}/E_{H2}	1.29	1.30	1.30	1.29	1.31	1.30	1.4 ± 0.3
E_{e1}/E_{e2}	0.33	0.28	0.30	0.30	0.29	0.28	0.3 ± 0.1

Second, the total number of relativistic particles doubles during the evolution of an "average blob" (while the total density of relativistic particles decreases by a factor of 1.8). This is probably associated with the interaction between the relativistic and thermal plasma. High-energy relativistic particles can collide with thermal plasma at the boundaries of shocks, giving rise to additional relativistic electrons.

3. MODELING OF SOURCES WITH A UNIFORM MAGNETIC FIELD

Strictly speaking, the above conclusions are valid only for spherically symmetrical sources with uniform magnetic fields. However, there are a number of indirect arguments that suggest a model with a spherically symmetrical source with a uniform magnetic field can be applied to the case of a shock. One recent work concerned with modeling synchrotron spectra of nonuniform sources [24] shows that the nonuniform distribution of the relativistic particles has virtually no effect on the shape of the observed spectrum. This means that, to a first approximation, the nonuniformity of the particles can be neglected. This work also shows that the spectral index in the

optically thick region reaches -2.5 for the case of an elongated source (jet) if the field is uniform throughout the source. According to [21], the spectral index of a shock is likewise -2.5 in the optically thick region.

In addition, an approach to numerically modeling the synchrotron spectra of nonuniform sources is described in detail in [24]. We decided to model the spectra of nonuniform sources in order to understand how much nonuniformity of the magnetic field limits the applicability of our method.

We selected power-law nonuniformities for the analysis. If plasma flows from the center of a radio source, as in the solar wind, the magnetic field becomes extended in the radial direction and the magnetic field falls off as $1/r^2$. If the magnetic field is dipolar, then $H \propto \dfrac{1}{r^3}$. There is also evidence that the magnetic fields in a number of sources fall off as $1/r$ [25]. Thus, based on physical considerations, we expect that the distribution of the magnetic field obeys a power law. Of course, the nonuniformity of the magnetic-field distributions in real sources can have some other form, and describing it using a power law is only a model approximation; however, such an approximation is certainly closer to the real situation than using a model with a uniform source.

Table 3. Parameters of flare spectra

Source	Name	Date	ν_{max}, GHz	S_{max}, Jy	Size, light years
3C 120	A	March 1985	15	2.5	3.6
	B	April 1985	8	2.7	3.7
	C	February 1986	5	2.0	4.0
3C 273	A	January 1984	15	21	8.5
	B	May 1984	8	12	8.7
	C	January 1986	5	7	9.0
3C 279	A	March 15, 1981	37	2.8	3.4
	B	January 20, 1982	22	4.3	3.7
	C	September 10, 1982	14	3.1	3.9
	D	April 20, 1983	8	2.0	4.1
3C 345	A	June 1982	15	9	11.4
	B	January 1984	8	8	12.3
	C	June 1985	5	6	12.9

We carried out calculations for two classes of models—with elongated (jetlike) and spherically symmetrical blobs. In the former case, we assumed that the axis of symmetry of the jet was directed along the line of sight and that the magnetic field falls off along this axis with approach toward the observer according to the law

$$H(x) = \frac{H_0}{1 + k_H \left(\frac{x}{L}\right)^m} \text{ for } 0 \leq x \leq L, \quad (18)$$

$$H(x) = 0 \text{ for } x < 0 \text{ and } x > L.$$

Here, H_0 is the maximum magnetic-field strength at the center of the source, L is the extent of the jet along the line of sight, k_H characterizes the difference in the magnetic field along the source from the distant to the near edge, and m can take the values 1, 2, and 3 (details are presented in [24]).

In the case of a spherically symmetrical source, the distribution of the magnetic field is specified by the function

$$H(r) = \frac{H_0}{1 + k_H \left(\frac{r}{R}\right)^m} \text{ for } r \leq R, \quad (19)$$

$$H(r) = 0 \text{ for } r > R,$$

where R is the radius of the source and k_H characterizes the difference of the magnetic field from the center to the edge of the source.

We adopted the parameters of an average flare from [22] for the source parameters for the modeling. In all the models, the ratio H_{01}/H_{02} was set equal to 2.3. As is described above, we obtained this value for

a uniform source (13). The results of the calculations are presented in Tables 1 and 2. The results of the numerical modeling are in good agreement with the conclusions we drew based on our method. Thus, our method is not very sensitive to which of the models actually describes the real source.

4. EVOLUTION OF A BLOB IN THE QUASAR 3C 279

We attempted to study the dynamics of the variation of the physical parameters of a flare in the quasar 3C 279. Detailed light curves at various frequencies are available for this source, and simultaneous spectra have been constructed for various stages in the evolution of flares. We adopted the flare of 1984−1986 for our analysis. Light curves of this source are presented in [26−32]. We selected the most reliable spectra for the blob from [21], shown in the figure. We took the spectral index in the optically thin region of the spectrum to be $\alpha = 0.3$. Note that this value differs from the spectral index for an average blob presented in [22], $\alpha = 0.2$.

The light curves enabled us to estimate the characteristic size of the blob at the frequencies corresponding to the maxima of the spectra. The parameters of the spectra and estimated sizes of the blob are presented in Table 3, whose first six columns present (1) the name of the source, (2) designation for the flare, (3) date, (4) peak frequency of the blob ν_{max}, (5) maximum flux density of the blob S_{max}, and (6) estimated typical linear size of the blob derived from the light curves. Table 3 shows that the expansion of the blob is not relativistic and that this expansion remained nearly constant from 1981 to 1983.

The low-frequency turnover in the spectrum could be due to thermal absorption in a nonrelativistic plasma. However, we consider this mechanism to be improbable for the flare of 3C 279, since the spectral index in the optically thick region is close to 2.5, suggesting that the turnover is due to synchrotron self-absorption.

We were able to estimate the relative variations in the magnetic-field strength $H_\perp$, number density of relativistic electrons n_e, and energies of the magnetic field E_H and relativistic particles E_e by applying the method described above. The results are presented in Table 4. Evidence for constancy of the speed of the relativistic motion in 3C 279 (and, consequently, constancy of δ) in the period from 1981−1983 is presented in [33].

As in the average blob considered above, the magnetic field decreases and the number density of relativistic particles increases as the blob in 3C 279 expands. The energy density of the magnetic field also decreases, while the energy density of the relativistic

Table 4. Relative variations of the physical parameters

Source	Ratio	H	n_e	N_e	E_H	E_e
3C 120	A/B	24.2	0.04	0.04	540	0.01
	B/C	4.21	0.42	0.34	14.0	0.16
3C 273	A/B	6.89	0.40	0.37	44.3	0.14
	B/C	3.11	0.81	0.73	8.76	0.41
3C 279	A/B	22.6	0.04	0.04	397	0.01
	B/C	4.03	0.48	0.41	13.9	0.20
	C/D	5.59	0.40	0.35	26.9	0.15
3C 345	A/B	13.5	0.11	0.09	145	0.02
	B/C	4.88	0.33	0.29	20.6	0.13

particles grows. The maximum of the 3C 279 flare spectrum shifted from 37 to 22 GHz over 10 months. Over this time interval, the magnetic-field strength decreased by a factor of 20 and the number density of relativistic particles increased by a factor of 25; the energy density of the magnetic field fell by a factor of 400, while the energy density of the relativistic plasma grew by a factor of 100.

We also attempted to study the evolution of flares in the quasars 3C 120, 3C 273, and 3C 345. The light curves of these sources are less complete than that of 3C 279, so that our estimates are primarily qualitative rather than quantitative. The available observations of these sources [34−36] shows the presence of a low-frequency turnover in the blob spectrum. The parameters of the blob derived from the light curves are presented in Table 3. We estimated the spectral indices at high frequencies (in the transparency region of the spectrum) to be $\alpha = 0.1$ for the blobs in 3C 120 and 3C 345 and $\alpha = 0.6$ for the blob in 3C 273. Our estimates of the relative variations in the physical parameters are presented in Table 4. Overall, the evolution of the physical conditions in the blobs in these sources is similar to that for 3C 279.

5. CONCLUSIONS

We have used a simple method to estimate variations in the physical parameters of a blob of relativistic particles as it expands. Only observable quantities are present in our final formulas. The method itself is not fundamentally new, but the range of applicability of this "modernized" version has been expanded. The method was applied to variable sources with characteristic variability timescales of the order of hundreds of days, but this does not represent a real restriction, since the only serious limitation is the condition that the speed of the relativistic motion of the blob be constant. We also carried out numerical modeling of the relative variations in the physical parameters in blobs with nonuniform magnetic fields obeying a power-law distribution. These calculations show that we are able to use a simple model of a source with a uniform magnetic-field distribution, since the modeling results were virtually identical to those obtained in the first part of this work.

Our most interesting conclusions are the following: (1) there is an absence of equipartition between the energies of the magnetic field and relativistic particles during the expansion of the blob; (2) new relativistic particles are generated during the expansion.

REFERENCES

1. G. B. Sholomitskii, Astron. Zh. **42**, 673 (1965) [Sov. Astron. **9**, 516 (1965)].
2. W. A. Dent, Science **148**, 1458 (1965).
3. A. Witzel, D. S. Heeschen, C. J. Schalinski, and T. R. Krichbaum, Mitt. Astron. Ges. **65**, 239 (1986).
4. J. Dennett-Thorpe and A. G. de Bruyn, Astrophys. J. **529**, L65 (2000).
5. R. Altschuler, Fundam. Cosm. Phys. **14**, 37 (1989).
6. J. N. Bregman, Astron. Astrophys. **2**, 125 (1990).
7. S. J. Wagner and A. Witzel, Astron. Astrophys. **33**, 163 (1995).
8. E. B. Fomalont, S. Frey, Z. Paragi, *et al.*, Astrophys. J., Suppl. Ser. **131**, 95 (2000).
9. H. van der Laan, Nature **211**, 1131 (1966).
10. A. P. Marscher and W. K. Gear, Astrophys. J. **298**, 114 (1985).
11. P. A. Hughes, H. D. Aller, and M. F. Aller, Astrophys. J. **341**, 54 (1989).
12. P. A. Hughes, H. D. Aller, and M. F. Aller, Astrophys. J. **341**, 68 (1989).
13. P. A. Hughes, H. D. Aller, and M. F. Aller, Astrophys. J. **374**, 57 (1991).
14. E. Valtaoja, H. Teräsranta, S. Urpo, *et al.*, Astron. Astrophys. **254**, 71 (1992).
15. V. S. Artyukh, in Preprint of the Lebedev Physics Institute, Academy of Sciences of the USSR 189, 289, 1988, http://lnfm1.sai.msu.ru/~patch/css/
16. A. P. Marscher, Astrophys. J. **264**, 296 (1983).
17. A. P. Marscher, *Conf. Superluminal Radio Sources*, Ed. by J. A. Zensus and T. J. Pearson (1987), p. 280.
18. V. A. Artyukh and M. A. Ogannisian, Pis'ma Astron. Zh. **14**, 706 (1988) [Sov. Astron. Lett. **14**, 301 (1988)].
19. V. I. Slysh, Nature **199**, 682 (1963).
20. A. G. Pacholczyk, *Non-thermal Processes in Galactic and Extragalactic Sources* (Freeman, San Francisco, 1970).
21. E. Valtaoja, S. Haarala, H. Lehto, *et al.*, Astron. Astrophys. **203**, 1 (1988).

22. E. Valtaoja and H. Teräsranta, Astron. Astrophys. **289**, 35 (1994).

23. S. A. Tyul'bashev and P. A. Chernikov, Astron. Astrophys. **373**, 381 (2001).

24. V. S. Artyukh and P. A. Chernikov, Astron. Zh. **78**, 1 (2001) [Astron. Rep. **45**, 16 (2001)].

25. A. P. Lobanov, Astron. Astrophys. **330**, 79 (1998).

26. T. J. Balonek, Ph.D. Thesis (Univ. of Massachusetts, 1982).

27. E. Salonen, H. Teräsranta, S. Urpo, *et al.*, Astron. Astrophys. **70**, 409 (1987).

28. A. M. Flett and C. Henderson, Mon. Not. R. Astron. Soc. **204**, 1285 (1983).

29. H. D. Aller, M. F. Aller, G. E. Latimer, and P. E. Hodge, Astrophys. J., Suppl. Ser. **59**, 513 (1985).

30. W. K. Gear, L. M. J. Brown, E. I. Robson, *et al.*, Astrophys. J. **304**, 295 (1986).

31. R. Barvainis and C. R. Predmore, Astrophys. J. **282**, 402 (1984).

32. T. F. Haddock, Ph.D. Thesis (Univ. of Michigan, 1984).

33. S. C. Unwin, M. H. Cohen, J. A. Biretta, *et al.*, Astrophys. J. **340**, 117 (1989).

34. T. L. Roellig, E. E. Becklin, C. D. Impey, and M. W. Werner, Astrophys. J. **304**, 646 (1986).

35. E. I. Robson, W. K. Gear, P. E. Clegg, *et al.*, Nature **205**, 194 (1983).

36. D. C. Backer, in *IAU Symp. No. 100: VLBI and Compact Radio Sources*, Ed. by R. Fanti, K. Kellerman, and G. Setti (Reidel, Dordrecht, 1984), p. 31.

Translated by D. Gabuzda

Magnetorotational Supernova Explosions and the Formation of Neutron Stars in Close Binary Systems

G. S. Bisnovatyĭ-Kogan[1] and A. V. Tutukov[2]

[1]*Space Research Institute, Moscow, Russia; Unified Institute of Nuclear Research, Dubna, Russia*
[2]*Institute of Astronomy, Moscow, Russia*
Received March 12, 2004; in final form, March 15, 2004

Abstract—The formation of neutron stars in the closest binary systems ($P_{orb} < 12$ h) gives the young neutron star/pulsar a high rotational velocity and energy. The presence of a magnetic field of 3×10^{11}–3×10^{13} G, as is observed for radio pulsars, enables the neutron star to transfer $\sim 10^{51}$ erg of its rotational energy to the envelope over a time scale of less than an hour, leading to a magnetorotational supernova explosion. Estimates indicate that about 30% of all type-Ib,c supernovae may be the products of magnetorotational explosions. Young pulsars produced by such supernovae should exhibit comparatively slow rotation ($P_{rot} > 0.01$ s), since a large fraction of their rotational angular momentum is lost during the explosion. The magnetorotational mechanism for the ejection of the envelope is also reflected by the shape of the envelope. It is possible that the Crab radio pulsar is an example of a product of a magnetorotational supernova. A possible scenario for the formation of the close binary radio pulsar discovered recently by Lyne *et al.* is considered. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

Studies of the physics of supernovae and modeling of supernovae represent a well-developed area of modern astrophysics [1]. In current models, a type-Ia supernova results from either the explosion of the core of a degenerate carbon–oxygen dwarf with a mass equal to the Chandrasekhar mass in a close, semidetached binary [2] or the merging of degenerate components in a binary system [3]. Type-II supernovae result primarily from the collapse of the core of a red or blue supergiant with an initial mass exceeding $\sim 8\ M_\odot$, accompanied by the formation of a neutron star (when the initial mass of the star is less than $\sim 25\ M_\odot$) [1]. The energy of these supernovae is primarily gravitational energy that is released during the formation of the neutron star. It is possible that some type-II supernovae in low-mass, metalpoor galaxies are associated with the thermonuclear explosion of the carbon–oxygen core of a red, or more often blue, supergiant with an initial mass of 6–$8\ M_\odot$ that does not leave behind any remnant. The low heavy-element abundance weakens the stellar wind of the red supergiant, making it easier for its carbon–oxygen core to attain the Chandrasekhar limit. Type-Ib,c supernovae exhibit a deficit, or even a total absence, of hydrogen in their spectra and probably result from the explosions of helium or carbon–oxygen stars, primarily the components of massive, close binary systems that have lost their extended,

hydrogen-rich envelopes in earlier stages of their evolution [4, 5]. It is thought that the mechanism for these supernovae is the same as that for type-II supernovae, since the core structures of their precursors are the same. Theoretical estimates of the masses of the helium precursors to type-Ib,c supernovae are in the range ~ 2–$10\ M_\odot$ when a neutron star is formed and $>10\ M_\odot$ when a black hole is formed [6].

The magnetorotational mechanism for a supernova explosion and the acceleration and ejection of the supernova envelope was proposed in 1970 [7, 8]. The idea behind this scenario is simple. A young neutron star rotating near the rotational-stability limit has a rotational energy of $\sim 3 \times 10^{52}$ erg. Therefore, the transfer of even a small fraction of this energy to the supernova envelope via its rather strong magnetic field can explain the energetics of the supernova explosion, even without the participation of the neutrinos that carry away most of the released gravitational energy. Recent two-dimensional computations fully confirm the efficiency of this mechanism [9]. The necessary magnetic field is initially a relict field that is rapidly amplified during the collapse and subsequent differential rotation of the neutron star [7, 10]. It is important that this mechanism can operate for both type-II and type-Ib,c supernovae whenever there is a collapse with the formation of a neutron star as long as the compact, degenerate cores of the presupernovas with radii of $\sim 10^8$–10^9 cm rotate sufficiently rapidly.

It is obvious that evaluating the perspectives for the magnetorotational mechansim requires information about the rotational speeds of the cores of presupernovas. Unfortunately, this information is quite limited. Direct observational data are completely lacking. Some numerical simulations have led to the conclusion that the the masses of rotating degenerate presupernova cores may exceed those of nonrotating cores [11]. Studies of the rotation of degenerate dwarfs provide indirect evidence about the rotational speeds of asymptotic-branch stars and their derivatives, and degenerate dwarfs have similar masses, sizes, and origins to the cores of presupernovas with initial masses of about 10 $M_\odot$. As a rule, spectroscopic studies of the observed rotational speeds of degenerate dwarfs yield values below 10 km/s [12]. More accurate astroseismological methods give even lower rotational speeds, down to 0.6 km/s [12]. Converting these speeds to the equatorial rotational speeds of their precursors, which had sizes two orders of magnitude larger than those of the degenerate dwarfs, leads to implausibly low rotational speeds for the main-sequence stars that produced them: $\sim 10-100$ m/s. Applying this procedure to young radio pulsars with rotational periods of ~ 0.01 s leads to estimated rotational speeds for their O−B main-sequence precursors of ~ 6 km/s, which is low compared to the characteristic observed rotational speeds for these stars: $100-500$ km/s. This may indicate the action of efficient braking of the core rotation of single stars, and probably also the components of wide binary systems, during the course of their evolution.

If the transfer of angular momentum is sufficiently efficient to support rigid-body rotation of the star, it will rotate so slowly when it becomes a supergiant that the young neutron star that forms as a result of the core collapse will rotate too slowly (with a period of ~ 100 s [13]) to become a radio pulsar. Ordinary radio pulsars may be produced by the evolution of components in massive close (interacting in the past) binaries with initial masses of $8-25\ M_\odot$ [14]. This can also explain the observed bimodal speed distribution of radio pulsars [15] as an effect of the disruption of close binary systems during the explosion of the secondary as a supernova [16]. The observed correlation between the axial (P_p) and orbital (P_{orb}) rotational periods of the nine currently known radio pulsars in close binaries with nonzero eccentricities ($P_p = 10^{-6} P_{orb}$) confirms the possibility of an orbital origin for the axial rotation of radio pulsars [14, 16] and the efficiency of spin−orbital synchronization of the rotations of the components of close binary systems.

The observed orbital eccentricities of known radio pulsars paired with neutron stars in binaries can be used to estimate the minimum masses of the neutron-star precursors, based on computations of the evolution of the components of massive close binary systems carried out earlier and assuming the absence of an appreciable "kick" velocity. PSR J1829+24, with an orbital period of 1.2 d and a rotational period of 0.04 s, has a neutron-star companion in an orbit with eccentricity $e = 0.14$ [17]. Taking the current total mass of this system to be 2.8 $M_\odot$, we find that the mass lost during the supernova explosion was $m = e \times 2.8 M_\odot = 0.4\ M_\odot$. This means that the mass of the helium presupernova at the epoch of the explosion was $\sim 1.8\ M_\odot$, and the initial mass of the main-sequence star was $\sim 8\ M_\odot$ [18]. This is thus an estimate for the minimum initial mass of a single star or a component in a wide binary that produces a supernova with a neutron star as its final remnant. This estimate is also supported by the absence of any star whose brightness exceeds that of a star with a mass of more than $\sim 9\ M_\odot$ in the location of the type-II SN 1999gi [19].

The maximum initial mass of stars forming neutron stars coincides with the minimum mass of stars producing black holes. According to estimates based on the radiation of observed close binaries with black holes, this latter mass is close to 25 $M_\odot$ [6]. Note that the position of this boundary is essentially independent of whether or not the star is a member of a close binary, due to the nearly completely isolated evolution of the helium core of the massive star [18]. One theoretical estimate of the rate of supernovae forming neutron stars, based on a semiempirical function for the formation of binary systems [3], yields the value ~ 0.03 per year [20]. This coincides with the observed rate of supernovae in the Galaxy if its luminosity is taken to be $\sim 3 \times 10^{10}\ L_\odot$ [21].

2. PULSARS IN CLOSE BINARY SYSTEMS

Some neutron stars form in close binary systems. The observed correlation between the axial and orbital rotations of neutron stars suggests the formation of young pulsars with rotational periods of 0.005−5 s [14, 16]. In another scenario, radio pulsars form with fairly long periods of 30−100 ms, since a large fraction of the rotational energy is lost during the magnetorotational supernova explosion [22]. This model is consistent with statistical studies of radio pulsars [23]. Millisecond pulsars represent a special class of spun-up pulsars, whose orbital rotation has been accelerated due to disk accretion in an X-ray binary stage, accompanied by a decrease in the magnetic field of the neutron star due to screening by the accreting matter [24]. Evidence for this scenario is provided by the facts that all millisecond pulsars have magnetic fields that are two to four orders of magnitude lower than the
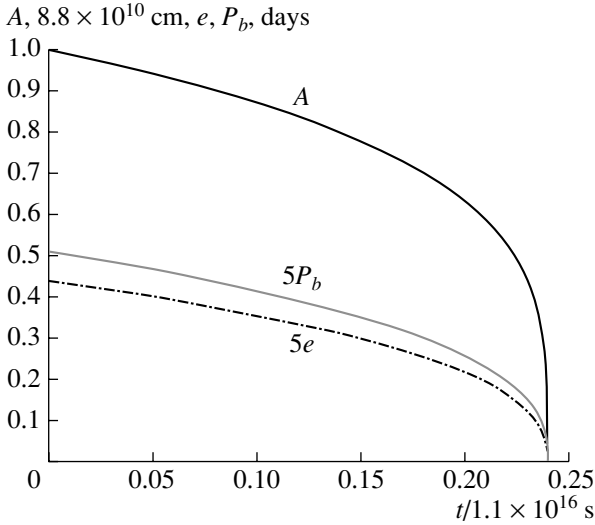
A, 8.8×10^{10} cm, e, P_b, days

Fig. 1. Evolution of the semimajor axis A, orbital period P_b, and eccentricity e of the binary system due to the action of gravitational radiation.

A, 8.8×10^{10} cm, e, P_b, days

Fig. 2. Possible past evolution of the semimajor axis A, orbital period P_b, and eccentricity e of the binary system due to the action of gravitational radiation.

mean value and single millisecond pulsars are concentrated in globular clusters, where the probability of losing a companion during collisions with other stars is most likely [25]. The discovery of a system consisting of two radio pulsars [26], one a spun-up millisecond pulsar (PSR J0737-3039A, $P = 22.7$ ms, $B = 6.3 \times 10^9$ G) and the other an ordinary pulsar (PSR J0737-3039B, $P = 2.8$ s, $B = 1.6 \times 10^{12}$ G), is in full agreement with this scheme. The millisecond pulsar formed as a result of the first supernova in the system, which then passed through an X-ray binary stage in which the pulsar's rotational speed was increased and its magnetic field decreased. The evolution of the companion also ended in a supernova, which led to the formation of the second pulsar, which did not pass through an X-ray binary phase and, therefore, acquired fairly standard properties. A scenario for the formation of such a pair of pulsars is considered in [24].

Knowing the current parameters of a given system, we can easily trace its subsequent and previous evolution, since the only physical process that changes the parameters of the binary system is the emission of gravitational radiation. The magnetic stellar winds and tidal interactions that can act in ordinary binary systems are not important in systems with two neutron stars. The variation of the semimajor axis A and eccentricity e of a binary with stars of masses m_1 and m_2 under the action of gravitational radiation is determined (after averaging over the period) by the system of equations [27]

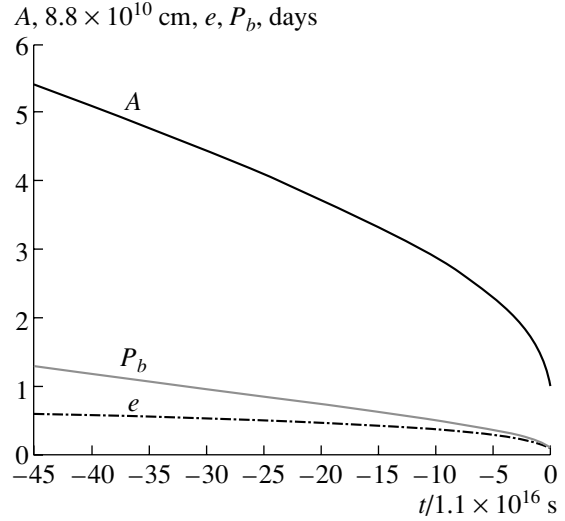$$\frac{dA}{dt} = -\frac{64}{5} \frac{m_1 m_2 (m_1 + m_2)}{a^3 (1 - e^2)^{7/2}} \tag{1}$$

$$\times \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right),$$

$$\frac{de}{dt} = -\frac{304}{15} \frac{m_1 m_2 (m_1 + m_2) e}{A^4 (1 - e^2)^{5/2}} \left(1 + \frac{121}{304} e^2 \right). \tag{2}$$

The current parameters of the binary are [26]

$$m_1 = 1.34 \, M_\odot \, (\text{millisecond pulsar}), \tag{3}$$

$$m_2 = 1.25 \, M_\odot, \quad e_0 = 0.0878,$$

$$P_{b0} = 0.102 \text{ days} = 8.83 \times 10^3 \text{ s},$$

$$A_0 = 8.8 \times 10^{10} \text{ cm}.$$

The last value in (3) is obtained from Kepler's law, taking the masses of the stars and the period of the binary P_{b0} to be known. Integrating system (1)–(2), we can obtain the dependences $A(t)$, $e(t)$, and $P_b(t)$ for the future (Fig. 1) and the past (Fig. 2). According to Fig. 1, the lifetime of the system until it coalesces is 84 million years, in agreement with the estimate presented in [26]. Integration into the past shows that the parameters of the system have changed very little over the characteristic lifetime of the ordinary (50 million years) and millisecond (210 million years) pulsars (Fig. 2). At the same time, it is difficult to imagine that the eccentricity e of the system has not increased after the collapse and explosion of the secondary, as takes place in other binary systems containing two neutron stars, such as PSR 1913+16 ($e = 0.617$) and PSR 2127+11C ($e = 0.68$). Thus, we must suppose either that ages of the pulsars based on measurements of the deceleration of their rotation are severely underestimated or that the second core collapse in the system occurred virtually without any ejection of

material, so that the eccentricity was not increased. If we are dealing with a rather old system with an age of several billion years, assuming the system's parameters have evolved only due to the emission of gravitational radiation, then seven billion years ago, the system had $A = 3.76A_0$, $P = 0.73$ days, and $e = 0.5$. If we consider the ages of pulsars based on the deceleration of their rotation to be valid, then we must suppose that the collapse and formation of the second neutron star occurred with virtually no ejection of material. It is possible that there was a complete equalization of the periods of the axial and rotational rotations in this close pair. In this case, the rotational period of the second neutron star after its formation would be $10-100$ ms, and its energy would not be sufficient for a magnetorotational explosion. Evidence for intense tidal relaxation is provided by the fact that both stars in the system are observed as pulsars, which is possible only if their beams are oriented in similar directions.

The rotational energy of a neutron star with a period of ~ 0.001 s exceeds the typical energy of a supernova explosion by about a factor of 10. If it has a strong magnetic field, the energy of a such a young neutron star is sufficient to account for the supernova, even if the flux of neutrinos preceeding the collapse were not able to do so. One-dimensional computations in a cylindrical model have shown that approximately 10% of the rotational energy is transformed into explosive energy (see, for example, [22]). Two-dimensioinal computations of a magnetorotational supernova explosion [9] yield an even higher coefficient for the transformation of rotational into explosive energy, suggesting it is possible to produce a fairly strong explosion with an initially relatively slowly rotating neutron star. One of the goals of the current paper is to elucidate the relative rate of magnetorotational explosions and to obtain a description of close binaries with compact, helium, nondegenerate companions—possible precursors of such supernovae.

3. HELIUM COMPONENTS IN CLOSE BINARIES

According to the empirical distribution of orbital semimajor axes for close binaries, $20-40\%$ of massive supernovae occur in such systems [3]. The corresponding presupernova stars do not have the extended, hydrogen-rich envelopes that are typical of type-II supernovae and are compact ($\sim R_\odot$) helium stars [18, 21]. The explosions of such stars, whose envelopes are (virtually) completely devoid of hydrogen and are enriched in carbon and oxygen, give rise to type-Ib,c supernovae. Their observed relative rate is estimated to be ~ 0.2 [28], consistent with theoretical

estimates within the uncertainties in the rates of both types of supernovae. SNIb,c are close to SNII in their host-galaxy distribution. A statistical study of 177 supernovae demonstrated that both types of supernovae occur only in spiral galaxies, i.e., in galaxies with active ongoing star formation [29]. This underscores the youngness and similarity of the various supernova precursor stars, which are all massive, short-lived stars.

Thus, we will assume that all presupernovas in massive close binary systems are helium stars, possibly with a relatively low-mass remnant hydrogen envelope if it has not been swept away by the stellar wind. If the initial masses of the components on the main sequence were $8-25$ $M_\odot$, the masses of the helium stars should be in the range $2-10$ $M_\odot$ [18]. Let us now consider what is known about the observational properties of systems containing helium stars with masses in this range. Helium stars with masses exceeding ~ 5 $M_\odot$ are well known as Wolf–Rayet stars, some of which are members of close binaries with known component masses [6]. Due to the intensity of their stellar winds, the sizes of their photospheres appreciably exceed the sizes of the relatively compact Wolf–Rayet stars ($\sim R_\odot$). This significantly lowers the effective temperature of the radiation and corresponds to a bolometric correction which, for the high luminosities of Wolf–Rayet stars ($\sim 10^5$ $L_\odot$), makes them detectable in a large volume of our own Galaxy and in many nearby galaxies. An example of a helium star with a mass in the relevant interval is the Wolf–Rayet star B32 [6], which has a mass of ~ 5 $M_\odot$ and forms a close pair with a main-sequence star with a mass of ~ 30 $M_\odot$. The orbital period of the system is 1.9 days. According to our current understanding of stellar evolution, the explosion of the Wolf–Rayet star as a type-Ib,c supernova is inevitable. In several cases, the intense (10^{37} erg/s) X-ray emission of type-Ib,c supernovae has been attributed to the interaction between the ejected envelope of the supernova and material that was blown out by the intense stellar wind of the Wolf–Rayet precursor [30]. The origin of this intense stellar wind is probably pulsational instability of the helium stars [6], possibly combined with the effect of radiation pressure.

The detection of nondegenerate helium stars with lower masses that are in the core helium-burning stage is hindered primarily by the absence of a strong stellar wind in these stars, which, given their high surface temperatures, leads to a large bolometric correction ($\sim 3^m - 5^m$) that makes them weak at visible wavelengths. However, a small group of such stars is known; these are enriched in helium and have spectra reminiscent of the spectra of Wolf–Rayet stars, and they may belong to close binaries with orbital periods of $5-12$ h [31]. One example is the SPH2 system,

whose helium component has a spectra that resembles the spectrum of a WN4 star [31]. In terms of numbers of atoms, the helium content of its envelope is comparable to the hydrogen content. It is possible that some stars in this family are contained in semidetached close binaries with degenerate dwarfs accreting matter at a rate of $\sim 10^{-7}$ $M_\odot$/yr. This accretion rate sustains stationary hydrogen burning in the envelope of the degenerate dwarf. Such systems are also supersoft X-ray sources. Another way to explain the periodic variability of the radial velocities of these stars (apart from the possibility that they are in binary systems) is to suppose that they are pulsating low-mass helium stars. However, the observed periods for the radial-velocity variability appreciably exceed the expected values for main-sequence helium stars (~ 10 min). It is therefore possible that this group of stars includes nondegenerate helium stars with masses of $2-5$ $M_\odot$ that are precursors of type-Ib,c supernovae in close binary systems. The estimated mass for one of these "dwarf" Wolf−Rayet stars, HD 45166, is ~ 3.5 $M_\odot$ [31], so that it is indeed a member of the group in which we are interested. The star WR46, whose brightness varies with a period of 2.3 h, may also belong to this group [31, 23]. It is probable that stars of this group make up the majority of type-Ib,c precursors.

As a rule, nondegenerate low-mass helium stars are known as sdB and sdO stars. The majority of known stars of this type (to $\sim 70\%$) are in close binaries with orbital periods of $0.1-10$ days [31–37]. The formation of such stars during the evolution of the close binaries has been considered in many studies [18, 39−41]. These computations have shown that the mass distribution of the helium remnants is continuous in the interval $0.5-49$ $M_\odot$. However, it is usual when estimating the parameters of sdB and sdO subdwarfs to adopt masses close to 0.5 $M_\odot$, similar to the masses of helium stars at the blue tip of the horizontal branch for globular clusters. In reality, the observed high rate of binarity of these stars suggests that their masses may reach $2-3$ $M_\odot$ [40]. The absence of a detectable stellar wind and the small number of such helium stars hinders searches for additional such stars in our Galaxy. However, several examples have been found. The observed parameters of many sdO stars ($\log g = 4.5$ and $\log T_e = 4.8$ [35, 42]) agree with the expected parameters for helium remnants in close binaries with initial masses of $5-$ 10 $M_\odot$ and current masses of $1-2$ $M_\odot$ [39]. In addition, the parameters of several systems with nondegenerate helium components of the desired type have been determined [43]. For example, ϕ Per contains a main-sequence star with a mass of ~ 21 $M_\odot$ and a helium star with a mass of ~ 3.4 $M_\odot$ with an orbital period of 127 days. The past evolution of this system

can be described in a picture with conservative mass-transfer in a system whose components had nearly equal initial masses of $12-13$ $M_\odot$. The system V Sgr, with an orbital period of 138 days [43], contains a helium star with a mass of ~ 2.5 $M_\odot$ and a main-sequence star with a mass of ~ 4 $M_\odot$. Judging from these component masses, the mass transfer during the previous evolution in this system was clearly nonconservative. The main conclusion we can draw from this section is that helium stars with masses in the range $2-10$ $M_\odot$ that are members of close binary systems, as is predicted by computations of the evolution of close-binary components, are indeed known. Further observations are needed to increase the numbers of known objects in this class and to refine the main parameters and the statistics of these stars.

4. MAGNETOROTATIONAL SUPERNOVA EXPLOSIONS

The formation of radio pulsars is usually considered a natural and inevitable consequence of a supernova explosion accompanied by the formation of a neutron star. The strong magnetic field and rapid rotation are considered to be the result of the collapse of a degenerate, Chandrasekhar-mass core to the density of a neutron star with a radius that is a factor of $\sim 300-1000$ smaller than the core radius. This speeds up the rotation and amplifies the magnetic field by factors of $\sim 10^5-10^6$. The amplification of the frozen-in magnetic field during the rapid compression of the core to values of $\sim 10^{12}$ G observed in young radio pulsars seems to be inevitable [7, 10, 44, 45].

The situation with the rotation of young neutron stars is less well defined. To obtain a neutron star with a rotational period shorter than ~ 1 s, the rotational period of the presupernova core must be shorter than ~ 10 days. However, if a single star or component of a wide binary that is a presupernova supergiant with a radius of ~ 1000 $R_\odot$ conserves the angular momentum of the original O,B main-sequence star, its rotational period should be ~ 30 yrs [14]. Consequently, in the case of rigid-body rotation of such a presupernova, the rotational period of the young neutron star that results from the colapse of the supergiant core will be ~ 1000 s, which precludes the appearance of a radio pulsar. Neutron stars with long rotational periods in wide binaries have been found, but the origin of their long periods is probably associated with accretion [46]. The X-ray pulsar GX1+4, which forms a pair with an M supergiant, has a rotational period of ~ 140 s [47]. The doubling time for this period is only ~ 120 yrs, which makes it unclear what the period for its rotation was initially. It is possible that long-period X-ray pulsars, like many pulsars with periods of about a second, are located in a state of nearly

equilibrium rotation and undergo irregular variations of their periods about their equilibrium values [48].

The rapid rotation of a young neutron star requires that the presupernova was a component in a close binary system with an orbital period shorter than ~ 10 days. Such supernovae are of type-Ib,c [5]. According to the known semimajor axis distribution for close binaries [39], events leading to the appearance of a radio pulsar with a rotational period shorter than ~ 6 s make up about $\sim 30\%$ of the total number of type-Ib,c supernovae, i.e., about 0.01 such events per year in our Galaxy. Observational estimates of this rate are in the range $0.003-0.03$/year [48–50] but remain uncertain due to uncertainties in the translation of the observed statistics for pulsars into the rate of their formation. The absence of a direct, unambiguous relationship between supernovae and pulsars is also reflected by the following fact: of two hundred carefully studied supernova remnants, radio pulsars have been detected in only nine [52]. Of course, some of this difference should be ascribed to the directional beams of the pulsar radio emission, but some is likely a consequence of a lower rate for the formation of radio pulsars than the rate of supernovae.

The possibility that the axial and orbital rotations of the presupernova are partially synchronized can explain the observed correlation between the axial and orbital rotations of radio pulsars in close binaries with eccentric orbits ($P_{rot} = 10^{-6} P_{orb}$) [14]. In addition, observations of radio pulsars and helium subdwarfs in close binaries show that both of these types of object are well represented in the required interval of orbital periods [14, 53]. Since the radii of helium stars with masses of $2-6\ M_\odot$ are [18]

$$R_{\text{He}}/R_\odot = 0.2(M_{\text{He}}/M_\odot)^{0.65}, \qquad (4)$$

the minimum orbital periods of semidetached systems with such donors should be $\sim 0.01(M_{\text{He}}/M_\odot)^{0.5}$ days. If the axial and orbital rotations of a compact presupernova are synchronized, the rotational period of the young pulsar that is formed will be $\sim 0.001(M_{\text{He}}/M_\odot)^{0.5}$. The rotational energy of a neutron star with a uniform density is

$$E_\omega = \frac{4\pi^2 M R^2}{5 P_p{}^2} = 2 \times 10^{46} P_p{}^{-2}\ \text{erg.} \qquad (5)$$

In this last estimate, we have assumed that the mass of the neutron star is $1.4\ M_\odot$ and its radius is 10^6 cm. It is obvious that, if $P_p < 0.0045$ s ($P_{orb} < 0.05$ days), the rotational energy of the neutron star will exceed 10^{51} erg—the energy associated with an ordinary type-Ib,c supernova. This admits the possibility, in the presence of a sufficiently intense magnetic field,

that an appreciable, if not the dominant, contribution to the energy of a supernova of this type is made by the rotational energy of the young neutron star. We should note that the observed energies of supernovae associated with the formation of neutron stars can range over two orders of magnitude [53–55]. On this basis, it is not ruled out that there exist weak supernovae with energies lower than the indicated limit. This would make it possible to raise the upper limit for the orbital periods of close binaries capable of producing magnetorotational supernova explosions. On the other hand, the latest stages in the evolution of the presupernova core are passed through rather rapidly, and there is not sufficient time for equalization of the axial and orbital rotations of the core. In this case, the angular velocity of rotation of the presupernova core can be much higher than the orbital velocity, and the requirement on the orbital periods in systems in which magnetorotational supernova explosions are possible is weakened.

When estimating the rate of magnetorotational explosions required for systems containing nondegenerate helium stars paired with lower-mass compact companions, it is sufficient to determine the minimum orbital period P_{min} that is able to provide the rotational energy of the young neutron star (more than 10^{51} erg). Adopting the radius of the helium star in accordance with (4), we find for the maximum orbital period of such systems with synchronous rotation $P_{max}^s = 0.0025(M_{\text{He}}/M_\odot)$ days. Let us suppose that the rotation of the presupernova core is a factor of 10 faster than the orbital rotation; this will increase the maximum orbital period by a factor of 10, to $P_{max} = 0.025(M_{\text{He}}/M_\odot)$ days. Now, comparing the two limits on the orbital periods of helium stars with masses of $3\ M_\odot$ indicated above, we find that binaries with $\delta \log P_{orb} = 1$ or $\delta \log(A/R_\odot) = 0.7$ (where A is the semimajor axis of the orbit) produce young neutron stars with rotational energies above $\sim 10^{51}$ erg. Using the known orbital major-axis distribution for binaries, $dN = 0.2 d\log A$ within $A = 10-10^6\ R_\odot$ [3], we find the relative rate of magnetorotational supernovae to be ~ 0.14. Since, as is indicated above, analogous estimates indicate the rate of type-Ib,c supernovae to be ~ 0.4, we conclude that about one-third of all supernovae of this type can have a magnetorotational nature. Given the uncertainties in these estimates, it is also possible that all type-Ib,c supernovae are produced by the magnetorotational mechanism.

The computations of a magnetorotational explosion carried out in [9] show that, for the typical magnetic fields of young pulsars, $B_{psr} \sim 10^{12}-10^{13}$ G [57], the time scale for the development of a magnetorotational explosion is $10-100$ s if the neutron star initially has a millisecond period and the initial field is quadrupolar. The bulk of the time is spent

amplifying the magnetic field via differential rotation, which proceeds linearly in time:

$$B = B_0 \left(1 + \frac{t}{P_s} \right). \qquad (6)$$

We can obtain a simple estimate of the time for the process to develop using the fact that the rapid stage of the explosion begins when the local magnetic-field pressure reaches the pressure of the matter in the region of the maximum field amplification. The computations of [9] show that the pressure and density in the region of maximum fied amplification are $\rho \sim 10^{12}$ g/cm^3 and $P \sim 4 \times 10^{30}$ dyne/cm^2. Comparing with (6), we obtain for the time scale for the development of the magnetorotational explosion

$$t_{mr} \sim \frac{P_s \sqrt{8\pi P}}{B_0}. \qquad (7)$$

This time scale is $10-100$ s for $P_s = 10^{-3}$ s, $B_0 = 10^{11}-10^{12}$ G, and the matter pressure indicated above. Since during differential rotation only the radial field component is amplified and the period grows with distance from the center of the star, this estimate corresponds to a pulsar magnetic field that is approximately an order of magnitude higher than the observed values.

Estimation of the rotational speed of the core of a single presupernova is made more difficult by the absence of data on the angular-velocity distribution in very nonuniform giants and supergiants with dense cores and tenuous extended envelopes. If rigid-body rotation of the entire star is maintained, an isolated presupernova will rotate very slowly, and the magnetorotational mechanism will be ineffective. However, if the growth of the magnetic field due to differential rotation at the boundary of the core is suppressed by dissipative processes in the region of strong convection, and the core becomes "detached" from the envelope, it can continue to rotate rapidly during the presupernova stage. The magnetorotational mechanism can then operate in this case of single stars collapsing to produce neutron stars, probably associated with type-II supernovae.

The efficiency of the braking of the core rotation for moderate-mass stars that end their evolution in the production of a white dwarf can be verified empirically by analyzing the observed rotational speeds of single degenerate dwarfs. The observed rotational periods of these stars are from several hours to several days, including dwarfs in star clusters whose main-sequence turnoff points correspond to masses of $2-3$ $M_\odot$ [58]. In the case of conservation of the local angular momentum and rigid-body initial rotation, the expected rotational periods of young degenerate dwarfs should be only a few minutes, nearly three

orders of magnitude smaller than the observed values. It follows that the cores of single stars give a large amount of their angular momentum to the envelope as they are compressed, probably via their internal magnetic fields. It seems possible that the magnetorotational mechanism could lead to the ejection of the envelope in the form of a planetary nebula, which carries away the angular momentum of the remaining degenerate dwarf. However, it is clear that special studies are required to estimate the core rotational speeds of red supergiants.

5. CONCLUSIONS

We cannot rule out the possibility that single stars and the components of noninteracting, wide binaries that are at the end of their evolution can produce only slowly rotating degenerate dwarfs, neutron stars, and black holes due to the efficient braking of the dense cores of the evolved stars by their magnetic field. It seems most plausible to obtain rapidly rotating final end products of stellar evolution from stars in close binary sytems. The rapid rotation of the cores of pre-type-Ib,c supernovae, which is forced by the orbital rotation of the binary, leads in this case to the formation of rapidly rotating neutron sars—pulsars [14]—and Kerr black holes [6]. The formation of a Kerr black hole surrounded by a massive, rapidly accreting disk is probably a necessary condition for the developmemt of a gamma-ray burst [59]. Such a system can result from the collapse of a Wolf−Rayet star that is a member of a close binary system [6]

This hypothesis is supported by the observational detection of signs of intense stellar winds surrounding the sources of prolonged (>2 s) bursts of gamma radiation. Short (<2 s) bursts could result from the merging of the components of close neutron-star binares under the action of the radiation of gravitational waves by these systems [61−63]. Alternative views of the relationship between short gamma-ray bursts and soft gamma-ray repeaters are presented in [64, 65].

Rapidly rotating neutron stars with rotational energies sufficient to explain supernovae produced via the magnetorotational mechanism are probably formed during core collapses in close binaries with helium presupernovas.

The rotational period of a young neutron star of this type can reach $\sim$0.001 s, and the orbital period of a binary sytem with a helium presupernova component with a mass of $2-10$ $M_\odot$ is less than several hours. The transfer of some of the rotational energy to the supernova envelope via the magnetic field can explain not only the energetics of supernovae associated with a core collapse that leads to the formation of a neutron star but also the morphology of the

supernova envelope (supernova remnant). The results of studies of the expansion of supernova remnants arising as a result of an anisotropic central explosion are presented in the review [66]. The formation of jet-like ejections during the collapse of a rapidly rotating neutron star with a strong magnetic field was considered in [67, 68]. It is possible that the Crab Nebula displays traces of such jets in its expanding envelope. This question is worthy of further investigation.

ACKNOWLEDGMENTS

REFERENCES

1. D. Arnett, *Supernovae and Nucleosynthesis* (Princeton Univ. Press, Princeton, 1996).
2. J. Whelan and I. Iben, Jr., Astrophys. J. **186**, 1007 (1973).
3. I. Iben, Jr. and A. Tutukov, Astrophys. J., Suppl. Ser. **54**, 335 (1984).
4. T. Shigeyama, K. Nomoto, *et al.*, Astrophys. J. **361**, L23 (1991).
5. A. Tutukov and N. Chugaĭ, Pis'ma Astron. Zh. **18**, 605 (1992) [Sov. Astron. Lett. **18**, 242 (1992)].
6. A. Tutukov and A. Cherepashchuk, Astron. Zh. **80**, 419 (2003) [Astron. Rep. **47**, 386 (2003)].
7. G. S. Bisnovatyĭ-Kogan, Astron. Zh. **47**, 813 (1970) [Sov. Astron. **14**, 652 (1970)].
8. G. S. Bisnovatyi-Kogan and S. I. Blinnikov, Astron. Zh. **59**, 876 (1982) [Sov. Astron. **26**, 530 (1982)].
9. S. G. Moiseenko, G. S. Bisnovatyi-Kogan, and N. V. Ardeljan, astro-ph/0310142.
10. D. L. Meier, R. I. Epstein, W. D. Arnett, and D. N. Schramm, Astrophys. J. **204**, 869 (1976).
11. R. Hirshi, G. Meynet, and A. Maeder, astro-ph/0309774.
12. S. Kawaler, astro-ph/0301539.
13. I. Iben, L. Tutukov, and A. Fedorova, Astrophys. J. **486**, 955 (1997).
14. A. V. Tutukov and A. V. Fedorova, Astron. Zh. **80** (2004, in press).
15. Z. Arzoumanian, D. Chernov, and J. Cordes, Astrophys. J. **568**, 289 (2002).
16. I. Iben, Jr. and A. Tutukov, Astrophys. J. **456**, 738 (1996).
17. M. McLaughlin, D. Lorimer, D. Champion, *et al.*, astro-ph/0310454.
18. A. V. Tutukov, L. R. Yungel'son, and A. Klyaĭman, Nauchn. Inform. Astron. Sov. Akad. Nauk SSSR **27**, 3 (1973).
19. S. Smartt, G. Gilmore, N. Trentham, *et al.*, Astrophys. J. **556**, L29 (2001).
20. A. V. Tutukov and L. R. Yungel'son, Astron. Zh. **79**, 738 (2002) [Astron. Rep. **46**, 667 (2002)].
21. E. van den Heuvel and J. Heize, Nat. Phys. Sci. **239**, 67 (1972).
22. G. S. Bisnovatyi-Kogan, *Stellar Physics. Stellar Evolution and Stability* (Springer, Berlin, 2002), Vol. 2.
23. D. Bhattacharya, *Neutron Stars*, Ed. by J. Ventura and D. Pines, NATO ASI Ser. C **344**, 103 (1991).
24. G. S. Bisnovatyi-Kogan and B. V. Komberg, Astron. Zh. **51**, 373 (1974) [Sov. Astron. **18**, 635 (1974)].
25. G. S. Bisnovatyi-Kogan, Astrofizika **32**, 313 (1990).
26. A. G. Lyne, M. Burgay, M. Kramer, *et al.*, astro-ph/0401086.
27. A. Lightman *et al.*, *Problem Book in Relativity and Gravitation* (Princeton Univ. Press, Princeton, 1975; Mir, Moscow, 1979).
28. E. Cappellaro *et al.*, Astron. Astrophys. **322**, 431 (1997).
29. S. van den Bergh, W. Li, A. Fillipenko, *et al.*, Astrophys. J. **556**, L29 (2001).
30. S. Immler, A. Wilson, and Y. Terashima, Astrophys. J. **573**, L27 (2002).
31. A. Oliveira, J. Steiner, and D. Cieslinski, Mon. Not. R. Astron. Soc. **346**, 963 (2003).
32. P. Maxted, Mon. Not. R. Astron. Soc. **333**, 231 (2002).
33. L. Moralez-Rueda, Mon. Not. R. Astron. Soc. **338**, 752 (2003).
34. M. Stark and R. Wade, Astron. J. **126**, 1455 (2003).
35. T. Williams, J. McGraw, and R. Gravhuis, Publ. Astron. Soc. Pac. **113**, 944 (2001).
36. D. Kilkenny and S. Kenris, Observatory **120**, 48 (2000).
37. L. Morales-Rueda, P. Maxted, and T. Marsh, astro-ph/0309088.
38. L. Morales-Rueda, P. Maxted, and T. Marsh, Mon. Not. R. Astron. Soc. **338**, 752 (2003).
39. I. Iben and A. V. Tutukov, Astrophys. J., Suppl. Ser. **58**, 661 (1985).
40. A. Tutukov and L. Yungel'son, Astron. Zh. **67**, 109 (1990) [Sov. Astron. **34**, 57 (1990)].
41. Z. Han, P. Maxted, and T. Marsh, Mon. Not. R. Astron. Soc. **341**, 669 (2003).
42. H. van Horn, in *Frontiers of Stellar Evolution,* ASP Conf. Ser. **20**, 265 (1991).
43. P. Podsiadlowski, in *Advances in Stellar Evolution,* Ed. by T. Rood and A. Renzini (Cambridge University Press, Cambridge, 1998), p. 261.
44. S. G. Moiseenko, N. V. Ardeljan, and G. S. Bisnovatyi-Kogan, Rev. Mex. Astron. Astrofis. **15**, 231 (2003).
45. D. Proga, A. MacFadyen, P. Armitage, *et al.*, astro-ph/0312319.
46. G. S. Bisnovatyi-Kogan, Astron. Astrophys. **245**, 528 (1991).
47. W. Cui and B. Smith, astro-ph/0310563.
48. R. V. E. Lovelace, M. M. Romanova, and G. S. Bisnovatyi-Kogan, Astrophys. J. **514**, 368 (1999).

49. D. Lorimer, astro-ph/0308501.

50. A. Ankay, O. Guseinov, and S. Tagieva, astro-ph/0305490.

51. N. Vranesevic, R. Manchester, G. Hobbs, *et al.*, astro-ph/0310201.

52. F. Camilo, *IAU Symp. No. 218: Young Neutron Stars and Their Environment* (2003), p. 176.

53. P. Maxted *et al.*, Mon. Not. R. Astron. Soc. **326**, 1391 (2001).

54. A. Pasterello, L. Zampieri, M. Turato, *et al.*, astro-ph/0309264.

55. D. Richardson, D. Branch, D. Casebeer, *et al.*, Astron. J. **123**, 745 (2002).

56. T. Pannuti, N. Duric, C. Lacey, *et al.*, Astrophys. J. **565**, 966 (2002).

57. N. Vranesevic, R. Manchester, G. Hobbs, *et al.*, astro-ph/0310201.

58. U. Heber, Napiwotzki, and I. Reid, Astron. Astrophys. **323**, 819 (1997).

59. A. I. MacFadyen and S. E. Woosley, Astrophys. J. **524**, 262 (1999).

60. E. Rykoff, D. Smith, and P. Price, astro-ph/0310501.

61. A. Tutukov and L. Yungelson, Mon. Not. R. Astron. Soc. **260**, 675 (1993).

62. A. Tutukov and L. Yungelson, Mon. Not. R. Astron. Soc. **268**, 871 (1994).

63. S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, and A. G. Polnarev, Pis'ma Astron. Zh. **10**, 422 (1984) [Sov. Astron. Lett. **10**, 177 (1984)].

64. G. S. Bisnovatyi-Kogan, Mem. Soc. Astron. Ital. **73**, 318 (2002).

65. G. S. Bisnovatyi-Kogan, astro-ph/0310361.

66. G. S. Bisnovatyi-Kogan and S. A. Silich, Rev. Mod. Phys. **67**, 661 (1995).

67. J. M. Leblanc and J. R. Wilson, Astrophys. J. **161**, 541 (1970).

68. G. S. Bisnovatyi-Kogan, N. V. Ardelyan, and S. G. Moiseenko, Mem. Soc. Astron. Ital. **73**, 1134 (2002).

Translated by D. Gabuzda

Comparison of the Luminosity Functions of Open Clusters Based on USNO-A1 Data

P. E. Zakharova and A. V. Loktin

Astronomical Observatory, Ural State University, Leninskiĭ pr. 51, Yekaterinburg 620083, Russia
Received June 27, 2002; in final form, March 15, 2004

Abstract—The luminosity and mass functions of a group of Galactic open clusters are constructed by applying a statistical method to photometric data from the USNO-A1 catalog. Despite some limitations, this catalog can be used for statistical analyses in Galactic astronomy. Pairwise comparisons of the derived cluster luminosity functions are performed for five age intervals. The differences between the luminosity functions of the open clusters are not statistically significant in most cases. It is concluded that the luminosity functions are approximately universal throughout a large volume in the solar neighborhood. Combined luminosity and mass functions are constructed for six age intervals. The slope of the mass spectrum may vary somewhat from cluster to cluster, and the mean slope may be somewhat higher than the Salpeter value. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

The luminosity function is among the most important tools used to analyze the composition and evolution of stellar systems and is especially effective in studies of open star clusters, since the absolute magnitudes and ages of their constituent stars can be determined fairly accurately. The Galactic subsystem of open clusters is homogeneous in terms of the spatial locations, kinematics, and chemical compositions of the clusters. The large age span of these objects makes them especially convenient for analyzing effects due to the evolution of the open-cluster luminosity function. Of special importance is the possibility of converting the luminosity function into the mass function, and the easy incorporation of evolutionary effects makes it possible to analyze the corresponding initial mass functions (IMFs), at least for young clusters that are not affected appreciably by dynamical evolution.

Many authors have analyzed and compared the luminosity functions ($\varphi(M_V)$), mass functions, and IMFs of open clusters. However, it is difficult to compare the luminosity functions constructed in different studies because they often refer to different volumes of the same cluster. The luminosity functions have usually been constructed without including the cluster coronas, thereby introducing uncertainties due to mass segregation in the clusters. The inhomogeneous nature of the observational data used to determine the cluster luminosity functions is a great hindrance to comparisons of the luminosity functions of different clusters.

The USNO-A1 catalog enables for the first time the construction of the luminosity functions of a large number of open clusters to faint limiting magnitudes ($V \approx 20^m$) via the application of a single technique to fairly homogeneous data. The resulting uniform series of luminosity functions can be used to address a number of questions which it had not been possible to answer unambiguously earlier.

Our principal aim is to compare the luminosity functions of a fairly extensive sample of open clusters, in order to determine whether the luminosity functions, and, consequently, the IMFs, of the open clusters are universal in the volume of the Galaxy occupied by the sample. Zakharova [1] performed a similar analysis using the luminosity functions of 26 open clusters determined by applying a single technique to homogeneous photometric data. She concluded that open clusters have sufficiently similar luminosity functions to justify the hypothesis that there is a universal IMF.

2. SELECTION OF THE CLUSTERS

The characteristics of the USNO-A1 catalog impose certain restrictions on the selection of clusters for luminosity-function determinations. The main constraint is imposed by the brightness limit of the catalog magnitudes: the catalog contains virtually no stars brighter than $V = 10^m$. As a result, clusters with distance moduli smaller than $V_0 - M_V \approx 9.0^m$ are almost indistinguishable in images based on the USNO catalog data, preventing the conduction of star counts. For example, we were unable to construct the luminosity function for the well-known cluster

Table 1. Coefficients in the transformation formulas for the V magnitudes

Cluster	a	σ_a	b	σ_b	c	σ_c	σ_V
NGC 188	14.63^m	0.08^m	1.05	0.01	0.38	0.05	0.17^m
NGC 1245	14.80	0.10	1.10	0.04	0.40	0.07	0.21
NGC 1912	14.80	0.17	0.98	0.04	0.48	0.08	0.17
NGC 2099	14.76	0.11	0.92	0.03	0.47	0.04	0.22
NGC 2420	15.10	0.10	0.98	0.02	0.14	0.05	0.12
NGC 6866	14.28	0.15	0.94	0.04	0.39	0.06	0.20
NGC 7160	13.95	0.27	0.86	0.08	0.32	0.09	0.28

Table 2. Coefficients of the transformation formulas for the B magnitudes

Cluster	a	σ_a	b	σ_b	c	σ_c	σ_B
NGC 188	14.70^m	0.12^m	1.07	0.02	-0.16	0.07	0.25^m
NGC 1245	15.09	0.11	1.04	0.04	-0.30	0.11	0.24
NGC 1912	14.83	0.16	0.95	0.04	-0.11	0.11	0.16
NGC 2099	15.16	0.10	0.94	0.03	-0.71	0.09	0.22
NGC 2420	14.75	0.17	0.90	0.04	-0.36	0.08	0.20
NGC 6866	14.24	0.18	0.94	0.05	-0.04	0.13	0.25
NGC 7160	14.01	0.32	0.79	0.10	-0.14	0.17	0.32

NGC 752. Some clusters, primarily the youngest, are lost against the stellar background in such images, since visually bright stars with extensive halos in the initial plates result in gaps in the catalog, and the presence of a few such stars is sufficient for a cluster to be replaced by a region of low stellar density. For example, we were not able to distinguish the young cluster NGC 2251, which was studied earlier photographically by Zakharova [2].

The second restriction arises from the need to verify the correspondence between the USNO magnitudes and the B and V magnitudes of the UBV system. Our experience shows that the coefficients for the transformation from the B and V magnitudes of the USNO catalog to the standard UBV magnitudes are different for different areas in the sky. Moreover, the photometric standard—preferably a photoelectric standard—must have a limiting magnitude of no less than $V = 15^m$. Table 1 lists the coefficients for the linear transformation $V_{st} = a + b(V_{\mathrm{USNO}} - 15.0) + c(B - V)_{\mathrm{USNO}}$, their standard errors, and the dispersions σ_V for seven rich open clusters. Table 2 lists the corresponding coefficients for the transformation of the B magnitudes for the same clusters.

It is clear from the tables that the transformation coefficients deviate significantly from their mean values in some cases, although the overall behavior of the transformation remains the same for different areas of the sky. Note that the mean residual dispersion is somewhat lower than the catalog-averaged photometric error of 0.3^m given in the description of the USNO-A1 catalog.

Wide neighborhoods around clusters should be analyzed when constructing the cluster luminosity functions. This results in fairly strong fluctuations in the stellar background at faint limiting magnitudes, preventing the identification of poorly populated clusters via star counts. In practice, we can construct luminosity functions for clusters containing no fewer than 100–150 members down to the indicated limiting magnitude.

As an additional criterion, we required that the selected clusters be listed in the latest version of the homogeneous catalog of open-cluster parameters [3], which now contains data for 425 clusters. This criterion ensures homogeneity in the scales of the cluster parameters—the color excesses, ages, and distance moduli—which is necessary if we wish to derive a combined luminosity function for clusters located at different distances from the Sun and with different interstellar extinctions.

Our initial sample contained 79 clusters whose parameters meet the above criteria; however, we kept only 63 of these for our further analysis. We excluded clusters with estimated populations of fewer than 100 members after the construction of the luminosity functions.

3. STAR COUNTS AND THE CONSTRUCTION OF LUMINOSITY FUNCTIONS

We began our analysis of the selected clusters by checking the agreement between the USNO magnitudes and the B and V magnitudes of the UBV photometric system. We chose photometric standards primarily from the Lausanne Institute of Astronomy WEBDA database, created under the supervision of Mermilliod (http://obswww.unige.ch/webda). For each cluster, we used standard stars to construct linear relations for the transformation of the USNO magnitudes to the B and V magnitudes of the UBV system; examples of the transformation coefficients can be found in Tables 1 and 2. The dispersion of the USNO magnitudes varied from 0.18^m to 0.36^m, depending on the area of the sky. These estimates indicate that (1) the photometric accuracy of the USNO-A1 catalog is sufficiently high that
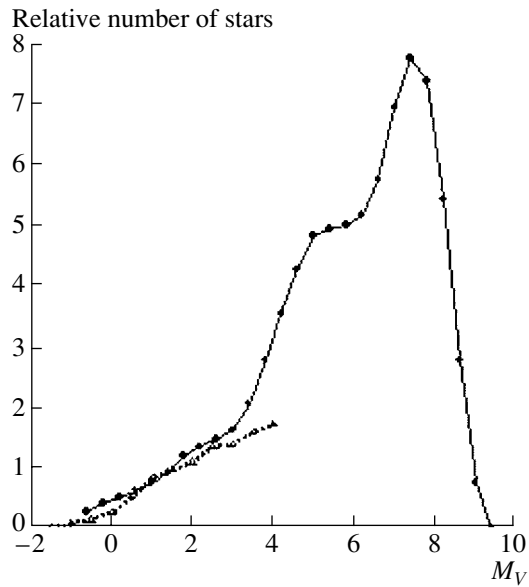
Relative number of stars

Relative number of stars

Fig. 2. Comparison of the luminosity functions of the open cluster NGC 2420 obtained in this paper (solid) and by Leonard [7] (dashed).

Fig. 1. Comparison of the luminosity functions of the open cluster NGC 1912 obtained in this paper (solid) and by Barkhatova and Zakharova [6] (dashed).

we can analyze the magnitude distributions using histograms with bins wider than 0.4^m and (2) the photometric accuracy of the catalog is nonuniform in terms of the random errors. However, this latter defect is of little importance for analyses of luminosity functions if the widths of the magnitude bins exceed the photometric errors.

We constructed the luminosity functions of the open clusters using the usual technique for this purpose employed at the Astronomical Observatory of Ural State University. We conducted star counts in the field of each cluster to determine the coordinates of the center and radius of the cluster. We determined the luminosity function by counting stars in magnitude bins (with widths of 0.4^m in our case) in the area bounded by the cluster radius and in a comparison area in the form of a ring surrounding the cluster. We performed the latter counts to statistically eliminate nonmembers. We chose the outer (background, or field-star) ring so that its area would be close to that of the cluster and its inner boundary would somewhat exceed the cluster radius, to allow for possible uncertainty in the estimated radius.

Figures 1 and 2 compare the luminosity functions of two clusters derived in this paper with the results obtained earlier by other authors. Figure 1 compares our luminosity function for NGC 1912 with the luminosity function constructed by Barkhatova and Zakharova [6] (shown by the dashed curve). For convenience, in this and other figures, we draw the luminosity functions as polygonal functions. We normalized all the luminosity functions to have the same
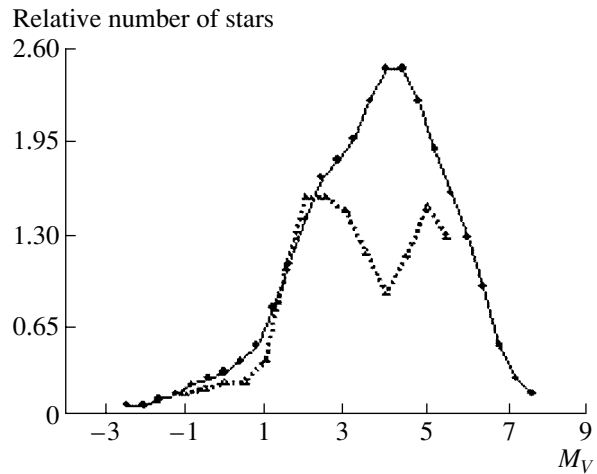
number of stars in the absolute-magnitude interval $0^m < M_V < 2^m$. This normalization was necessary because the luminosity functions cannot contain the same number of stars in all intervals, e.g., due to random errors in the estimated cluster radius or a different choice of comparison area. The vertical axis plots the number of stars as a fraction of the number of stars in the above normalization interval. It is clear from the figure that the two luminosity functions are in excellent agreement throughout nearly the entire magnitude interval studied in [6]. Note that both the luminosity functions in Fig. 1 were obtained using the same technique. It is also obvious that the USNO catalog yields a much wider range of magnitudes.

Figure 2 compares our luminosity function for NGC 2420 with the luminosity function obtained by Leonard [7]. The luminosity functions are normalized in the interval $-1.0^m < M_V < +1.5^m$. The two luminosity functions agree less well than in the previous case. However, this discrepancy can be explained by the fact that Leonard [7] analyzed the region of the cluster core, so that the deficit of faint stars in his luminosity function can be accounted for by the lack of corona stars and the loss of faint stars from the photometric measurements of the central area the cluster, which is crowded with bright stars. In addition, the choice of a different comparison area to account for field stars can result in substantial random errors.

It follows from the above comparisons of the two luminosity functions that the USNO catalog data can be used to analyze luminosity functions of rich open clusters that are not too nearby.

Table 3. Comparison of luminosity functions of clusters with $6.8 < \log t < 7.4^*$

No.	Cluster	2	3	4	5	6	7	8	9
1	IC1848	−	−	−	+	−	−	+	+
2	NGC 884		+	+	+	+	+	+	+
3	NGC 869			+	+	+	+	+	−
4	NGC 654				+	+	+	+	−
5	NGC 663					+	+	+	+
6	NGC 7128						+	+	+
7	NGC 2439							+	+
8	NGC 2489								−
9	NGC 2421								

* Here and in other table captions, the age t is in years.

Table 4. Comparison of luminosity functions of clusters with $7.4 < \log t < 8.0$

No.	Cluster	2	3	4	5	6	7	8	9	10	11	12	13
1	NGC 1960	+	+	+	−	−	+	+	+	+	+	+	+
2	NGC 659		+	+	−	+	+	+	+	+	−	+	+
3	NGC 6649			+	−	+	+	+	+	+	+	+	+
4	NGC 7510				−	+	+	+	+	+	+	+	+
5	NGC 6755					+	−	−	−	−	−	−	−
6	NGC 7790						+	+	+	−	+	+	+
7	NGC 7654							+	+	+	+	+	+
8	NGC 146								+	+	+	+	+
9	NGC 2345									+	+	+	+
10	King 14										−	+	−
11	NGC 436											−	+
12	NGC 2353												−
13	NGC 2168												

4. PAIRWISE COMPARISON OF LUMINOSITY FUNCTIONS

One of the important questions in studies of the luminosity function of clusters is whether there is a universal luminosity function for clusters throughout observed Galaxy. The answer depends on how different the star-formation parameters are in different parts of the Galactic disk. To resolve this question, we must carry out a statistical comparison of different luminosity functions. If the luminosity functions of young clusters do not differ in a statistically significant way, we can conclude that star formation is governed by the same laws in clusters born in different regions of the Galaxy out to distances from the Sun of about 3.5 kpc, i.e., in the volume occupied by the sample clusters. We should bear in mind that the effects of dynamical evolution and differences in chemical composition can be important for old clusters. The luminosity functions of clusters of different ages differ not only due to possible differences in their initial luminosity functions but first and foremost due to the evolution of their stars. We accordingly subdivided the sample clusters into five age groups and compared the luminosity functions within these groups (age t is given in years throughout this paper). We excluded all clusters with fewer than 150 members due to the relatively low accuracy of their sample distribution functions (i.e., their luminosity functions). The final sample contained 60 clusters.

We performed a pairwise comparison of the luminosity function using the criteria of Pearson and Kolmogorov at the 5% significance level. A description of the procedure used to apply these criteria can be found in [1]. The use of two statistical criteria enables a more trustworthy determination of the statistical significance of differences between the two distributions. Note that Pearson's χ^2 criterion is somewhat more reliable due to the specific characteristics of our data. The reason is that the Kolmogorov criterion was developed for the comparison of distribution functions and not probability densities. In our case, the distribution function has the form of the integrated luminosity function; however, despite the large photometric depth of the USNO catalog, the USNO data cover a limited interval of magnitudes, and the resulting luminosity function (the density of the distribution of apparent magnitudes) is also defined in a limited magnitude interval. This is why we adopted the Pearson test as our main criterion and the Kolmogorov test as a secondary check.

Tables 3−7 summarize the results of the pairwise comparison of the luminosity functions of the sample clusters divided into five age intervals. The first column in each of these tables, which all have the same layout, gives the running number of the cluster. The header row also shows the number of the cluster. The intersection of the row and column gives the result of the luminosity-function comparison for the clusters with the corresponding running numbers. Here "+" and "−" denote differences between the corresponding luminosity functions that are insignificant and significant, respectively, at the 5% significance level. The two statistical criteria used yielded virtually identical results, and we present only the results based on Pearson's criterion. The Kolmogorov criterion is

Table 5. Comparison of luminosity functions of clusters with $8.0 < \log t < 8.5$

No.	Cluster	2	3	4	5	6	7	8	9	10	11	12	13
1	NGC 5168	+	−	−	−	−	−	+	−	+	−	+	+
2	NGC 2323		+	−	−	+	+	+	+	+	−	+	+
3	NGC 6192			−	−	+	−	+	+	−	−	+	−
4	NGC 103				−	+	−	−	−	−	−	−	−
5	NGC 7086					−	−	−	−	−	+	−	+
6	NGC 6709						+	+	+	−	−	+	+
7	NGC 1027							−	−	−	−	−	−
8	NGC 1039								+	+	+	+	+
9	Mel 71									+	−	+	+
10	NGC 1912										−	+	+
11	NGC 1664											−	+
12	NGC 7062												+
13	NGC 6494												

Table 6. Comparison of luminosity functions of clusters with $8.5 < \log t < 9.0$

No.	Cluster	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	IC 2714	+	−	+	+	+	+	+	+	+	+	+	+	+	−	+
2	NGC 2236		+	+	−	+	+	+	+	+	+	+	+	+	−	+
3	NGC 2099			−	−	−	+	−	−	+	−	−	−	+	−	+
4	NGC 1528				+	−	−	−	−	−	−	−	+	−	−	−
5	NGC 1907					+	+	+	+	+	+	+	+	−	−	+
6	NGC 6866						−	+	+	−	+	−	+	−	−	−
7	NGC 2818							+	+	+	+	+	+	+	−	+
8	NGC 2324								+	+	+	+	+	+	−	+
9	IC 1369									−	+	−	+	−	−	+
10	NGC 1245										+	+	+	+	−	+
11	NGC 6811											−	+	+	−	+
12	NGC 2477												+	+	−	+
13	NGC 6802													−	−	+
14	NGC 2533														−	+
15	NGC 2204															−
16	NGC 6134															

slightly "harder" and yields a somewhat higher percentage of significant differences.

The luminosity function of only one cluster in Table 3, IC 1848, deviates significantly from the other luminosity functions. The good agreement between the luminosity functions of young clusters is somewhat surprising, since the clusters in this subsample span a wide range of ages and the luminosity functions of young clusters are difficult to determine from the USNO catalog data for the reasons mentioned above.

In Table 4 as well, the luminosity function of only one cluster, NGC 6755, deviates from those of the other clusters. The situation is reversed in Table 5: of the 78 comparisons, 40 have minus signs (the Kolmogorov criterion yielded 45 minus signs in this case), and it is difficult to identify a single cluster that shows the worst agreement.

Table 6 shows a much better agreement between the luminosity functions of different clusters than in the previous case, with only one cluster, NGC 2204, having a luminosity function that differs significantly from those of the other clusters of this subsample.

Only one cluster in Table 7, NGC 2506, mars the picture of almost complete agreement between the luminosity functions of the clusters of this subsample. Note that the effects of dynamical evolution and differences in chemical composition should show up most strongly in this group, so that the good agreement between the luminosity functions is somewhat surprising.

Table 7. Comparison of luminosity functions of clusters with $9.0 < \log t < 9.5$

No.	Cluster	2	3	4	5	6	7	8	9
1	NGC 2243	+	−	−	+	−	−	−	+
2	NGC 2420		+	+	+	+	+	+	+
3	NGC 2506			−	−	−	−	−	−
4	NGC 2658				+	+	+	+	+
5	NGC 6819					+	−	+	+
6	NGC 2243						−	+	+
7	NGC 7789							+	+
8	NGC 7142								+
9	NGC 6939								

Of the 348 pairs compared in all the tables, 120 (34%) showed statistically significant differences. If we discard the few "bad" clusters indicated above (IC 1848, NGC 6755, NGC 6709, NGC 7086, NGC 1528, NGC 2204, NGC 2506), only 59 of the remaining 268 pairs (22%) show statistically significant differences in their luminosity functions.
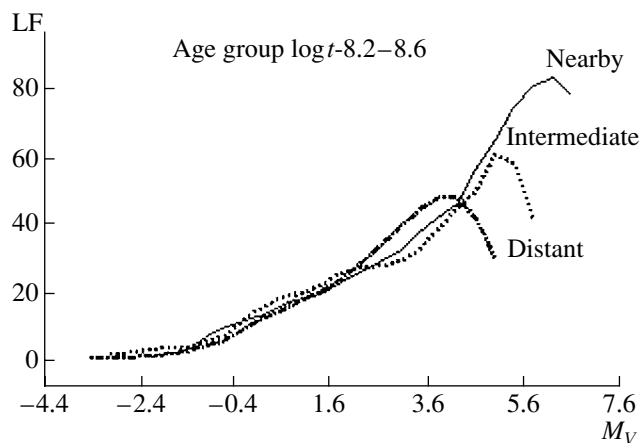
Fig. 3. Examples of combined luminosity functions of clusters in the same age interval subdivided into three groups according to distance from the Sun.
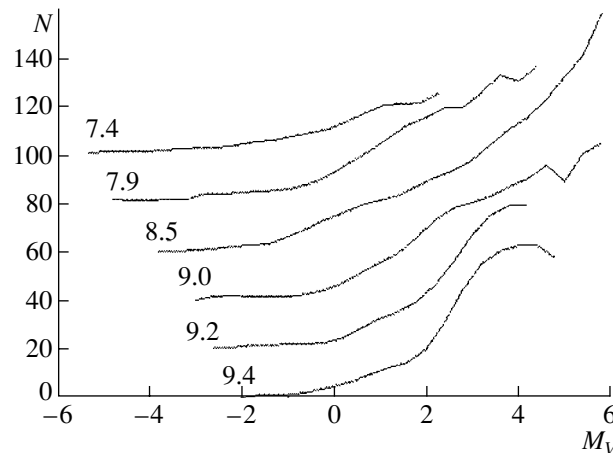

Fig. 4. Combined luminosity functions for the sample clusters for six age intervals (the age t is in years).

Given the considerable widths of the age bins and various factors hindering the construction of cluster luminosity functions, we conclude that the luminosity functions of most of the clusters do not differ significantly from each other. The star counts cover fairly large areas of sky, even for the distant open clusters, and the fluctuations of the density of Galactic field stars can be quite significant in some cases. Note that the fraction of coinciding luminosity functions may be somewhat underestimated due to the effect of random errors in the distance moduli and color excesses of the open clusters, which we adopted from the current version (2.1) of the homogeneous catalog of open-cluster parameters [3]. In some cases, the agreement of the luminosity functions may also be marred by differences between the mean heavy-element abundances in the atmospheres of the stars. However, even for the most metal-poor clusters, the corrections to the absolute magnitudes do not exceed the width of the binning interval used to construct the luminosity functions [3]. In any case, the results of the luminosity-function comparisons presented in the tables can be taken to represent a lower limit for the degree of agreement between the luminosity functions of these open clusters. The fact that, on the whole, the luminosity functions of the open clusters agree well with each other, despite the appreciable random errors in the luminosity functions derived for both the youngest and oldest open clusters, supports the possibility that the initial luminosity function is universal in the volume of the Galaxy considered. The significant deviations of the luminosity functions of a few rich clusters could be due to the incompleteness of the photometric measurements of faint stars in crowded fields near the cluster cores. This is a

characteristic shortcoming of all photometric studies, especially those based on photographic photometry.

Given the large number of factors contributing to the errors in the constructed luminosity functions, we conclude that the lack of statistically significant differences between the luminosity functions of most of the clusters suggests that the conditions for star formation in clusters of all the observed ages are uniform throughout the volume of the Galaxy considered, independent of the spatial location of the clusters.

5. COMBINED LUMINOSITY FUNCTIONS FOR AGE INTERVALS

Experience shows that statistically useful luminosity functions can be obtained only for a relatively small number of rich open clusters that are surrounded by uniform star fields. The errors in the constructed luminosity functions increase both with the area of the sky occupied by the cluster (the dispersion of the stellar density is proportional to the area considered [4]) and the limiting magnitude (the nonuniformity of the stellar field increases with the limiting magnitude). In addition, the fraction of stars whose images are suitable for photometry decreases if the stellar density is high. There is, however, another way to tackle this problem: analyzing several clusters and combining their luminosity functions, neglecting the possible differences between the luminosity functions of different clusters. This last assumption is reasonable in view of our above conclusion that the luminosity functions of open clusters have an approximately universal form. In this case, the effect of the random errors decreases, as is usual when averaging over samples, although this is achieved at the expense of losing information about the behavior of the luminosity functions of individual clusters. However,
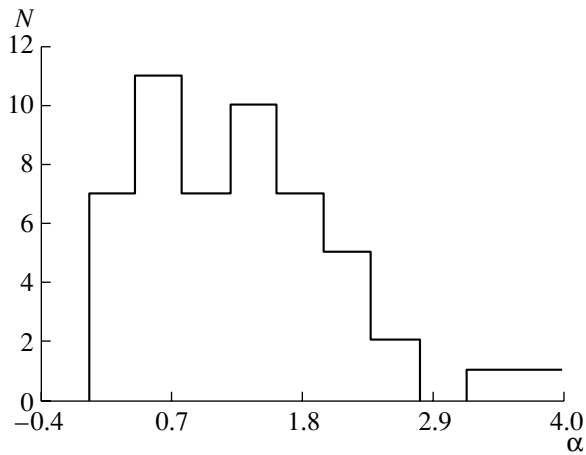
Fig. 5. Distribution of the slopes α of the cluster mass functions derived in this paper.

Fig. 6. Distribution of the estimated slopes of the mass functions compiled by Tarrab [8].

this approach makes it possible to explore the general properties of the luminosity functions with greater certainty.

To construct the combined luminosity function, we subdivided our sample into six age intervals, excluding a few youngest and oldest clusters because of the small number of such objects in our sample. We subdivided the clusters within each age interval into two or three groups according to their distance moduli. This makes it possible to extend the interval of absolute magnitudes covered by the averaged luminosity function: nearby clusters lack the bright end of the luminosity function due to the bright magnitude limit of the initial catalog, but such clusters show the faint end of the luminosity function. The reverse is true for distant clusters. To illustrate this point, Fig. 3 shows the combined luminosity functions for the same age interval but for three different intervals of distance from the Sun. The difference of the limiting magnitudes and the small scatter of the luminosity functions over most of the interval considered are clearly visible.

Figure 4 shows the appropriately normalized combined luminosity functions for six age groups over the common intervals of absolute magnitudes. We can see the intervals of absolute magnitudes in which it is possible to construct the luminosity functions based on the USNO catalog data. The combined luminosity functions in Fig. 4 are shifted along the N axis by 20 units for ease of viewing. Note that the slope of the luminosity functions increases somewhat in the transition from young clusters to older open clusters.

6. MASS FUNCTIONS

We calculated the mass functions from the luminosity functions of the individual clusters by translating the M_V values into stellar masses using the
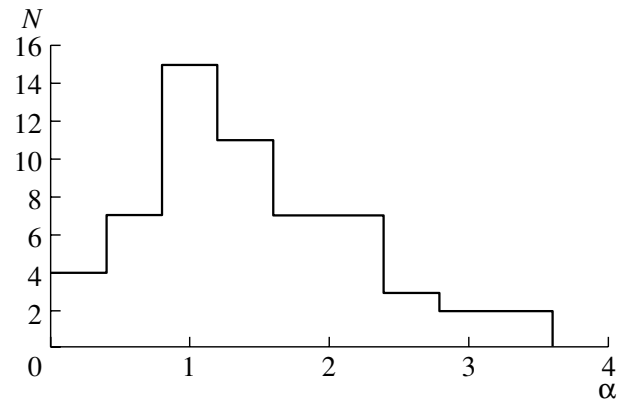
theoretical isochrones of Bertelli *et al.* [5], which are given in M_V—mass coordinates. This is a convenient way to take into account evolutionary effects, which is important when constructing the mass functions of open clusters. We computed the mass function for each cluster of the sample using the usual procedure for the transformation of the distribution functions: transforming the density function into the distribution function, replacing the magnitude scale with the mass scale, and numerically differentiating with respect to the new variable to carry out the inverse transformation to the probability density—the mass function. We retained clusters whose luminosity functions contained a linear segment for further analysis. For each of the 51 clusters selected in this way, we estimated the slope of its mass spectrum α from the relation $dN/d\log m \propto m^{-\alpha}$. Figure 5 shows the distribution of the estimated slopes α for the sample clusters. For comparison, Fig. 6 shows the distribution of the slopes of the open-cluster mass functions obtained by Tarrab [8]. It is evident that our distribution is somewhat more concentrated toward a specific interval in which the estimated α are located. The distribution of α values for our clusters has an almost rectangular shape, implying that the actual slopes for various open clusters are not exactly equal and are distributed over some interval. Indeed, if the mass function of all the clusters had the same slope, the narrow distribution of α values convolved with the error distribution would produce a bell-shaped distribution function. Unfortunately, the small size of the sample and difficulties with estimating the real errors of the mass-function slopes prevented us from correcting the distribution for the effect of random errors. The mean slope of the cluster mass functions calulated from the distribution in Fig. 5 is 1.27 ± 0.12.

Table 8 lists the individual estimates of α for the sample clusters. The standard errors of the α estimates given in this table are the estimated errors of

Table 8. Comparison of luminosity functions of clusters with $9.0 < \log t < 9.5$

Cluster	α	$\log t$	Cluster	α	$\log t$
NGC 146	0.12 ± 0.12	7.82	NGC 2439	1.49 ± 0.28	7.25
NGC 381	3.72 ± 0.61	8.57	NGC 2658	1.64 ± 0.30	9.15
NGC 436	0.64 ± 0.16	7.93	NGC 2818	0.19 ± 0.14	8.63
NGC 457	1.29 ± 0.32	7.32	NGC 6134	0.22 ± 0.20	8.97
NGC 654	0.63 ± 0.12	7.15	NGC 6192	1.49 ± 0.14	8.13
NGC 659	1.36 ± 0.19	7.55	NGC 6494	3.26 ± 0.70	8.48
NGC 663	0.91 ± 0.04	7.21	NGC 6649	0.92 ± 0.13	7.57
NGC 744	1.82 ± 0.97	8.25	NGC 6709	2.28 ± 0.22	8.18
NGC 869	0.39 ± 0.03	7.07	NGC 6755	2.08 ± 0.14	7.72
NGC 884	0.16 ± 0.06	7.03	NGC 6802	2.28 ± 0.07	8.87
NGC 1027	2.15 ± 0.15	8.20	NGC 6811	0.78 ± 0.44	8.80
NGC 1039	0.86 ± 0.24	8.25	NGC 6834	1.60 ± 0.40	7.88
NGC 1245	2.29 ± 0.67	8.70	NGC 6866	1.56 ± 0.16	8.58
NGC 1528	2.49 ± 0.12	8.57	NGC 7062	0.51 ± 0.14	8.47
NGC 1778	0.69 ± 0.28	8.16	NGC 7086	1.77 ± 0.27	8.14
NGC 1907	1.50 ± 0.22	8.57	NGC 7128	1.68 ± 0.38	7.25
NGC 1912	0.47 ± 0.11	8.46	NGC 7235	1.20 ± 0.68	7.07
NGC 1960	0.20 ± 0.13	7.47	NGC 7510	0.64 ± 0.20	7.58
NGC 2168	0.81 ± 0.13	7.98	NGC 7654	0.95 ± 0.21	7.76
NGC 2236	1.30 ± 0.42	8.54	NGC 7790	2.63 ± 0.18	7.75
NGC 2323	0.33 ± 0.18	8.10	IC 1369	1.42 ± 0.17	8.64
NGC 2324	1.00 ± 0.09	8.63	IC 1848	0.99 ± 0.16	6.84
NGC 2345	0.49 ± 0.13	7.85	IC 2714	1.55 ± 0.24	8.54
NGC 2353	1.67 ± 0.13	7.97	King 14	0.57 ± 0.10	7.92
NGC 2420	0.46 ± 0.11	9.05	Mel 71	0.77 ± 0.08	8.37
NGC 2421	1.89 ± 0.10	7.37			

the regression coefficients based on a nearly linear segment of the mass function. Therefore, the errors listed in Table 8 are lower limits for the real errors, which are difficult to estimate.

We constructed the combined mass functions for six age intervals, shown in Fig. 7, as we did for the luminosity functions. The mass interval in which we can use the USNO catalog data to analyze the cluster mass functions is visible. It is immediately apparent from this figure that the slopes of the mass functions are virtually independent of the mass and age inside this interval. The mean slope of the mass function derived from the combined mass functions is $\alpha = 1.50 \pm 0.05$, which somewhat exceeds the classical Salpeter value of $\alpha = 1.35$. The slight discrepancy between the distribution-averaged slope given above and the
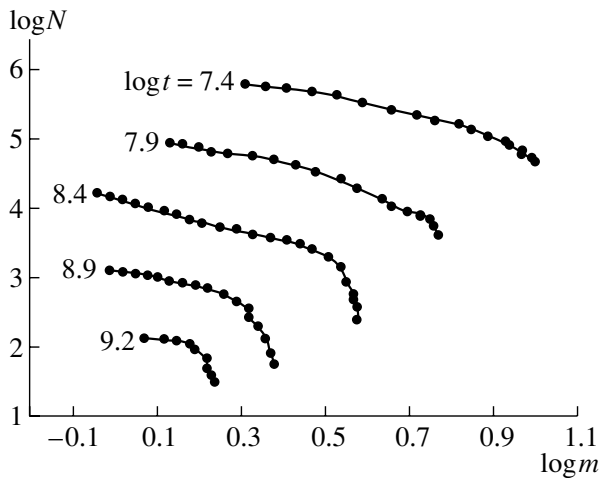
Fig. 7. Combined mass functions of the clusters in our sample in five age intervals (shown is the logarithm of the age in years).

mean slope of the combined cluster luminosity functions can be explained by the different procedures used to identify the linear segments of the mass functions: the linear segments for individual clusters are chosen in the presence of fluctuations due to random errors. Since linear segments in the combined mass functions are easy to identify, we consider the mean slope of the open-cluster mass spectrum derived from the combined mass functions to be more trustworthy. Our estimates of the mean slopes of the cluster mass functions indicate that this mean slope does not differ appreciably from the Salpeter value.

We used the slopes of the mass functions derived for various age groups to determine whether this parameter depends on the cluster age by estimating the coefficients of the linear regression model $\alpha = a + b \log t$. The coefficient $b = -0.28 \pm 0.19$ is formally not significantly different from zero, but the indicated possible weak age dependence of the slope could be due to the well-known fact that the IMF is steeper in the region of massive stars.

7. CONCLUSIONS

Our results show that the USNO-1A catalog can be used to determine the luminosity and mass functions of open clusters, provided that the clusters are carefully selected in accordance with the criteria discussed above and that photoelectric standards are used to reduce the catalog magnitudes to the standard UBV system. The number of such clusters will increase as a result of further analyses.

Our technique for the determination of the luminosity functions is essentially unaffected by mass segregation—the tendency of massive members to concentrate toward the cluster center—since the counts include stars both in the core region and in the cluster coronas. The only factor to somewhat affect the resulting luminosity functions is the impossibility of performing photometric measurements for faint stars in the dense cores of rich clusters, which is always the case for photographic photometry. Note that experience shows that the main factors resulting in the fairly large errors of statistically derived luminosity functions are (1) the nonuniform distribution of the density of field stars and (2) errors in the calibration of the magnitudes in the initial catalog. Further increase of the accuracy of the luminosity functions of open clusters can be achieved only via direct identification of cluster members using kinematical criteria.

In most cases, the luminosity functions of the open clusters do not differ significantly, possibly indicating that the process of star formation has the same properties throughout the observed part of the Galaxy.

The slope of the open-cluster mass spectra may vary from cluster to cluster, and the mean slope may somewhat exceed the Salpeter value. The slope of the open-cluster mass functions at logarithmic masses from 0.0 to 1.0 is independent of both mass and age.

ACKNOWLEDGMENTS

REFERENCES

1. P. E. Zakharova, Astron. Zh. **58**, 1170 (1981) [Sov. Astron. **25**, 665 (1981)].
2. P. E. Zakharova, *Astrogeodesic Studies* (Sverdlovsk, 1986), p. 134 [in Russian].
3. A. V. Loktin, T. P. Gerasimenko, and L. K. Malysheva, Astron. Astrophys. Trans. **20**, 607 (2001).
4. A. E. Vasilevskii, *Methods of Stellar Statistics* (Ural Gos. Univ., Sverdlovsk, 1985) [in Russian].
5. G. Bertelli, A. Bressan, *et al.*, Astron. Astrophys., Suppl. Ser. **106**, 275 (1994).
6. K. A. Barkhatova and P. E. Zakharova, *Star Clusters* (Sverdlovsk, 1979), p. 90.
7. P. J. T. Leonard, Astron. J. **95**, 108 (1988).
8. I. Tarrab, Astron. Astrophys. **109**, 285 (1982).

Translated by A. Dambis

The Symbiotic System Z Andromedae during the Flare of 2000−2002

O. G. Taranova[1], N. A. Tomov[2], M. T. Tomova[2], and V. I. Shenavrin[1]

[1]*Sternberg Astronomical Institute, Universitetskiĭ pr. 13, Moscow, 119992 Russia*

[2]*Institute of Astronomy, Rozhen National Astronomical Observatory, P.O. Box 136, 4700 Smolyan, Bulgaria*

Received June 6, 2003; in final form, March 15, 2004

Abstract—Optical and infrared photometric observations of the classical symbiotic binary Z And obtained in quiescent and active phases in 1999−2002 are analyzed. *RI* photometric data from the literature are used to estimate the parameters of the binary system's components. The data as a whole suggest that the hot compact component expands and the emission measure of the surrounding nebula grows during the flare. Possible interpretations of the flare based on several mechanisms are considered. © *2004 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

Symbiotic stars represent a group of interacting binary systems whose photometric behavior is characterized by alternating quiescent periods and phases of eruptive activity. In some cases, the activity is accompanied by mass loss from the accreting star in the form of isolated shells, and in other cases, with high-velocity gas outflows (stellar winds). The energy distribution in the ejected shells changes; as a rule, this results in a strong increase in the optical brightness. The Z And system is considered to be a prototype of the class of symbiotic stars; its components are a normal, cool M4.5 giant [1]; a hot, compact object with a temperature exceeding 10^5 K [2]; and an expanding nebula formed by the wind of the giant and partly ionized by the hot component. The orbital period of this binary, 758 days, has been obtained using photometric [3] and radial-velocity [4] data.

In the quiescent state, variations with an amplitude of $\sim 0.5^m$ in B and V and $\sim 1^m$ in U are visible in the orbital light curve. In the quiescent phase and during a small flare of 1997, 28-min pulsations with an amplitude of $0.002^m - 0.005^m$ were detected in the B band; it is thought that they arise from the rotation of an accreting magnetized white dwarf [5].

In active states (1915, 1939, 1960, 1984, and 2000), increases of the optical brightness of Z And have reached $2^m - 3^m$. During the large flare of 1939, the visual brightness of the star was nearly 8^m; only P Cygni lines were observed in the spectrum, and the emission lines corresponded to those of an A star. The spectra of the M giant and the surrounding nebula were not visible [6]. The optical spectrum of Z And after the activity of 1960 was observed by Boyarchuk [7]; these data testify to a redistribution of the energy

in the continuum due to the expansion and cooling of the hot component as its brightness rose. In their analysis of radio spectra and low- and high-resolution UV spectra of Z And obtained during its active phase in 1984−1985, Fernández-Castro *et al.* [8] found that the increase in the visual brightness is closely associated with a minimum in the UV and radio. They explained this anticorrelation between the visual and UV fluxes as being due to damping of the hot stellar radiation in a false photosphere formed during the ejection of a dense envelope.

The last active phase of Z And began at the end of August 2000 [9]; the star reached its optical maximum in December, when its V brightness was 8.9^m. After some time at brightness maximum, the 28-min pulsations vanished, possibly because the white dwarf's flux was blocked by the optically thick envelope [10]. The radiation of the hot star dominated over the radiation of the giant and surrounding nebula.

In this paper, we estimate the main parameters of the components of the Z And system during its optical flare of 2000−2002, based on optical and infrared photometry of the star just before, during, and in the decline of the optical flare. In the analysis, we also used spectral data on the symbiotic system Z And in its quiescent and active phases.

2. OBSERVATIONAL DATA

Infrared photometry of Z And was carried out in the fall of 1999 (JD 2451453.498), fall and winter of 2000 (JD 2451869.337 and JD 2451903.241), and winter of 2002 (JD 2452310.197) using an InSb-photometer mounted on the 125-cm telescope of the Crimean Laboratory of the Sternberg Astronomical

Table 1. Multicolor photometry of Z And

Date	JD 2451000+	Phase	U	B	V	R[a]	I[a]	J	H	K	L	M
September 15, 1999, quiescent state	437.383	0.560	10.81 ± 0.03	11.53 ± 0.02	10.43 ± 0.03	8.80 ± 0.02	7.20 ± 0.01	6.23 ± 0.02	5.28 ± 0.03	4.97 ± 0.01	4.76 ± 0.03	4.99 ± 0.08
November 22, 2000, active phase	871.313	0.131	8.60 ± 0.01	9.41 ± 0.01	8.94 ± 0.01			6.05[b]	5.22[b]	4.90[b]	4.64 ± 0.01	4.91 ± 0.02
December 6, 2000, active phase	885.229	0.150	8.40 ± 0.01	9.30 ± 0.01	8.86 ± 0.01			6.02[b]	5.20[b]	4.86[b]	4.61 ± 0.01	4.93 ± 0.02
January 27, 2002, active phase	1302.224	0.699	10.01 ± 0.03	10.93 ± 0.02	10.17 ± 0.03			6.21 ± 0.02	5.31 ± 0.01	4.98[b]	4.76 ± 0.01	5.03 ± 0.03

[a] Data of Belyakina [18].

[b] The internal error is $< 0^m 01$.

Institute. The star BS8860 ($J = 1.46^m$, $H = 0.62^m$, $K = 0.37^m$, $L = 0.25^m$, $M = 0.67^m$ [11]) served as the photometric standard; its H magnitude was estimated from its spectral type, based on the data of Koorneef [12].

UBV photometry of Z And was carried out in the fall of 1999 (JD 2451437.383), winter of 2000 (JD 2451885.229), and winter of 2002 (JD 2452302.224) using a single-channel photoelectric photometer at the Cassegrain focus of the 60-cm ($F/12.5$) telescope of the Rozhen National Astronomical Observatory (Bulgaria). BD+48°4090 ($V = 9.01^m$, $B - V = 0.48^m$, $U - B = -0.06^m$ [13]) was used as a comparison star.

The optical and infrared observations were obtained on nearby dates. Given the insignificant variability of the infrared brightness of the system, even during the optical flare, we combined the obtained infrared photometric data and assigned them to the date of the UBV observations. Table 1 lists the results of our photometric observations.

Due to unfavorable atmospheric conditions, our UBV observations in November 2000 had large errors; however, Skopal *et al.* [14] observed the star with a higher accuracy at the same time (November 22), and their UBV magnitudes are also listed in Table 1. Our UBV data for December agree very well with the photometry of Skopal *et al.* [14], and we present the average of these two sets of magnitudes

in Table 1. Thus, Table 1 collects $UBVJHKLM$ photometric data obtained in the quiescent state of 1999 and during the flare.

Figure 1a shows the V brightness variations, and Fig. 1b, variations of the JL brightnesses and $J - K$ color index for 1999–2002.

The availability of nearly simultaneous optical and infrared photometric observations of Z And has enabled us to analyze the brightness of the binary system over a broad range from 0.36 to 5 μm during various activity phases of the hot component of the system.

3. ANALYSIS OF THE SPECTRAL ENERGY DISTRIBUTION

To analyze the spectral energy distribution (SED) of Z And at 0.36–5 μm, we converted the observed magnitudes to fluxes using the calibration of Johnson [15]. The observed U fluxes (0.36 μm) were corrected taking into account the SED of Z And near the Balmer jump, which is almost centered on the U band. In addition, strong emission lines are observed in the optical spectrum of Z And, and their contribution was taken into account in the observed UBV fluxes. The correction for interstellar extinction was performed using the value $E(B - V) = 0.30$. The technique and results of the correction of the UBV photometric data for Z And in its quiescent and active phases are described in detail in [16].
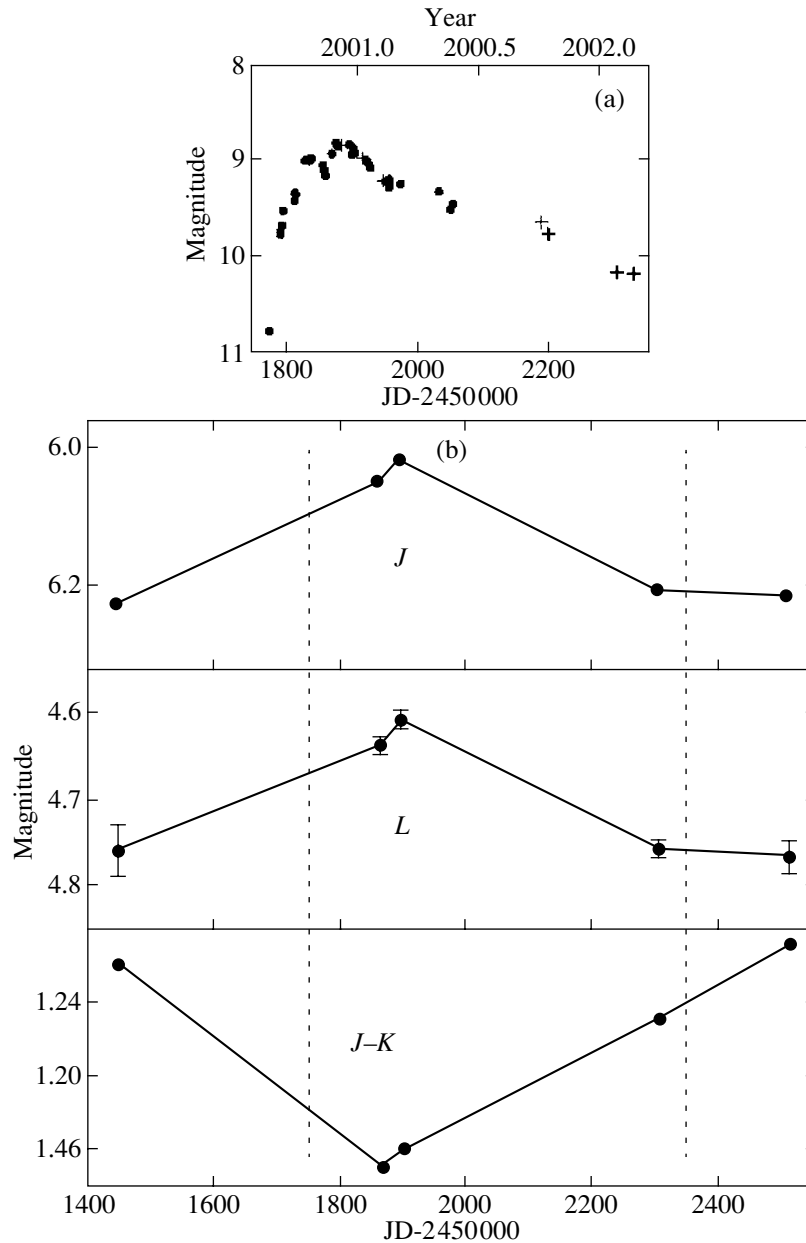
Fig. 1. (a) Variations of the V brightness of Z And during the flare of 2000−2002. Shown are the data of Skopal *et al.* [14] (circles) and our own previously unpublished data (crosses); the epochs at which the *JHKLM* photometry of the binary system was carried out are shown by the arrows. (b) Variations of the *JL* brightnesses and $J − K$ color index in 1999−2002.

The ephemeris of the orbital motion was estimated using the elements of [3, 4, 17], Min = JD 2442666 + $758.8^d E$.

The main purpose of our present analysis is to estimate the parameters of the hot star and the surrounding nebula during the optical flare of Z And in 2000−2002. These parameters can be determined from the radiation of the system's components. The radiations of each component, e.g., in the quiescent state, were estimated as follows:

(1) We derived the SED of the cool component using the infrared photometry in the quiescent state, when the cool star is the sole source of radiation at 1.25−5 μm. We assumed that the radiation of the cool star did not change during the optical flares.

(2) We determined the SED of the hot component using its UV radiation and temperature estimates, assuming that the hot star radiates as a blackbody.

(3) We then obtained the SED of the nebula by taking the difference of the observed fluxes and the total flux from the stellar components.

The SEDs Z And in its quiescent state and the active phase of 2000−2002 are presented in Fig. 2.

3.1. Quiescent State

To analyze the SED in the quiescent state and estimate the parameters of the cool component, we supplemented our data with R (0.7 μm) and I (0.8 μm) observations of Z And [18] obtained near the brightness maximum of JD 2447598. The effective temperature of the cool giant was determined based on its spectral class [1] and the calibration for normal giants of van Belle *et al.* [19], which indicates that the effective temperature of an M(4.5 ± 0.2) star is 3400 ± 50 K. However, the calibration itself has a standard error of ±270 K. Therefore, we took the effective temperature of the cool component to be 3400 ± 270 K.

The SED of the red giant was determined from its K magnitude (where only the giant contributes) correcting for the interstellar extinction, assuming that the giant's color index was $V - K = 5.15^m$. This value is given by Fluks *et al.* [20] for the star HD5316 (M4III), which, according to van Belle *et al.* [19], has an effective temperature of 3400 K. The U, B, R, and I fluxes of the giant were derived from its V flux and the flux ratios for HD5316, taken from the SIMBAD database and Kharitonov *et al.* [21]. The observed integrated flux of the giant was estimated from its K flux (Table 2) based on the data of [22]: $(2.25 \pm 0.05) \times 10^{-8}$ erg cm^{-2} s^{-1}, where the uncertainty corresponds to the observational errors. The radius of the giant was estimated from the relationship $R_{\text{cool}} = r(F_{\text{int}}/\sigma T^4)^{0.5}$, where $r = 1.12$ kpc is the distance to the system [2, 8] and $R_{\text{cool}} = (85 \pm 14)R_{\odot}$.

We determined the UBV fluxes of the hot star in the quiescent state from the observed flux at $\lambda 1270$ Å, where the nebula's radiation can be neglected. The flux at $\lambda 1270$ Å was obtained via a simple average of the values for the quiescent state of Z And in 1978−1982, which are listed in Table 2 of Fernández-Castro *et al.* [8]. The mean flux at $\lambda 1270$ Å was corrected for interstellar reddening: $F_{\text{hot}} = (1.55 \pm 0.11) \times 10^{-12}$ erg cm^{-2} s^{-1}, where the uncertainty is the standard error. The temperature of the hot component during the quiescent state was determined using various methods [2, 23, 24]; the spread of resulting values is in the range $(1-2) \times 10^5$ K. For our estimates, we adopted the average value 1.5×10^5 K. The UBV fluxes of the hot star in the quiescent state (assuming it radiates as a blackbody) are listed in Table 2. The adopted temperature and UV flux of the hot component at $\lambda 1270$ Å yield a radius for this star of $R_{\text{hot}} = r[F_{\text{hot}}/\pi B(\lambda 1270, 1.5 \times 10^5)]^{0.5} \approx 0.06 R_{\odot}$ (Table 3), with an uncertainty of less than $0.01 R_{\odot}$.
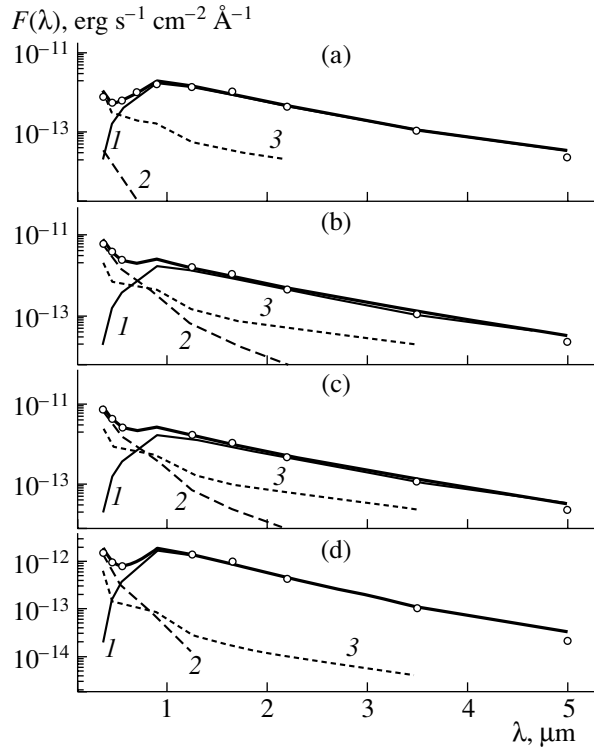

Fig. 2. SED of Z And in its quiescent state and the active phase of 2000−2002: (a) quiescent state in September 1999, (b) active phase in November 2000, (c) active phase in December 2000, (d) active phase in January 2002. Shown are the observed fluxes (open circles), the SED of the M4.5 giant (thin solid curves (*1*), the SEDs of the hot star and nebula (dashed curves (*2*) and dotted curves (*3*), respectively), and the SED of the total radiation of the three components (bold solid curves).

The quantity $\pi B(\lambda 1270, 1.5 \times 10^5$ K) is the flux from a blackbody at $\lambda 1270$ Å for a temperature of 1.5×10^5 K.

The U, B, and V radiation of the nebula was obtained by subtracting the stellar fluxes from the observed total fluxes. We then calculated the spectra of the nebula, varying the electron temperature and the emission measure and fitting the results to the observed SED. To calculate the spectrum of the nebula, we must know the ionization state of helium, which we estimated using the fluxes in the HeI $\lambda 4471$ Å and HeII $\lambda 4686$ Å lines from our high-resolution optical spectra, which were obtained at orbital phase 0.228 in the quiescent state [25]. It turned out that doubly ionized helium dominates in the nebula in the quiescent state, so that the radiation of the nebula is mainly due to the emission by hydrogen and ionized helium [26]. In addition, the calculations require the values of γ_ν—the volume emissivities of these elements, which are determined by recombinations and free−free transitions. Due

Table 2. Continuum fluxes of the components of Z And in units of 10^{-12} erg cm^{-2} s^{-1} Å^{-1}

Date	Component	U	B	V	R	I	J	H	K	L	M
September 15, 1999, quiescent state	Cool	0.020	0.160	0.376	0.710	1.755	1.343	0.856	0.439	0.113	0.034
	Hot	0.030	0.014								
	Neb	1.012	0.293	0.257	0.195	0.157	0.056	0.032	0.020		
	TF	1.062	0.467	0.633	0.905	1.912	1.399	0.888	0.459	0.113	0.034
	OF	0.988 ± 0.022	0.497 ± 0.010	0.625 ± 0.017	1.093 ± 0.020	1.870 ± 0.017	1.347 ± 0.025	0.877 ± 0.023	0.432 ± 0.004	0.105 ± 0.003	0.023 ± 0.001
	d	7	−6	1	−17	2	4	1	6	8	48
November 22, 2000, active phase	Cool	0.020	0.160	0.376	0.710	1.755	1.343	0.856	0.439	0.113	0.034
	Hot	5.537	2.934	1.336	0.730	0.340	0.063	0.022	0.007		
	Neb	1.983	0.717	0.631	0.547	0.432	0.147	0.081	0.051	0.020	
	TF	7.540	3.811	2.343	1.987	2.527	1.553	0.959	0.497	0.133	0.034
	OF	7.205 ± 0.054	3.848 ± 0.036	2.492 ± 0.022			1.591 ± 0.002	0.927 ± 0.001	0.461 ± 0.001	0.117 ± 0.001	0.024 ± 0.001
	d	5	−1	−6			−2	3	8	14	42
December 6, 2000, active phase	Cool	0.020	0.160	0.376	0.710	1.755	1.343	0.856	0.439	0.113	0.034
	Hot	6.257	3.315	1.510	0.826	0.384	0.071	0.024	0.008		
	Neb	2.382	0.861	0.758	0.657	0.519	0.176	0.098	0.061	0.024	
	TF	8.659	4.336	2.644	2.192	2.658	1.590	0.978	0.508	0.137	0.034
	OF	8.662 ± 0.064	4.257 ± 0.039	2.682 ± 0.023			1.635 ± 0.001	0.944 ± 0.001	0.478 ± 0.001	0.120 ± 0.001	0.024 ± 0.001
	D	0	2	−1			−4	4	6	14	42
January 27, 2002, Active phase	Cool	0.020	0.160	0.376	0.710	1.755	1.343	0.856	0.439	0.113	0.034
	Hot	1.247	0.629	0.274	0.146	0.066	0.012				
	Neb	0.622	0.140	0.123	0.107	0.084	0.029	0.016	0.010	0.004	
	TF	1.889	0.929	0.773	0.963	1.905	1.384	0.872	0.449	0.117	0.034
	OF	2.067 ± 0.044	0.863 ± 0.018	0.795 ± 0.022			1.372 ± 0.025	0.853 ± 0.007	0.428 ± 0.001	0.105 ± 0.001	0.022 ± 0.001
	D	−9	8	−3			1	2	5	11	54

Note: Cool, cool component; Hot, hot component; Neb, the nebula; TF = Cool + Hot + Neb is the total flux; OF, observed flux; $d = (\mathrm{TF} - \mathrm{OF})/\mathrm{OF}$ in per cent.

to blending of the high-order Balmer lines, symbiotic systems with strong hydrogen-emission spectra (e.g., AG Peg and AG Dra) have almost identical continua on both sides of the series limit (see, e.g., [27, 28]). This may also be true for Z And, since its Balmer lines are fairly strong and are visible up to H$_{32}$ [24]; therefore, we used the hydrogen-emission coefficient on the short-wavelength side of the Balmer jump for the value at $\lambda3650$ Å [29]. We first used the adopted electron temperature to calculate γ_ν for the

Table 3. Parameters of the hot stellar and nebular components of Z And derived from the observations

Date	Hot component		Nebula	
	T_{eff}, K	R, $R_\odot$	T_e, K	n_e^2V, 10^{59} cm^{-3}
September 15, 1999, quiescent state	150 000	0.06	20 000 ± 1000	4.7 ± 0.3
November 22, 2000, active phase	35 000 ± 1000	2.22 ± 0.06	20 000 ± 1000	17.4 ± 0.6
December 6, 2000, active phase	35 000 ± 1000	2.36 ± 0.07	20 000 ± 1000	20.9 ± 0.5
January 27, 2002, active phase	58 000 + 2000/−3000	0.72 + 0.04/−0.03	20 000 ± 2000	3.4 ± 0.4

centers of the photometric bands. We then estimated the emission measure using these values and the derived fluxes of the nebula. Finally, the emission coefficients γ_ν were obtained by fitting the computed spectrum of the nebula to the observed spectrum. As a result, we have used the following values of γ_ν (in units of 10^{-40} erg cm^3 s^{-1} Hz^{-1}) in the infrared: $\gamma_J(\mathrm{H}^0) = 5.63$, $\gamma_H(\mathrm{H}^0) = 5.49$, $\gamma_K(\mathrm{H}^0) = 6.29$, $\gamma_L(\mathrm{H}^0) = 6.42$, $\gamma_J(\mathrm{He}^+) = 20.7$, $\gamma_H(\mathrm{He}^+) = 21.7$, $\gamma_K(\mathrm{He}^+) = 23.0$, $\gamma_L(\mathrm{He}^+) = 24.3$.

We used the relationship between the observed flux F_λ and the emission measure n_e^2V in the form

$$F_\lambda = \{[1 + 2\alpha(\mathrm{He})]/4\pi r^2\}[\gamma_\nu(\mathrm{H}^0, T_e)$$
$$+ \alpha(\mathrm{He})\gamma_\nu(\mathrm{He}^+, T_e)]n_e^2V(c/\lambda^2) \times 10^{-8}.$$

We assumed the value $\alpha = 0.1$ for the helium abundance [26]; the distance to Z And r is 1.12 kpc. The electron temperature $T_e = 20\,000 \pm 1000$ K and emission measure $n_e^2V = 4.7 \pm 0.3 \times 10^{59}$ cm^{-3} (Table 3) best reproduce the radiation of the circumstellar nebula in the quiescent state of Z And at 0.36–0.8 μm (the *UBVRI* bands). The uncertainties in T_e and n_e^2V are determined by the observational errors.

3.2. Active Phase

Considerable changes took place in the SED of Z And in the active phase, which can also be noted in the infrared (Table 1). The radiation of the nebula is clearly visible in the *L* band.

As a first approximation, we estimated the emission measure of the nebula using the *L* flux observed in the active phase (after subtracting the giant's flux) and the electron temperature derived for the quiescent state. The parameters (emission measure and electron temperature) were then refined based on the photometric data.

At our last observational epoch (Table 1, last line), the *HKL* fluxes of Z And had reached values typical of the quiescent state, when only the cool giant (Table 2)

contributes in these bands. However, the *UBVJ* fluxes still exceeded their values for the quiescent state. It is natural to attribute the excess *UBVJ* radiation to the radiation of the nebula. Using the estimated excess *J* flux, we determined an approximate value for the emission measure, which was then refined using the *UBV* data. The optimal parameters of the nebula during the active phase are listed in Table 3. The electron temperature was again 2×10^4 K, though ionized helium no longer dominated in the nebula. The ionization state of helium was determined using the same lines as we had for the quiescent state [25] but using the observations of November 17 and December 6, when we obtained our photometric observations. Singly ionized helium dominated in the nebula in this phase, and the continuum radiation of the nebula was most likely formed by hydrogen and neutral helium.

As is noted above, we used the emission coefficient of the hydrogen continuum for the short-wavelength limit of the Balmer jump when calculating the *U* flux of the nebula in the quiescent state. The situation with the *U* flux of the nebula during the optical maximum was different. The Hγ linewidth decreased [30] at that time, probably indicating that the widths of the higher-order Balmer lines also decreased. For the same reason, the blending of lines located near the Balmer jump also decreased, resulting in a decreased flux at 3650 Å. Therefore, we took a simple average of the values for the short- and long-wavelength sides of the Balmer jump for the hydrogen emission coefficient in the active phase [29].

Further, we calculated the IR emission coefficients of neutral hydrogen and helium and of singly ionized helium for the *JHKL* photometric bands (1.25–3.5 μm). If the abundance of helium is 0.1 and most of the helium is doubly ionized, its contribution to the continuum radiation is comparable to that of hydrogen. However, if most of the helium is singly ionized, its contribution is about 10% of the hydrogen continuum. The free–free continuum radiation dominates in the infrared [29]. Therefore, for the active phase,

when singly ionized helium dominates in the nebula and the main contribution to the continuum radiation comes from neutral helium, we calculated the IR emission coefficients only for free–free transitions and used these as the total emission coefficients. This results in errors in the calculated fluxes of the nebula of no more than 5%. The IR emission coefficients were calculated using the formulas from [31, 32]. As a result, we used the following values for neutral helium (in units of 10^{-40} erg cm^3 s^{-1} Hz^{-1}): $\gamma_J(\mathrm{He}^0) = 3.64$, $\gamma_H(\mathrm{He}^0) = 4.18$, $\gamma_K(\mathrm{He}^0) = 4.67$, $\gamma_L(\mathrm{He}^0) = 5.27$. The resulting estimate of the emission measure of the nebula in the active state is almost a factor of 4.4 greater than in the quiescent state at similar phases of the orbital motion (Table 3).

We determined the *UBVJHK* fluxes of the hot component during the active phase by subtracting the radiation of the giant and nebula from the observed total fluxes. The radiation of the hot component was then fitted to a blackbody curve for a specified temperature, and the radius of the hot component, estimated from the observed fluxes. The resulting estimates of the temperature and radius of the hot component in the active phase are listed in Table 3. The data in Table 3 testify to a considerable increase in the radius of the hot component in the active phase compared to the quiescent state. The parameters of the hot component can be refined using the recently obtained UV fluxes at $\lambda\lambda1059$ and 1104 Å [10], where this component is the only source of radiation. The fluxes at these wavelengths (corrected for interstellar extinction) were 36.2×10^{-12} and 34.5×10^{-12} erg cm^{-2} s^{-1} Å^{-1}, respectively. These values are approximately a factor of 3 lower than corresponding blackbody fluxes for a temperature of 3.5×10^4 K and a radius of $2.36R_\odot$ (Table 3).

This discrepancy means that the optical and infrared data, on the one hand, and the UV data, on the other hand, cannot be explained using a single set of parameters for the hot component. At the epoch of the optical maximum, the phase of the orbital motion of the binary system was 0.150, so that the hot component was behind the giant; it may be possible to explain the above discrepancy as an effect of Rayleigh scattering of the hot component's radiation by neutral hydrogen in the extended atmosphere of the cool component [8]. However, Z And is a noneclipsing system [4], and Rayleigh scattering has been detected only in spectra obtained during the flares of 1984 and 1985. The discrepancy may be a consequence of the expansion of the hot component during the active phase (see below), with some of its radiation passing through denser atmospheric layers of the cool giant.

4. DISCUSSION

Our photometric data obtained before, during, and in the decline of the optical flare of Z And in 2000–2002 indicate that a considerable redistribution of energy took place in the observed spectrum during the optical maximum; such a redistribution is typical of systems containing a luminous hot star and no accretion disk [33].

Mikołajewska and Kenyon [4] concluded that the hot component of Z And expands by a factor of almost 100 in the active phase, while its effective temperature drops. The hot-component radius we found for the epoch of maximum brightness suggests an expansion by approximately a factor of 40; this may be related to the ejected envelope with its photospheric characteristics.

The radius and temperature of the hot component at the epoch of the optical maximum yield a luminosity of $\sim 7500L_\odot$. A smaller radius for the hot component is not consistent with our spectral data [30], which indicate the presence of absorption components in the $\lambda\lambda4471$ and 4713 Å triplet helium lines. The residual intensity of the HeI $\lambda4471$ Å line was 0.46 in November and 0.60 in December 2000. At that time, the contribution of the cool giant to the total radiation of the system was 0.07–0.08, and the appearance of these lines could be due to the hot component. In this case, the depth of the HeI $\lambda4471$ Å line should be smaller than the contribution of the hot component determining the radius. If we suppose that the temperature of the hot component during the active phase was lower than the temperatures typical for O stars (i.e., $\sim 3 \times 10^4 - 4 \times 10^4$ K), it is difficult to explain the presence of a strong $\lambda4686$ Å HeII emission line [30] in the spectrum of Z And. The blackbody temperature consistent with the observed SED is indirectly confirmed by the interpretation of the absorption components of some other lines in the spectrum of Z And. The presence of only absorption lines of the helium triplet has been observed in the spectra of some O stars with $T_{eff} > 4 \times 10^4$ K and $\log g \sim 3.5$–3.8; this is explained as an effect of blocking of the UV flux at 228–912 Å by lines of metals and decreased absorption of the flux in the helium singlet lines [34]. It has been suggested that the same mechanism could be responsible for the presence of the triplet absorption lines in the spectrum of Z And [30].

The fluxes of Z And we observed during the optical flare prevent us from reducing the inferred radius of the hot component, since, in turn, this would be in contradiction with the observed depth of the absorption lines. On the other hand, the hypothesis that the hot component has a much lower temperature

encounters difficulties in connection with explaining the presence of certain lines in the spectrum. Because of the uncertainty in the interstellar reddening, the luminosity of the hot component of Z And during its optical flare has an error of about 100%.

The temperature and radius of the hot component obtained for the quiescent state imply a bolometric luminosity of $1700L_\odot$. Nussbaumer and Vogel [26] found that estimates of the temperature of the hot component in this period lie in the range $10^5 - 2 \times 10^5$ K. The temperature of the hot component of Z And in the quiescent state may be higher than the hot components of some other symbiotic systems (e.g., AG Peg and AG Dra) and close to the value 2×10^5 K, since the nebula is dominated by the He^{++} rather than He^+ radiation (as is the case in the AG Peg and AG Dra systems). If we use the temperature 2×10^5 K and take the observed UV flux at 1270 Å to be $\sim 1.55 \times 10^{-12}$ erg cm^{-2} s^{-1} Å^{-1}, the bolometric luminosity of the hot component is $\sim 3700L_\odot$. The inferred luminosities in the quiescent and active phases admit evolution with an approximately constant bolometric luminosity. However, we can resolve the question of whether such evolution is present only based on UV data obtained in both the quiescent and active phases.

Our last set of observations (Tables 1 and 2) implies that the radius of the hot component during the optical flare considerably exceeds its value for the quiescent state, but the effective temperature is much lower. These parameters may correspond to a pseudophotosphere formed during the active phase as a result of the outflow of material from the flaring star. This outflow should have an optical depth of about unity in the continuum at a distance equal to the obtained radius.

Paczyński and Żytkow [35] and Fujimoto [36] have suggested that the burning of hydrogen on the surface of an accreting degenerate dwarf proceeds in a stationary regime within a narrow range of accretion rates and that small variations in the accretion rate can result in the expansion of the accretor with the bolometric luminosity remaining approximately constant [33]. This scenario was adopted by Fernández-Castro *et al.* [8] as the most probable one for the optical flares of Z And.

Bisikalo *et al.* [37] have modeled the evolution of the gas flows in the vicinity of the hot compact component of Z And, which accretes material from the stellar wind of the cool giant. They calculated the accretion rate as a function of the wind velocity. One of their results is that the accretion disk surrounding the compact source exists as long as the velocity has certain low values. When the velocity becomes slightly higher, the disk ceases to exist, and material is accreted directly from the flow. There exists a range of velocities at which the accretion changes: the disk decays, and its mass is accreted over a time covering less than 0.1 of the orbital period of Z And after the epoch of the velocity increase. As a result, the accretion rate increases by a factor of 20−30. It is thought that the increase in the accretion rate is accompanied by a change in the nuclear-burning regime and can lead to the ejection of matter by the accreting component. The correct explanation of our photometric data should probably be sought for in the framework of these theoretical considerations.

5. CONCLUSIONS

Let us summarize the main results of this paper.

(1) The effective temperature of the cool component (M4.5 giant) of the binary system Z And is 3400 ± 270 K, the observed integrated flux is 2.25×10^{-8} erg cm^{-2} s^{-1}, and the radius of the star is $(85 \pm 14)R_\odot$.

(2) In the quiescent state of Z And, the electron temperature of the nebula and its emission measure are 2×10^4 K and 4.7×10^{59} cm^{-3}.

(3) During the flare of 2000−2002, the emission measure of the nebula increased by approximately a factor of 4.4, to the value 20.9×10^{59} cm^{-3}; at the same time, the electron temperature of the nebula remained virtually unchanged.

(4) During the active phase of Z And, the radius of the hot stellar component of the system increased by approximately a factor of 40, reaching $2.36R_\odot$ at the epoch of maximum brightness of the star. This radius probably corresponds to the ejected envelope, which imitates a photosphere. The temperature of the hot component in that period was almost 3.5×10^4 K.

ACKNOWLEDGMENTS

REFERENCES

1. U. Murset and H. M. Schmid, Astron. Astrophys., Suppl. Ser. **137**, 473 (1999).
2. T. Fernandez-Castro, A. Cassatella, A. Gimenez, and R. Viotti, Astrophys. J. **324**, 1016 (1988).
3. L. Formiggini and E. M. Leibowitz, Astron. Astrophys. **292**, 534 (1994).

4. J. Mikolajewska and S. J. Kenyon, Astron. J. **112**, 1659 (1996).

5. J. L. Sokoloski and L. Bildsten, Astrophys. J. **517**, 919 (1999).

6. P. Swings and O. Struve, Astrophys. J. **93**, 356 (1941).

7. A. A. Boyarchuk, Astron. Zh. **44**, 1016 (1967) [Sov. Astron. **11**, 818 (1968)].

8. T. Fernandez-Castro, R. A. Gonzalez-Riestra, A. Cassatella, *et al.*, Astrophys. J. **442**, 366 (1995).

9. A. Skopal, D. Chochol, T. Pribulla, and M. Vanko, Inform. Bull. Var. Stars, 5005 (2000).

10. J. L. Sokoloski, S. J. Kenyon, A. K. H. Kong, *et al.*, in *The Physics of Cataclysmic Variables and Related Objects,* Ed. by B. T. Gansicke, K. Beuermann, and K. Reinsch, ASP Conf. Ser. **261**, 667 (2002).

11. H. L. Johnson, R. I. Mitchel, B. Iriarte, and W. Z. Wisniewski, Comm. Lunar Planet. Lab. **4**, 99 (1966).

12. J. Koorneef, Astron. Astrophys. **128**, 84 (1983).

13. T. S. Belyakina, Izv. Krym. Astrofiz. Obs. **41**, 275 (1970).

14. A. Skopal, M. Vanko, T. Pribulla, *et al.*, Contrib. Astron. Obs. Skalnate Pleso. **32**, 1 (2002).

15. H. L. Johnson, Ann. Rev. Astron. Astrophys. **4**, 163 (1966).

16. N. Tomov, O. G. Taranova, and M. T. Tomova, Astron. Astrophys. **401**, 669 (2003).

17. T. Eiroa, H. Hefele, and Qian Zhong-Yu, in *The Nature of Symbiotic Stars,* Ed. by M. Friedjung and R. Viotti (Reidel, Dordrecht, 1982), IAU Coll. **70**, 43 (1982).

18. T. S. Belyakina, Izv. Krym. Astrofiz. Obs. **84**, 49 (1992).

19. G. T. van Belle, B. F. Lane, R. R. Thompson, *et al.*, Astron. J. **117**, 521 (1999).

20. M. A. Fluks, B. Plez, P. S. Thé, *et al.*, Astron. Astrophys., Suppl. Ser. **105**, 311 (1994).

21. A. V. Kharitonov, V. M. Tereshchenko, and L. N. Knyazeva, *Spectroscopy Star Catalogue* (Nauka, Alma-Ata, 1988) [in Russian].

22. H. M. Dyck, G. W. Lockwood, and R. W. Capps, Astrophys. J. **189**, 89 (1974).

23. U. Murset, H. Nussbaumer, H. M. Schmid, and M. Vogel, Astron. Astrophys. **248**, 458 (1991).

24. A. A. Boyarchuk, Izv. Krym. Astrofiz. Obs. **38**, 155 (1967).

25. N. Tomov and M. T. Tomova (2004, in press).

26. H. Nussbaumer and M. Vogel, Astron. Astrophys. **213**, 137 (1989).

27. N. Tomov, M. T. Tomova, and D. V. Raikova, Astron. Astrophys., Suppl. Ser. **129**, 479 (1998).

28. M. T. Tomova and N. A. Tomov, Astron. Astrophys. **347**, 151 (1999).

29. D. E. Osterbrock, *Astrophysics of Gaseous Nebulae* (Freeman, San Francisco, 1974), p. 73.

30. N. Tomov, M. T. Tomova, and R. K. Zamanov, in *Symbiotic Stars Probing Stellar Evolution*, Ed. by R. L. M. Corradi, J. Mikolajewska, and T. J. Mahoney, ASP Conf. Ser. (2003, in press).

31. R. L. Brown and W. G. Mathews, Astrophys. J. **160**, 939 (1970).

32. A. Skopal, Contribs Astron. Observ. Skalnate Pleso. **31**, 119 (2001).

33. J. Mikolajewska and S. J. Kenyon, Mon. Not. R. Astron. Soc. **256**, 177 (1992).

34. J. Herrero, J. Puls, and M. R. Villamariz, Astron. Astrophys. **354**, 193 (2000).

35. B. Paczynski and A. Zytkow, Astrophys. J. **222**, 604 (1978).

36. M. Y. Fujimoto, Astrophys. J. **257**, 767 (1982).

37. D. V. Bisikalo, A. A. Boyarchuk, E. Yu. Kil'pio, and O. A. Kuznetsov, Astron. Rep. **46**, 1022 (2002).

Translated by G. Rudnitskiĭ

The Eclipsing Variable CM Dra:
Surface Activity and Orbital Elements

A. V. Kozhevnikova, V. P. Kozhevnikov,
P. E. Zakharova, T. S. Polushina, and M. A. Svechnikov

Astronomical Observatory, Ural State University, pr. Lenina 51, Yekaterinburg, 620083 Russia
Received October 10, 2003; in final form, January 9, 2004

Abstract—A long series of observations of the low-mass eclipsing system CM Dra (M_1, $M_2 < 0.25 M_\odot$, dM4.5e, $V = 12.9^m$, $P_{orb} = 1.27^{\rm d}$) was obtained in 1996−1997 as part of the international TEP (Search for Transits of Extrasolar Planets) project. The observations were obtained on the 70-cm telescope of the Astronomical Observatory of Ural State University with an automatic two-star photometer. The total duration of the data series was 155 h. We construct and analyze a master light curve, including calculation of the photometric and absolute orbital elements. Evidence for surface activity was detected. The light curve of CM Dra outside eclipses shows low-amplitude (0.024^m) brightness variations, suggesting the presence of a starspot on one of the components. The amplitude of these variations stayed the same over 20 years, but the phase of the brightness maximum shifted by one-third of the orbital period, possibly reflecting changes in the longitude or asymmetry of a single large polar spot. We detected four flares whose energies were factors of several hundred higher than those of solar flares. The derived flare rate agrees with data for CM Dra in the literature but is much lower than is typical for Population I stars in our Galaxy.
© *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The eclipsing variable CM Dra is a low-mass spectroscopic binary on the lowest part of the main sequence. Both its components have masses below 0.25 solar masses; their spectral types are dM4.5e + dM4.5e. The system's apparent magnitude is 12.9^m, and the orbital period is $P = 1.27$ days [1]. In 1993, Strassmeier *et al.* [2] included this system in their catalog of binaries displaying chromospheric activity, and evidence for solar-type activity in the system was found. It is currently classified as a BY Dra star [3]. These are late-type flare stars with magnetic fields, which show quasi-sinusoidal brightness variations due to the presence of spots on their surfaces [4, 5]. The CM Dra system was studied in detail in 1977 [1], when its photometric and absolute elements were determined and low-amplitude, sinusoidal, brightness variations characteristic of starspot activity were detected. Until recently, there were no data confirming the presence of these variations. The photometric and absolute orbital elements of CM Dra were determined only once, in 1977 [1].

Certain parameters of CM Dra, primarily its small components (their total surface brightness is about 12% of that of the solar surface) and its high orbital inclination (close to 90°), make this star a suitable object for searches for terrestrial-size planets via differential photometry. To accumulate a sufficient number of observations, the international Transits of Extrasolar Planets (TEP) project was started to conduct photometric searches for transits of extrasolar planets across the disks of the stars in the CM Dra system [6]. Our observations of CM Dra at the Astronomical Observatory of Ural State University in 1996−1997 were part of this program.

In addition to the primary goal of these studies (searches for planetary transits), we decided to use the large amount of accumulated data to improve the fundamental orbital parameters for this binary system. CM Dra by itself is of interest as an extremely-low-mass red star among studied close binaries on the main sequence. In this paper, we present the results of our photometric observations, our studies of the light curve of CM Dra based on our observations, an analysis of flare and starspot activity on the star, and a new determination of the system's fundamental parameters.

2. OBSERVATIONS

We observed CM Dra using an automatic two-star photometer mounted on the 70-cm telescope of
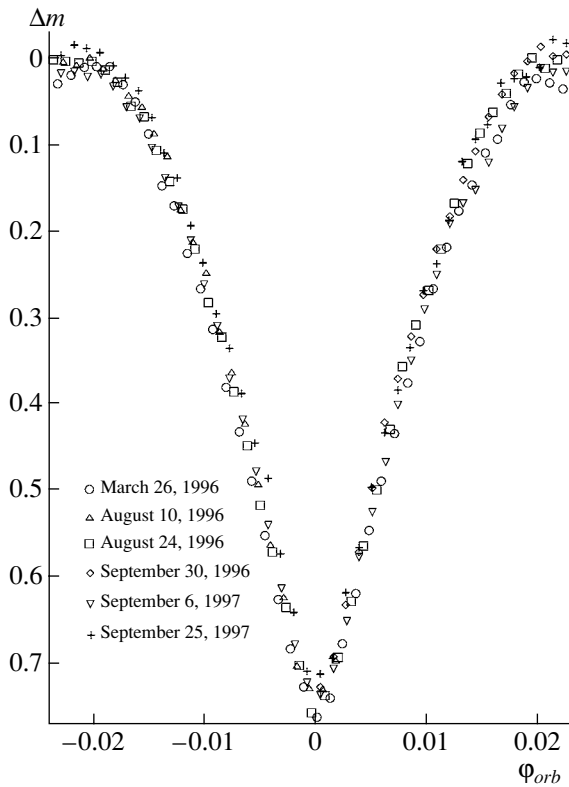
Fig. 1. The primary minimum of CM Dra from our observations. Different symbols represent data points for different nights of observations.
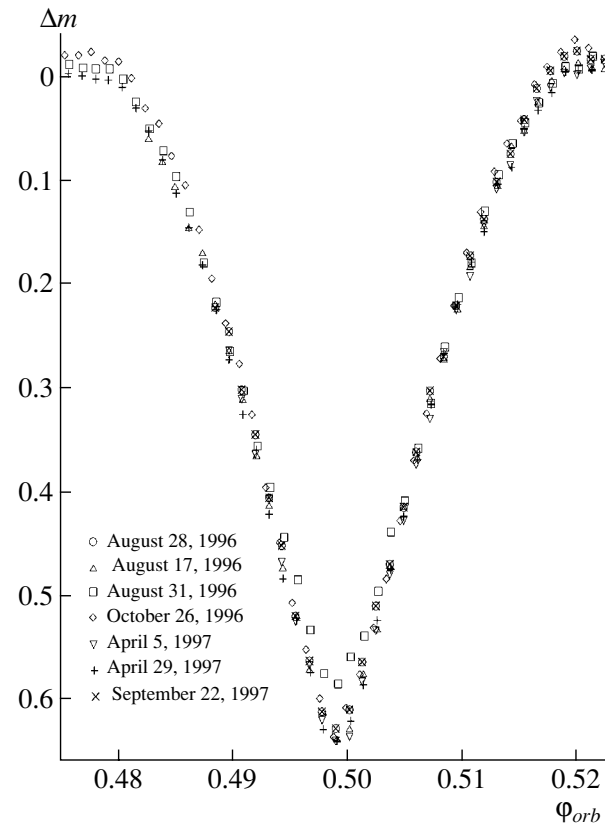

Fig. 2. The secondary minimum of CM Dra. Different symbols represent data points for different nights of observations.

the Astronomical Observatory of Ural State University [7]. We carried out R differential photometry, in accordance with the requirements of the TEP project. CM Dra and the comparison star were kept in their respective diaphragms by means of offset guiding. The counts from the photometer were directly recorded by a computer. Periodically (once each 34 min), the stars were automatically removed from the diaphragms, and the sky background was measured for 32 s. The sky background was approximated with a third-power polynomial and subtracted from the readings for CM Dra and the comparison star. We then found the differences in the magnitudes of CM Dra and the comparison star, taking into account differences in the sensitivities of the two photometer channels. Since the angular separation of CM Dra and the comparison star was only a few minutes of arc, these differential magnitudes are already corrected for first-order atmospheric extinction and the absorption of light by thin clouds that sometimes appeared during the observations. For our subsequent analysis, we used differential magnitudes averaged over 128-s time intervals.

We obtained the observations during 43 nights (155 observing hours, 68 of them in 1996 and 87 in 1997), yielding 4359 brightness estimates (magnitude differences between CM Dra and the comparison star). The rms error of a single measurement was 0.005^m.

3. CONSTRUCTION OF THE LIGHT CURVE

It was difficult to construct the mean light curve of CM Dra using all the observations directly, because the individual light curves for each night demonstrated slow variations due to differential extinction, since CM Dra and the comparison star have very different colors. The participants of the TEP project removed these changes by subtracting a polynomial chosen as the best approximation of the outside-eclipse parts of individual light curves for one night [8]. When plotting a master light curve using our own data at the Astronomical Observatory of Ural State University, we decided not to exclude these slow variations in the individual light curves, since that should average out thanks to the large number of observations (155 h). Apart from extinction, such variations can reflect brightness variations due to
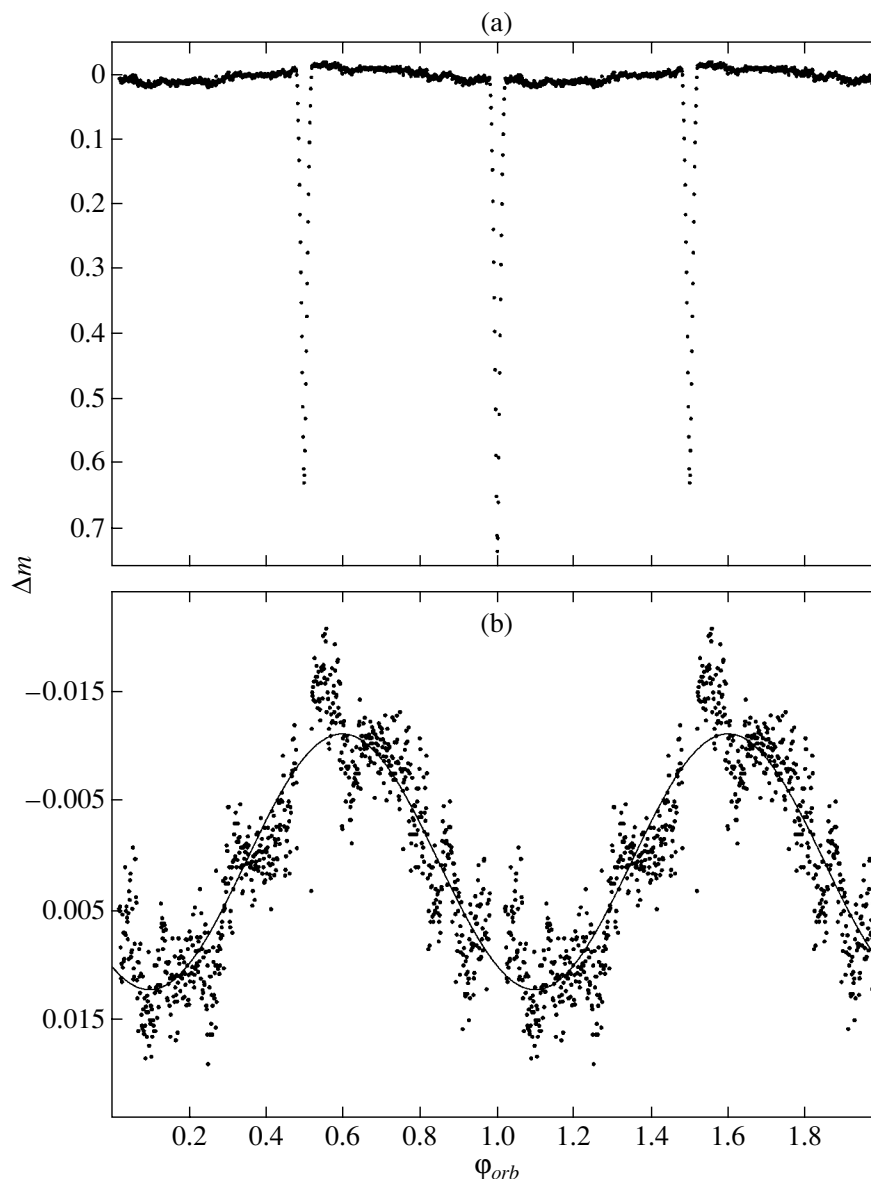
Fig. 3. (a) Master light curve of CM Dra containing 856 data points, obtained by averaging all 155 hours of observations. Each point is an average of four to six individual brightness estimates. (b) Outside-eclipse brightness variations on an expanded scale (minima excluded). The solid curve is the approximating sine function.

starspots, if they are present on the binary components. Brightness variations of this kind should be manifest on approximately the same time scale as extinction; it is believed that the rotational period for the CM Dra components is synchronized with the orbital period, 1.27 days. The CM Dra system is very old, and its components are assumed to exhibit completely synchronized rotation [9]. Thus, having retained the trends for the individual light curves, we were able to detect the outside-eclipse brightness variations of CM Dra that are discussed in the following section.

Unfortunately, our instrumental color system changed slightly during the almost two-year period of our observations: at different times, we used different photomultipliers (FEU-106, FEU-136, FEU-79); in addition, we used ordinary glass filters as well as dichroic filters. This could result in the mean differential magnitude of CM Dra being somewhat different in different observing seasons. In addition, it is known that the magnitudes of many BY Dra stars can vary by about 0.2^m between seasons, which is explained as an effect of the changing pattern of the spotted regions [10]. In our case, we were not able

Table 1. Parameters of the outside-eclipse variations of the CM Dra light curves

Time interval	Number of nights	Duration, h	Half-amplitude, mag	Phase of maximum
Apr.−Oct. 1996	22	68	0.010 ± 0.002	0.517 ± 0.023
Apr.−Oct. 1997	21	87	0.013 ± 0.001	0.635 ± 0.009
All data	43	155	0.012 ± 0.001	0.597 ± 0.008

to distinguish and separate these two types of variations. Therefore, we derived averaged light curves for each observing season (April−May 1996, August−October 1996, April−May 1997, August−October 1997) and subtracted them from the individual light curves for each night during the corresponding season.

The times of minima were computed using the ephemeris from [6]:

$$\text{Min I} = \text{HJD} 2449830.756967(01) \qquad (1)$$
$$+ 1.268389851(10)E,$$

$$\text{Min II} = \text{HJD} 2449831.990019(01) \qquad (2)$$
$$+ 1.268389851(10)E.$$

Figures 1 and 2 present the primary and secondary minima of CM Dra indicated by our observations. Different symbols are used to plot data points for different observing nights. The data acquired for the secondary minimum on August 31, 1996, were excluded from consideration, since they deviate strongly from the mean, possibly due to a stellar flare.

The mean light curve of CM Dra obtained by averaging all our observations is presented in Fig. 3a.

Table 2. Flare parameters of CM Dra

Date	Duration, min	Maximum amplitude, Δm	Maximum energy, $(E_{star} + E_{flare})/E_{star}$
1996	No flares detected		
April 7, 1997	92	0.21 ± 0.01	1.210
August 9, 1997	70	0.10 ± 0.01	1.096
August 29, 1997	21	0.03 ± 0.01	1.028
September 6, 1997	34	0.21 ± 0.01	1.215

Each data point is the result of averaging four to six brightness estimates for different observing nights. The two narrow minima with nearly equal depth are evident: the depth of the primary minimum is 0.74^m, and that of the secondary minimum, 0.63^m. Both minima are of short duration, indicating that the components are strongly detached. The shape of the light curve shows no ellipsoidality effects for the components.

4. EVIDENCE FOR ACTIVITY ON THE COMPONENTS OF CM Dra

4.1. Outside-eclipse Brightness Modulations of CM Dra

Our light curve of CM Dra reveals an interesting feature: the brightness outside eclipses demonstrates smooth, low-amplitude variations (Fig. 3a). Figure 3b presents the outside-eclipse brightness variations (with minima excluded) on an expanded scale. We obtained a least-squares fit to these variations using a sine curve; the best-fit amplitude is 0.024^m and the maximum brightness occurs at orbital phase 0.60.

As is noted above, sinusoidal brightness variations in the outside-eclipse parts of the light curve of CM Dra with the same period and almost the same amplitude ($A = 0.022^m$) were found in 1977 [1]. Such low-amplitude, sinusoidal variations are characteristic of BY Dra stars and can be explained by the presence of dark spots on the surface of the rotating star [4].

However, there is an interesting difference between our results and those of 1977. The phase of maximum brightness derived 20 years before our observations was 0.28 of the orbital period, whereas the phase of maximum brightness indicated by our observations is 0.60. This phase shift is about one-third of the orbital period.

Since our observations cover almost two years, we also searched for a change in the phase of maximum brightness between 1996 and 1997. For this purpose, we plotted separate light curves for 1996 and 1997, then determined the phase of maximum brightness
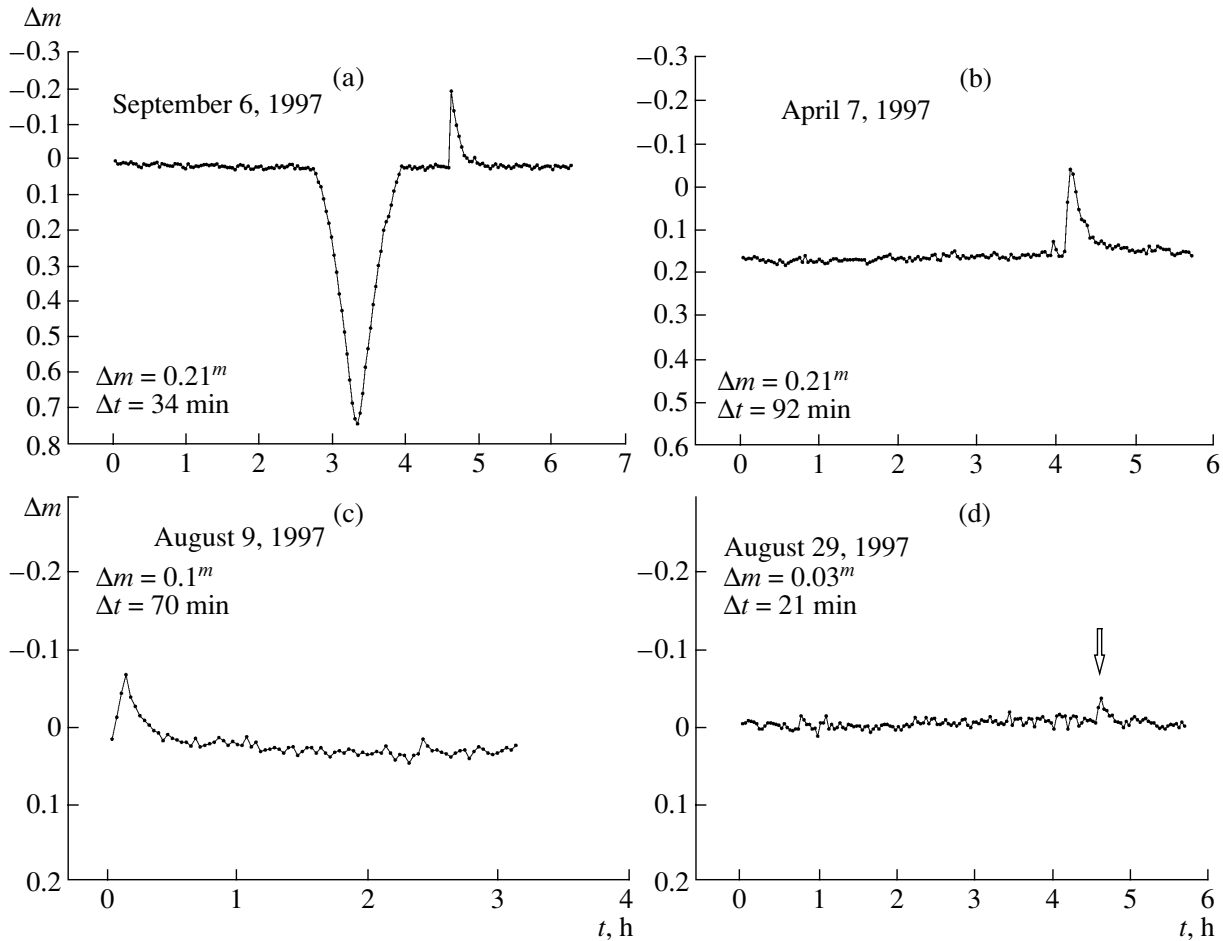
Fig. 4. Flares detected in individual light curves of CM Dra over the entire period of the observations; the amplitudes are between 0.03^m and 0.21^m and the durations between 21 and 92 minutes.

using individual sine-curve approximations for each of these light curves.

Unfortunately, as is noted above, the distribution of the data points around the sine curve is not normal. To obtain the most realistic estimates of the rms errors as given by the least-squares technique, we combined data points in the light curves within phase bins of 0.02. We then fit the outside-eclipse parts of these light curves with sine curves; each part contained 50 data points. The results with the corresponding rms errors are collected in Table 1.

Though Table 1 shows that the difference between the phases of maximum brightness for our 1996 and 1997 observations slightly exceeds 3σ, it would be premature to conclude that there was a real change in the phase of maximum brightness between 1996 and 1997. We can currently conclude only that this phase can vary considerably over decades, as follows from the comparison of our data with those obtained 20 years earlier.

The similar shapes and amplitudes for the brightness variations of CM Dra in 1977 and in 1996–1997 may indicate that these variations are due to the same long-lived starspot feature. The shift of the maximum-brightness phase is difficult to explain as a result of asynchronous rotation of the components, because the system is very old, quite detached, and believed to be synchronized [9].

Figure 3a shows that the outside-eclipse brightness variations of CM Dra are smooth, without a flat minimum or maximum. This shape for the variations may indicate that the spotted area is close to the rotational pole of one of the stars; otherwise, it would have become invisible at some point in the orbit [1]. The low amplitude of the brightness variations may provide further evidence that the spotted area is near the pole. The catalog of chromospherically active stars [2] shows that brightness variations due to starspots on other BY Dra stars usually have large amplitudes, up to several tenths of a magnitude. If the size of the spotted area on CM Dra is similar to the sizes of the

Table 3. Photometric orbital elements of CM Dra

Orbital elements	Our data (1997)	Lacy's data [1] (1977)
k	0.899 ± 0.004	0.93 ± 0.02
r_1	0.0681 ± 0.0002	0.0669 ± 0.0005
r_2	0.0612 ± 0.0002	0.0624 ± 0.0005
X_1	0.8 ± 0.02	0.5 ± 0.03
X_2	0.5 ± 0.03	0.5 ± 0.03
e	0.005 ± 0.003	0.00 ± 0.02
w	$183° \pm 14°$	$0°$
L_1	0.557 ± 0.002	0.534 ± 0.002
L_2	0.443 ± 0.002	0.466 ± 0.002
i	$89°59'$	$89°49'$

spotted areas of most other BY Dra stars, then we can easily explain the low amplitude of the brightness variations as being due to the spot's somewhat asymmetric shape, assuming that the spotted area is almost exactly at the pole.

The change in the phase of maximum brightness by one-third of the orbital period over the 20 years between 1977 and 1996–1997 can be explained as an effect of changes in the shape or asymmetry of a large, long-lived polar spot, i.e., by variations in its longitude distribution.

4.2. Flare Activity of CM Dra

Our analysis revealed four flares in individual nightly light curves of CM Dra in 1997. No flares were detected in 1996. Table 2 presents the data for the detected flares: dates, approximate durations of the flares in minutes, maximum amplitudes (in magnitudes), and maximum relative energies calculated using the formula

$$(E_{star} + E_{flare})/E_{star} = 2.512^{-\Delta m}. \quad (3)$$

We used the following flare-detection criteria, suggested in [6]: (a) a flare should last for several minutes and contain more than one data point, since a peak represented by a single data point could be due to cosmic rays from the direction of the star; (b) the amplitude of the intensity peak should be no less than

0.015^m—flares can be reliably identified only when this condition is fulfilled.

Figure 4 presents the photometric profiles of the detected flares. Their characteristic shape is clear: a rapid brightness increase (the impulsive phase) followed by a gradual decline. This shape is also charasteristic of solar flares, but the energy of the flares detected for CM Dra is much larger. The influence of the flares on the total magnitude of CM Dra was 3–20%, whereas the highest amplitudes for solar flares do not exceed several thousandths of a magnitude. Two flares with different strengths one after another were apparently observed within a single night on April 7, 1997. Table 2 shows that the flare duration and energy are correlated: stronger flares are more prolonged.

We estimated the flare rate, i.e., the mean number of flares per hour, which turned out to be 0.026 flares/h for the entire time interval of our observations. Such a low flare rate is not typical of Population I stars in the Galaxy. Low rates of flare activity were likewise derived for CM Dra in 1977 (0.02 flares/h) [1] and 1994–1996 (0.025 flares/h) [6]. Given the system's high spatial velocity of 163 km/s [1] and its old age, it seems likely that the low flare rate reflects the fact that CM Dra is a Population II star.

On the other hand, the flare rate we observed may be low due to the flare-detection criterion we used (see above); i.e., faint flares could simply have been undetected according to this criterion.

5. THE PHOTOMETRIC AND ABSOLUTE ORBITAL ELEMENTS OF CM Dra

To reliably determine the photometric orbital elements, it is necessary to exclude the outside-eclipse brightness variations. We applied a procedure for excluding the sinusoidal variations similar to that used in the earlier study of CM Dra [1] and the study of the eclipsing spotted system YY Gem [11]. The variations were approximated with a least-squares-fit sine function, and the resulting values, subtracted from the light curve.

We determined the photometric orbital elements using the code of Lavrov based on the method of differential corrections [12, 13]. The derived orbital elements are presented in Table 3, where r_1 and r_2 are the radii of the system's stars in units of the radius of the relative orbit, a ($a = 1$); $k = r_2/r_1$ is the ratio of their relative radii; i is the orbital inclination; w is the longitude of periastron; e is the orbital eccentricity; L_1 and L_2 are the brightnesses of the system's stars in units of the total brightness of the system ($L_1 + L_2 = 1$); and X_1 and X_2 are the limb-darkening coefficients of the stellar disks.
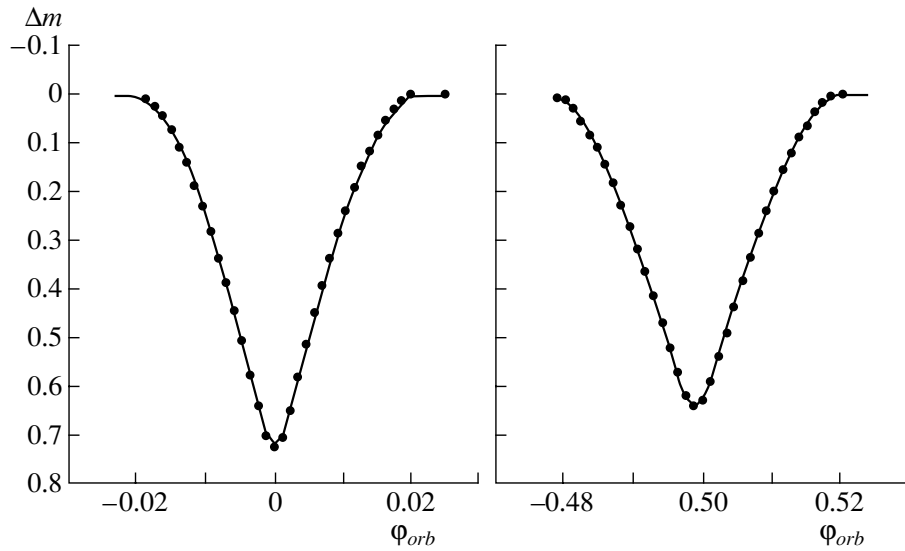
Fig. 5. Theoretical light curve of CM Dra based on the calculated photometric orbital elements (solid curve) and the observed light curve (points). The left panel shows the primary minimum, and the right panel, the secondary minimum.

Figure 5 shows the theoretical light curve (solid) plotted using the calculated photometric orbital elements together with the observed light curve (points). The theoretical and observed light curves are in good agreement.

We used the derived photometric orbital elements to construct a model for the system and determined the degree of filling of the inner critical Roche lobes by the components of CM Dra. The result is presented in Fig. 6. Both components of the system are deep inside their Roche lobes; the degrees of filling are 14.7 and 13.9% for the primary and secondary, respectively, confirming that the system is fully detached.

We determined the absolute orbital elements using the spectroscopic observations of [1], the effective temperatures $T_1 = T_2 = 3150 \pm 100$ K, and the radial-velocity amplitudes $K_1 = 70.0 \pm 2$ km/s and $K_2 = 80.1 \pm 2$ km/s for the primary and secondary, respectively. The resulting absolute orbital elements are given in Table 4, where $M_{1,2}$, $R_{1,2}$, $L_{1,2}$ are the components' masses, radii, and luminosities; A is the semimajor axis of the orbit; and $M_{1,2\,bol}$ are the absolute bolometric magnitudes of the components. Within the uncertainties, our results are in good agreement with the 1977 data.

The main results of this study are the following.

(1) We obtained detailed observations of CM Dra in 1996–1997 (155 h during 43 nights). These observations were acquired with the 70-cm telescope of Ural State University equipped with a two-star photometer. The observations were used to construct a master light curve for CM Dra.

(2) We analyzed the light curve of CM Dra to derive the photometric orbital elements. Within the observational uncertainties, our results are in good agreement with the data obtained by Lacy in 1977 [1], but we find a nonzero orbital eccentricity.

(3) We derived the absolute orbital elements of CM Dra using the spectroscopic results of [1].

(4) We detected low-amplitude (0.024^m) brightness variations of CM Dra in the outside-eclipse part of its light curve. The parameters of these variations (period, amplitude, phase) were determined by finding a least-squares sine-curve fit for the variations. The shape and amplitude of these brightness variations agree with data from the literature and can be explained by the presence of a spot near the pole of one of the components of CM Dra. However, a comparison of our observations (1996–1997) with those for 1977
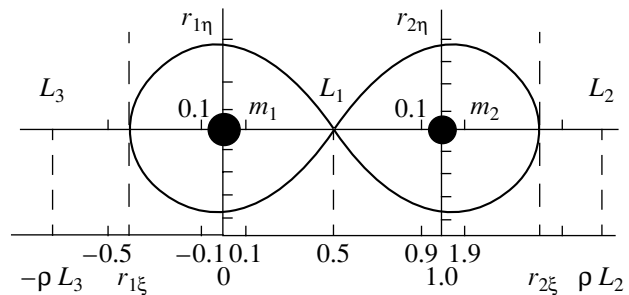

Fig. 6. Degree of filling of the inner critical Roche lobes by the components of CM Dra.

Table 4. Absolute orbital elements for CM Dra

Orbital elements	Our data (1997)	Lacy's data [1] (1977)
$M_1/M_\odot$	0.237 ± 0.017	0.237 ± 0.011
$M_2/M_\odot$	0.207 ± 0.015	0.207 ± 0.008
$R_1/R_\odot$	0.256 ± 0.007	0.252 ± 0.008
$R_2/R_\odot$	0.233 ± 0.007	0.235 ± 0.007
$A/R_\odot$	3.8 ± 0.1	–
$M_{1\,bol}$	10.29 ± 0.20	10.39 ± 0.11
$M_{2\,bol}$	10.56 ± 0.20	10.54 ± 0.11
$L_1/L_\odot$	$(5.8 \pm 0.9)10^{-3}$	$(5.5 \pm 0.8)10^{-3}$
$L_2/L_\odot$	$(4.8 \pm 0.9)10^{-3}$	$(4.8 \pm 0.7)10^{-3}$

revealed a phase shift of the maximum brightness relative to the orbital period. This may provide evidence for changes in the longitude or shape (asymmetry) of a single, long-lived, large spot on one of the components of CM Dra.

(5) We detected four flares with energies two orders of magnitude higher than the energies of solar flares. The derived flare rate agrees with literature data for the CM Dra system but is much lower than is typical for Population I stars of our Galaxy.

ACKNOWLEDGMENTS

REFERENCES

1. C. H. Lacy, Astrophys. J. **218**, 444 (1977).
2. K. G. Strassmeier, D. S. Hall, F. C. Fekel, and M. Scheck, Astron. Astrophys., Suppl. Ser. **100**, 173 (1993).
3. J. P. Caillault, S. Drake, and D. Florkowski, Astron. J. **95**, 887 (1988).
4. R. E. Gershberg, *Solar Activity in the World of Stars* (Znanie, Moscow, 1990), Ser. Kosmonavt. Astron., No. 7, p. 63 [in Russian].
5. L. N. Mavridis, S. I. Avgoloupis, J. H. Seiradakis, and P. P. Varvoglis, Astron. Astrophys. **296**, 705 (1995).
6. H. J. Deeg, L. R. Doyle, V. P. Kozhevnikov, *et al.*, Astron. Astrophys. **338**, 479 (1998).
7. V. P. Kozhevnikov and P. E. Zakharova, *Disks, Planetesimals and Planets,* ASP Conf. Ser. **219**, 381 (2000).
8. L. R. Doyle, H. J. Deeg, V. P. Kozhevnikov, *et al.*, Astrophys. J. **535**, 338 (2000).
9. S. Viti, H. R. A. Jones, A. Schweitzer, *et al.*, Mon. Not. R. Astron. Soc. **291**, 780 (1997).
10. I. Yu. Alekseev, *Spotted Low-Mass Stars* (Astroprint, Odessa, 2001) [in Russian].
11. G. E. Kron, Astrophys. J. **115**, 301 (1952).
12. M. I. Lavrov, *Computer Analysis of the Eclipse Variables of Light Curves* (Kazan, 1980), Part I [in Russian].
13. M. I. Lavrov, Tr. Kazan. Gos. Astron. Obs. **53**, 43 (1993).

Translated by N. Samus'

The Influence of Nonthermal Particles and Radiation on the Charge State of Heavy Ions in Solar Cosmic Rays

Yu. Yu. Kartavykh[1], V. M. Ostryakov[2], E. Möbius[3], and M. A. Popecki[3]

[1]Ioffe Physical Technical Institute, St. Petersburg, Russia

[2]St. Petersburg State Polytechnic University, St. Petersburg, Russia

[3]Center for Space Research, University of New Hampshire, Durham, New Hampshire, USA

Received April 3, 2003; in final form, March 15, 2004

Abstract—The influence of various types of nonthermal electron and proton distributions and photoionization on the charge state of energetic heavy elements moving in a plasma is investigated. The mean charges of Mg, Si, and Fe are calculated for a bi-Maxwellian distribution of the background electrons and for electron and neutral beams with power-law energy distributions. An anomalously high density of the nonthermal component is required to obtain substantial deviations of the equilibrium mean charges of these elements (a few charge units) from the case when they interact with a purely Maxwellian plasma. In this context, the mean charges for O, Ne, Mg, Si, and Fe ions are also calculated for a model with charge-consistent acceleration. The results indicate that photoionization does not significantly influence the charge state of solar cosmic rays if the parameters of the plasma are those characteristic of impulsive solar events. © 2004 MAIK "Nauka/Interperiodica".

1. INTRODUCTION

The numerous available observations of fluxes of particles in interplanetary space due to solar events enable us to divide these events into at least two groups: impulsive and gradual. In impulsive events, the accompanying soft X-ray radiation has a relatively short duration (tens of minutes), while the associated flux of particles is observed over several hours. These events are characterized by an unusual chemical composition compared to the ordinary solar wind and corona: the material is enriched in electrons, ^{3}He (the ^{3}He/^{4}He ratio exceeds the coronal ratio by a factor of 10^3–10^4), and heavy elements (the mean Fe/O $\sim$ 1 exceeds the coronal value by a factor of $\sim$10)[1]. The most recent data obtained with the "Wind" spacecraft show as well an appreciable enrichment in superheavy elements in nine impulsive events. For example, the content of particles with nuclear charges of 34–40 and 50–56 relative to oxygen exceeded the coronal values by factors of about 100 and 1000, respectively [2]. The particle fluxes from gradual events have durations of several days, while the time scale for the fall of the associated X-ray radiation is several hours. These particle fluxes typically have the coronal chemical composition. Data for several gradual events indicate that the Fe/O ratios in the initial phase of the events were close to the values observed in impulsive events, after which this ratio

decreased, until it was even lower than the coronal value at the end of the event [3, 4]. This has been explained as an effect of interaction with turbulence generated by the accelerated protons [5].

The two groups of events are also accompanied by different microwave radiation. As a rule, impulsive events are characterized by type III radio emission, while type II and IV outbursts accompany prolonged flares [6]. Note that type II and IV emission is associated with the propagation of shocks in the solar corona and interplanetary space, respectively. It is usually thought that different mechanisms are responsible for the acceleration of solar cosmic rays in impulsive and gradual events. For example, the particles in impulsive events are generated by second-order Fermi acceleration, or stochastic acceleration. In the case of gradual events, which are associated with the formation of shocks, particles from the solar corona undergo regular acceleration (see, for example, [1]). However, observations of high-energy gamma rays in a number of flares (both impulsive and gradual) led to the conclusion that the fluxes of particles interacting with the solar atmosphere in these events were enriched in heavy elements and ^{3}He and that the particles were accelerated by a single (stochastic) mechanism [7–9]. Therefore, first, it can clearly not be ruled out that the combined action of two types of acceleration is involved and, second, there are differences in the chemical compositions

of the fluxes of particles that are captured and that escape into interplanetary space. However, we will leave these questions outside the framework of our current study, and we will assume that only stochastic acceleration of heavy particles acts in impulsive events and that these particles immediately leave the acceleration region.

In addition to the properties noted above, impulsive and gradual events differ in the mean charges of the heavy-ion cosmic rays that are accelerated in them. For example, the mean charge of Fe in impulsive events is much higher than in gradual events ($\sim$20 as opposed to $\sim$11–14) [10–13]. The same is true of the charges of other elements, such as Si. Further, it is thought that the type III microwave radiation characteristic of impulsive events is generated by energetic (nonthermal) electrons propagating in the solar atmosphere. Miller and Vinas [14] suggested that this nonthermal component of the plasma could lead to additional collisional ionization of the accelerated ions (over that due to the thermal component), right to the observed charge $Q_{Fe} \sim 20$. They also suggested that this could explain the difference in the mean heavy-ion charges of the two classes of solar events. Our current study is concerned with a quantitative study of this question based on the behavior shown by several elements (O, Ne, Mg, Si, and Fe).

The mean charge of an energetic ion moving in a plasma is bounded from above by the so-called equilibrium charge Q_{eq}, which corresponds to the case when the particles are confined in the medium for a fairly extended time [15–17]. A balance is reached between the ionization and recombination of ions with a specified energy E_i. The mean charge of an element subject to acceleration will be between its injected (most likely, thermal) and equilibrium values, since, in reality, such a balance will not be reached due to the comparably small time that the particles are accelerated and confined. In our previous computations, we considered only thermal electrons and protons of the background plasma as the main agents ionizing energetic ions [16–20]. Here, we will also include the nonthermal component and consider its influence on the mean equilibrium charges of Mg and Fe and on the mean charges of O, Ne, Mg, Si, and Fe in a charge-consistent model for the stochastic acceleration. As examples of such nonthermal distributions, we considered a bi-Maxwellian electron distribution and a power-law energy distribution for the electrons and protons.

In [20], we analyzed data for the mean charges of O, Ne, Mg, Si, and Fe for eight impulsive events obtained by the SEPICA instrument on the ACE spacecraft in 1998. For five of these events, the mean charge of O and/or Si could be explained by suggesting that these elements were accelerated in regions with lower temperatures than the regions in which Ne, Mg, and Fe were accelerated. Here, we will investigate whether such temperature differences could arise due to the presence of nonthermal particles in the plasma or due to photoionization, since, together with collisional ionization, the appearance of multiply charged ions could be due to interactions with X-ray radiation (see also [21]). X rays are generated during powerful solar flares, and the hypothesized connection between this radiation and the mean charges of various elements assumes that the X rays could influence the charge states of different elements in different ways.

2. METHOD FOR COMPUTATIONS OF THE ION-CHARGE DISTRIBUTIONS. CASE OF A NONTHERMAL PLASMA

As in our previous studies [16–20, 22], we will include among the processes that can change the charge states of ions ionization (stripping) due to collisions with electrons and heavy particles (protons and, no less important, helium) of the background plasma, as well as dielectronic and radiative recombination. In the general case, the rate of ionization (recombination) $S(R)$, which depends on the ion energy E_i, can be written as

$$S(E_i) = \int_0^\infty v \sigma_{ion} f(\mathbf{v}) d\mathbf{v}_{e,p}, \qquad (1)$$

$$R(E_i) = \int_0^\infty v \sigma_{rec} f(\mathbf{v}) d\mathbf{v}_e,$$

where $v = |\mathbf{v}_i - \mathbf{v}_{e,p}|$ is the relative speed of the interacting particles, $f(\mathbf{v})$ is the distribution function for the relative speeds, and σ_{ion} and σ_{rec} are the corresponding cross sections.

It is traditional to assume that the electrons and ions in the solar corona obey a Maxwellian distribution. However, under certain conditions (for example, in the presence of density and/or temperature gradients), appreciable deviations from such a distribution are possible (see [23] and references therein). Various types of nonthermal (non-Maxwellian) distributions have also been invoked to explain certain observations. For example, if a Maxwellian distribution is assumed for electrons in the corona, the charge composition of the solar wind is not consistent with measurements of coronal temperatures [24]. Therefore, in their computations of the charges of C, O, Mg, Si, and Fe, Esser and Edgar [24] used a two-temperature (bi-Maxwellian) distribution for the coronal electrons, as is often observed in the solar wind near the Earth's orbit. This is a superposition
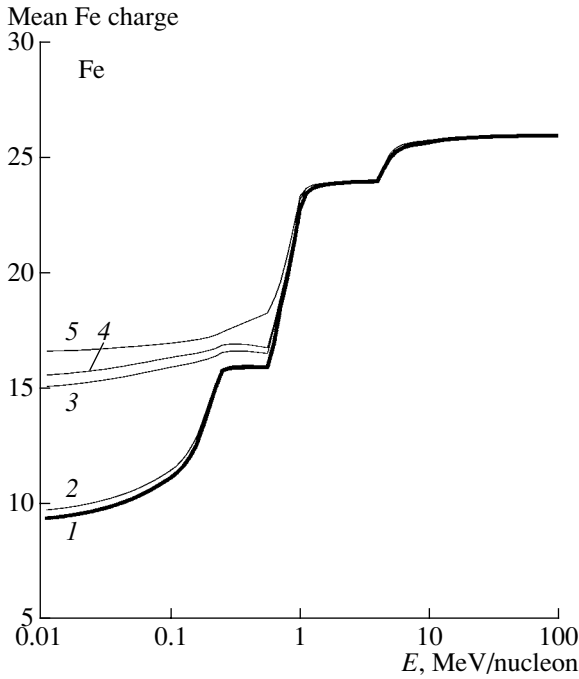
Mean Fe charge

Fig. 1. Mean equilibrium charge of energetic Fe ions for various types of particle distributions for the ambient plasma: (*1*) the electrons and protons have Maxwellian distributions with temperatures $T = T_p = T_e = 10^6$ K; (*2 and 3*) bi-Maxwellian distributions for the background plasma with $T_c = 10^6$ K and $T_h = 2.3 \times 10^7$ K for $\delta = 0.01$ and 0.1, respectively; (*4*) along with the Maxwellian particles ($T = T_p = T_e = 10^6$ K), there are beams of electrons with a power-law energy distribution with $\xi = 5$ and $\delta = 0.1$; (*5*) same as case (*4*) but with a neutral beam (of electrons and protons) with a power-law energy distribution with $\xi = 5$ and $\delta = 0.1$.

Mean Mg charge
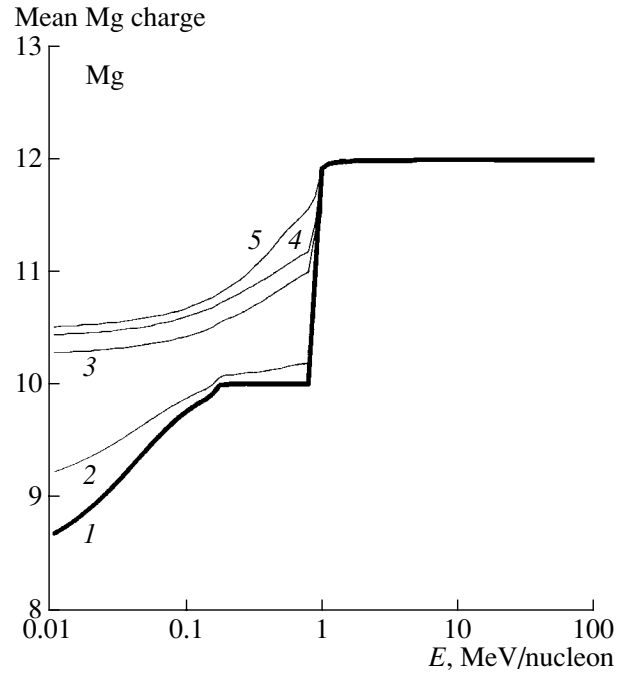
Fig. 2. Same as Fig. 1 for Mg.

of two Maxwellian distributions (so-called core and halo distributions) with different temperatures T_c and T_h and different number densities N_c and N_h, which probably correspond to the longitudinal and transverse directions relative to the regular magnetic field in interplanetary space.

Here, we considered an analogous distribution for the background electrons in the regions where the heavy particles are accelerated on the Sun, where the magnetic field is also fairly strong:

$$f(v) = N_c \left(\frac{m_e}{2\pi T_c} \right)^{3/2} \exp(-m_e v_e^2 / 2T_c) \quad (2)$$
$$+ N_h \left(\frac{m_e}{2\pi T_h} \right)^{3/2} \exp(-m_e v_e^2 / 2T_h).$$

We computed the ionization and recombination rates for this distribution in the rest frame of the moving ions. In this case, we can carry out the integration over the angular variable in (1) explicitly. After this

procedure, the integrand $f(v)$ in (1) will have the form [15]

$$f(v) = N_c \sqrt{\frac{m_e}{2\pi T_c}} \frac{v}{v_i} \left\{ \exp\left(-\frac{m_e}{2T_c}(v - v_i)^2 \right) \right. \quad (3)$$
$$\left. - \exp\left(-\frac{m_e}{2T_c}(v + v_i)^2 \right) \right\} + N_h \sqrt{\frac{m_e}{2\pi T_h}} \frac{v}{v_i}$$
$$\times \left\{ \exp\left(-\frac{m_e}{2T_h}(v - v_i)^2 \right) \right.$$
$$\left. - \exp\left(-\frac{m_e}{2T_h}(v + v_i)^2 \right) \right\}.$$

The subsequent integration over the speed v was done numerically. Thus, the parameters of function (3) are the core and halo temperatures and the ratio of the numbers of electrons in these components ($\delta \equiv N_h / N_c$). The main parameter influencing the computations is δ, while the (halo) temperature is less important. For example, the mean equilibrium charge of Fe for an energy of 0.01 MeV/nucleon and $T_c = 10^6$ K varies from 14.6 to 15.4 as T_h varies from 1.5 to 5 keV and is equal to 14.7 for a halo temperature of 10 keV. Figures 1 and 2 show the mean equilibrium charges for Fe and Mg ions for $\delta = 0.01$ and $\delta = 0.1$ for the case when the core and halo electron temperatures are 10^6 K and 2.3×10^7 K (~ 2 keV), respectively. The case of a purely Maxwellian distribution ($\delta = 0$) is also shown for comparison. We can see that the variations in Q_{eq} are substantial only if the number density of halo electrons is rather high

Equilibrium charge of Fe for an energy of 0.01 MeV/nucleon and $\delta = 0.2$ as a function of the lower energy limit E_0 (for $\xi = 4$) and the power-law index ξ (for $E_0 = 2$ keV)

$\xi = 4$	
E_0, keV	Q_{eq}
1.5	16.9
2.0	17.0
3.0	16.9
$E_0 = 2$ keV	
ξ	Q_{eq}
3.5	15.1
5.0	17.5
7.0	17.5

($\delta \sim 0.1$). As another example of a nonthermal distribution, we considered a power-law energy distribution for the electrons in the regions of the solar flares (see, for example, [25]). It is usually believed that such electrons generate the type III microwave outbursts and/or hard X-ray outbursts that are characteristic of impulsive solar cosmic-ray events. By analogy with (2), we have

$$f(v) = N_0 \left(\frac{m_e}{2\pi T}\right)^{3/2} \exp(-m_e v_e^2 / 2T) \qquad (4)$$
$$+ N_b \frac{\xi - 3}{2\pi v_0^3} \left(\frac{v_0}{v_e}\right)^\xi F(\mu),$$

where $F(\mu)$ is the pitch-angle distribution of the electrons ($\int F(\mu) d\mu = 1$) and the power-law index ξ should be greater than three due to the normalization condition. It is assumed that the second term in (4) is present only in the speed interval $v_0 \leq v_e \leq \sqrt{10} v_0$, which thereby limits the region of considered energies for the nonthermal particles to only one order of magnitude (v_0, corresponds to the lower energy limit E_0). With regard to the function $F(\mu)$, the characteristic time for the isotropization of the beam electrons is much shorter than the characteristic time for the acceleration of ions in impulsive events. Consequently, we can assume in a first approximation that these electrons are distributed isotropically, i.e., $F(\mu) = 0.5$. Note also that, in the case of a power-law distribution for the electrons, the integration over the relative speed of the colliding particles in (1) was carried out numerically, over both the angular and energy variables. Thus, the parameters of this distribution will be the power-law index ξ, the energy limits for the spectrum of nonthermal electrons, and the number

density of these electrons relative to the Maxwellian particles $\delta = N_b/N_0$.

Since there are quite a few input parameters, it is necessary to investigate their influence on the results obtained. First and foremost, we computed the equilibrium charges of Fe and Mg for various energy ranges for the electron beam: 1.5–15, 2–20, and 3–30 keV. In the first and last of these versions, the equilibrium charge was only slightly lower than when the electron beam had energies in the range 2–20 keV, which we chose as our basis case for the subsequent computations. We also considered the influence of variations in the power-law index ξ in (4). The results of these tests are presented in the table. Overall, we can conclude that the variations in Q_{eq} were not significant; they become more appreciable compared to the Maxwellian case when $\xi = 5$ or 6. Figures 1 and 2 show the equilibrium Fe and Mg charges obtained using the above method for several values of δ with $\xi = 5$ and energy limits of 2 and 20 keV. We can see that, as in the case of the bi-Maxwellian distribution, the influence of the nonthermal electrons becomes important when their number density is high ($\delta \sim 0.1-0.2$). Note also that the addition of a nonthermal power-law beam with fixed δ exerts a stronger influence on the increase of the equilibrium charge and, consequently, also on the charges of accelerated heavy ions.

Porquet *et al.* [26] present the mean equilibrium charges for heavy ions at rest for a hybrid distribution of the thermal electrons, which have a Maxwellian distribution below a certain threshold energy E_b and a power-law energy distribution above E_b, $f(E) \propto E^{-\alpha}$, where $\alpha = (\xi - 1)/2$. As a check of our computations, we compared our results for the mean equilibrium Fe charge for energies $E_i \sim 10$ keV/nucleon with the results of [26]. We selected from [26] results for $\alpha = 1.5$ ($\xi = 4$) and $x_b = E_b/kT = \alpha + 1/2$ for the two temperatures $T = 10^6$ K and $T = 10^7$ K. Note that the threshold value x_b determines the ratio of the numbers of nonthermal to thermal electrons for a specified α ($\delta \approx 1.2$, $\alpha = 1.5$, and $x_b = 2$). In both cases, we obtained good agreement between the mean iron charges (~ 17.2 and ~ 23.8 for $T = 10^6$ K and $T = 10^7$ K, respectively).

Above, we considered various types of nonthermal electron distributions and their influence on the mean charges of heavy elements moving in the plasma. However, in some solar flares, the contribution of nonthermal protons to the generation of the hard (>20 keV) X-ray radiation can be substantial (the so-called neutral-beam case). For example, protons contributed $\geq 60\%$ as much radiative energy as electrons in two of the 19 flares analyzed in [27]. Since the role of thermal protons in ionizing energetic
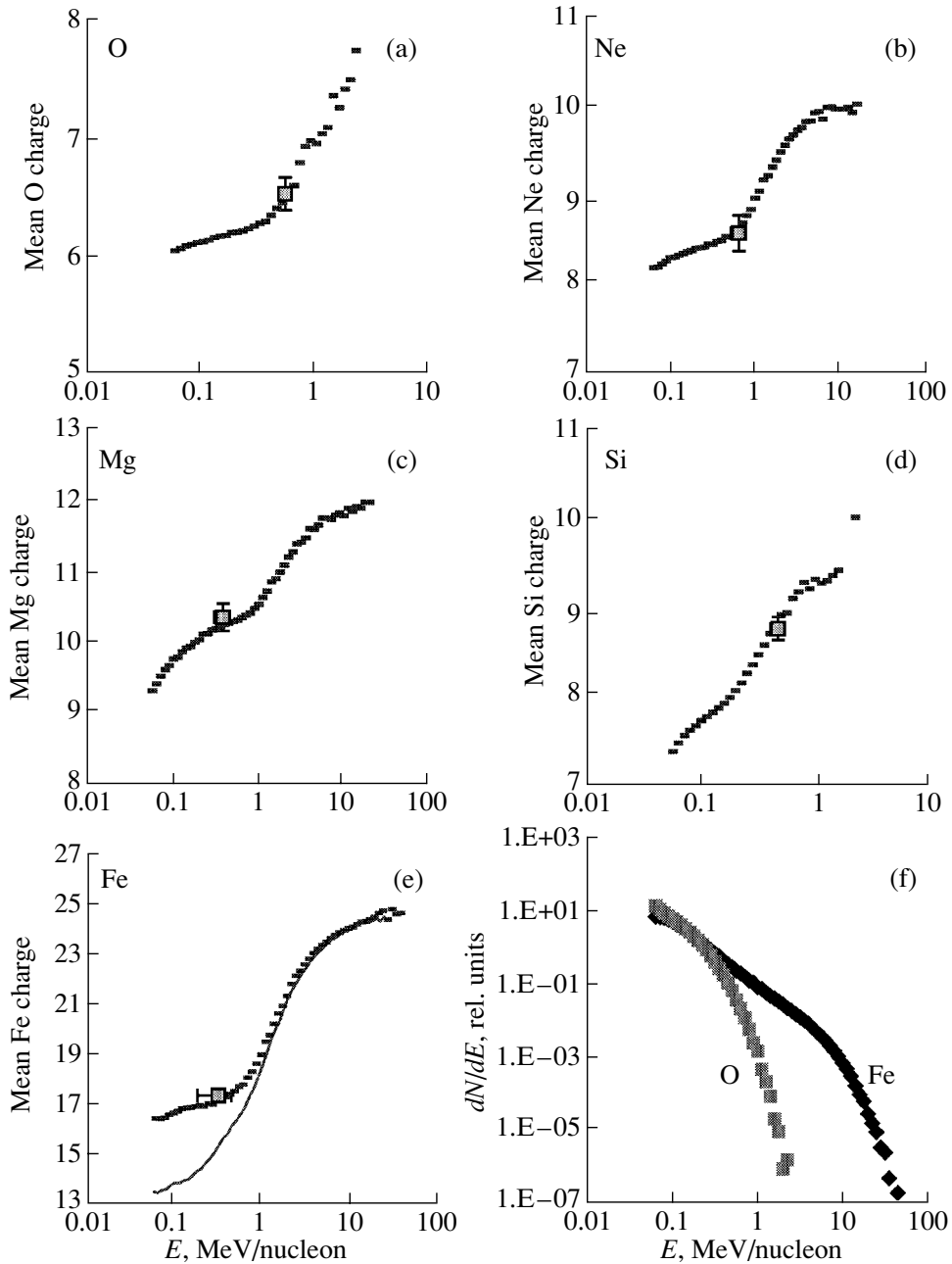
Fig. 3. Computed mean charges of various ions in the model with charge-consistent acceleration (dashes). The squares show observational measurements of the mean charges of (a) O, (b) Ne, (c) Mg, (d) Si, and (e) Fe for event 11 using the ACE/SEPICA instrument (according to the measurements presented in [13]). (f) Energy spectra for O and Fe summed over charge (see text for the computational parameters). The solid curve in plot (e) shows the behavior of the mean Fe charge in the absence of a neutral beam in the acceleration region for the same values of the remaining parameters.

(≥ 1 MeV/nucleon) heavy ions is dominant [16, 17], the contribution of energetic (nonthermal) protons should also be included when computing the ionization states of the accelerated particles. Therefore, we also considered the case when we have a beam containing both energetic electrons and energetic protons, specifying $\delta_e = \delta_p = \delta$ to ensure electrical neutrality of the beam. According to [27], the energy

of monoenergetic electrons in such a beam comprises ~ 0.018 of the corresponding energy of the protons. Consequently, a beam of electrons with energies from 2 to 20 keV will correspond to a beam of protons with energies from 0.11 to 1.1 MeV. Figures 1 and 2 (curve 5) show the equilibrium charges for energetic Mg and Fe ions when a neutral plasma beam is present ($\delta = 0.1$). We can see that the largest vari-

ations in the mean equilibrium charges of these elements compared to the case of a pure electron beam are attained for energies $E_i \sim 0.56$ MeV/nucleon (the difference is about $1-2$ charge units for Fe), with the difference for $E_i \sim 0.3$ MeV/nucleon (the mean energy of iron ions detected by ACE/SEPICA) being a factor of 2 lower.

To investigate the influence of the nonthermal distributions on the charges of the accelerated ions, we chose values of ξ and E_0 for the neutral beam that led to the largest variations in the equilibrium charge (an "optimistic" evaluation of the effect; see the table). Namely, we took the beam electrons to have energies of $2-20$ keV, the beam protons to have energies of $0.11-1.1$ MeV, the power-law index to be $\xi = 5$ for both the electrons and protons, and the relative number density to be $\delta = 0.2$. Assuming the presence of nonthermal particles with these parameters in the plasma, we then computed the charges and energy characteristics of accelerated O, Ne, Mg, Si, and Fe ions. We tried to explain the temperature discrepancy obtained by applying a uniform model to event 11 in [20] by varying the main parameters of our acceleration model (see below). Note also that this event (like impulsive events $12-18$, probably with the exception of event 16) was accompanied by type III and/or V radio emission, which indicates the presence of particle beams in the flare plasma.

Figure 3 presents the computed mean charges for the accelerated O, Ne, Mg, Si, and Fe for the case with $T = 10^6$ K and $\tau_a^0 N = 10^{11}$ s/cm^3, and with a neutral beam with the parameters presented above present in the plasma. Here, τ_a^0 is the characteristic time to accelerate the protons to energies of 1 MeV/nucleon; for ions, $\tau_a(Q, E) = \tau_a^0 (Q/A)^{S-2} E_i^{(3-S)/2}$, where $S = 1.5$ is the power-law index for the Alfven-wave turbulence and A is the atomic number of the specified element. The acceleration model corresponded to the charge-consistent approach for stochastic acceleration developed by us earlier (see, for example, [17, 20, 22]). The mean charges of the accelerated ions obtained with our computations were in agreement with the observational values for energies $E_i \sim 0.18-0.75$ MeV/nucleon for the ratios $\tau_a^0/\tau_d^0 = 0.1$ for Fe, 0.3 for Mg, 1.0 for Ne, 11.0 for O, and 15.0 for Si. Here, τ_d^0 is the characteristic time for diffusion (or confinement) of protons with energies of 1 MeV/nucleon in the acceleration region; for ions, $\tau_d(Q, E) = \tau_d^0 (Q/A)^{2-S} E_i^{(S-3)/2}$. It is difficult to explain such differences in the efficiency of accelerating different elements using only the differences in their Larmor radii, since the ratio of the Larmor radii for Fe^{+9} and O^{+6} is 2.3, while the ratios τ_a^0/τ_d^0 for these

ions differ by a factor of 100. Figure 3f presents the energy spectra of the oxygen and iron ions summed over all charge states. The resulting hardness of the Fe spectrum compared to the O spectrum is in qualitative agreement with ACE/ULEIS spectra for the impulsive events of March 21, 1999, and September 27, 2000 [28]. Figure 3e also shows for comparison the computed mean Fe charges for the same plasma parameters for the case when there is no neutral beam in the acceleration region. We can see that the influence of the beam is stronger at low plasma energies (<1 MeV/nucleon), which is where most direct charge measurements obtained with spacecraft are made. The observed mean Fe charge (Fig. 3e) can be explained in the case when only a thermal plasma is present in the acceleration region only with a much higher temperature [20].

Another way to explain the observed mean charges of various elements for the case of a fixed temperature of the acceleration region would be to suppose the presence of plasma density inhomogeneities (different values for the parameter $\tau_a^0 N$), with the ratio τ_a^0/τ_d^0 being the same for different elements (see below).

3. INFLUENCE OF PHOTOIONIZATION ON THE MEAN CHARGES OF HEAVY IONS

According to the estimates of [21], the photoionization of ions by the soft X-ray radiation generated in solar flares could also influence the charge composition of solar cosmic rays. This additional process leading to the loss of an electron by an ion should lead to an increase in the effective temperature in the acceleration region. We investigated this effect, making use of the more complete data on the photoionization cross sections of various elements and their ions presented in [29]. Let us write the full cross section for the loss of an electron due to the interaction of an ion with a photon with energy ε (in eV) as the sum of partial cross sections:

$$\sigma_{ph}(\varepsilon) = \sum_{nl} \sigma_{nl}(\varepsilon), \quad \sigma_{nl}(\varepsilon) = 0, \qquad (5)$$

$$\varepsilon < E_{th} \text{ and } \sigma_{nl} = \sigma_0 F(y), \varepsilon > E_{th}.$$

Here, E_{th} is the ionization threshold, below which a photon's energy is not sufficient to strip off the corresponding electron; σ_0 is a constant; and the function $F(y)$ can be written as

$$F(y) = \left[(y-1)^2 + y_w^2 \right] y^{-Q} (1 + \sqrt{y/y_a})^{-P}, \quad (6)$$

where $y = \varepsilon/E_0$ and $Q = 5.5 + l - 0.5P$, and E_0, y_w, y_a, and P are parameters [29] (do not confuse the parameters Q and l with the similarly denoted
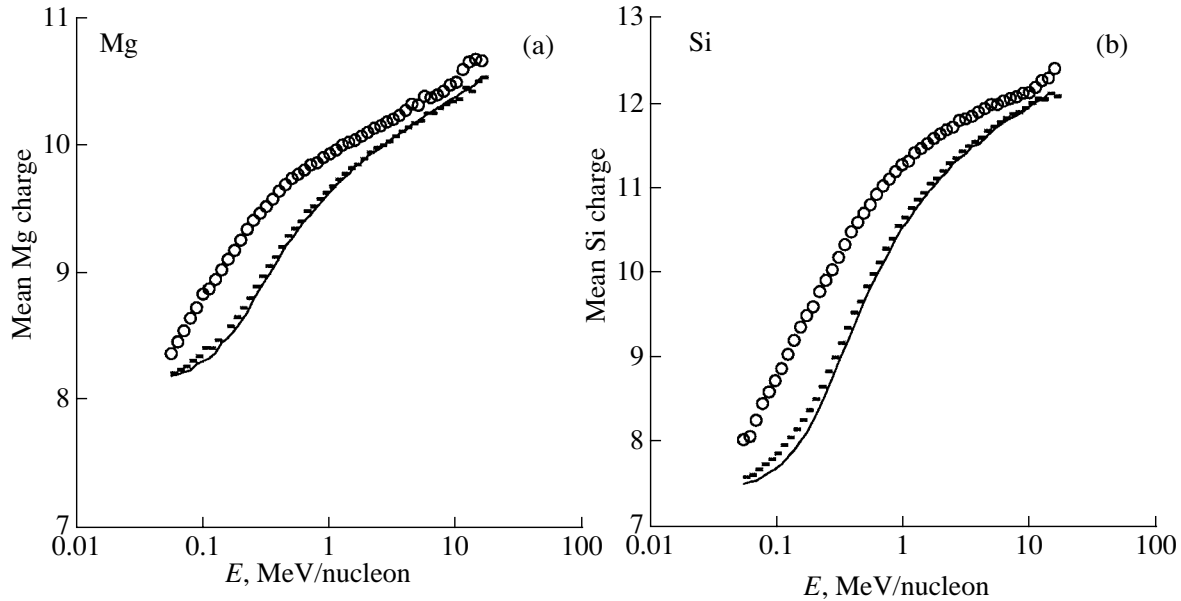
Fig. 4. Mean charges of accelerated (a) Mg and (b) Si ions computed taking into account photoionization (see text) for the characteristic acceleration times $\tau_a^0 = 10$ and 100 s (dashes and circles, respectively) and $T = 10^6$ K, $\tau_a^0 N = 10^{10}$ s cm^{-3}, $\tau_a^0/\tau_d^0 = 0.5$. The solid curves show the results for the computations without photoionization.

parameters described in the main text of the paper). We can now write the photoionization rate:

$$S_{ph} = \int (I/\varepsilon)\sigma_{ph}(\varepsilon)\, d\varepsilon, \qquad (7)$$

where I is the flux of soft X-ray radiation in the region where the particles are accelerated. Analysis of the observational data of [30] shows that, for most of the studied flares, this flux has the power-law form $I \sim \varepsilon^{-\gamma}$ with $\gamma > 2$. Here, we chose $\gamma = 2$, although, as we will see below, even in this most optimistic case, the influence of photoionization is not very significant.

As a rule, impulsive solar cosmic-ray events are associated with weak soft X-ray flares, which occur in very compact regions. In our estimates, we assumed that the acceleration in a flare of class M1.0 (an X-ray flux of 10^{-2} erg cm^{-2} s^{-1} at 1–8 Å) occurred in a region with a size of $L \sim 10^8$ cm. Given the distance from the Earth to the Sun, this corresponds to a flux through the surface bounding the active region of about 2.3×10^8 erg cm^{-2} s^{-1}. We find from formula (7) that, in this case, the characteristic time for the photoionization of heavy elements varies from thousandths of a second for neutrals and ions with low charges to hundreds or even thousands of seconds for multiply charged ions.

We also checked whether photoionization could explain the differences in the corresponding plasma temperatures obtained in [20] based on an analysis of the mean charges of various elements. In particular,

the mean charges of O and Si for event 11 provided evidence for temperatures in the acceleration regions for these ions of $T \sim (1{-}2) \times 10^6$ K, while the data for Fe indicated a temperature of $T \sim 6 \times 10^6$ K. Let us suppose that the ions could be subject to ionization by X-ray radiation during the entire time they are in the acceleration region. In this case, turning on photoionization in the acceleration model leads to the appearance of a new parameter: the ratio of the characteristic photoionization time for a specific ion to the time for this ion to accumulate energy (or the time it is confined in the plasma). An important assumption in this scenario is that the flux of X-ray radiation remains constant (and equal to its maximum value) during the entire particle acceleration (and confinement) process.

If there were an increase in the effective temperatures for Mg, Ne, and Fe due to photoionization, we would expect the charges of these ions to be appreciably higher compared to the charges of O and Si, other acceleration parameters being the same. Figure 4 presents the results of computations of the mean charges for the accelerated Si and Mg ions for the characteristic acceleration times $\tau_a^0 = 10$ and 100 s. We can see that it is not possible to obtain a dominant increase in the Mg charge compared to the Si charge due to photoionization. For $\tau_a^0 = 1$ s (which is close to the properties of impulsive solar cosmic-ray events), there are essentially no variations in the mean charges of these elements compared to the

case without photoionization. Note that increasing the power of the flare by an order of magnitude (i.e., by one magnitude) influences the mean charges of the accelerated ions to the same extent as does increasing the acceleration time by the same amount.

4. INTERPRETATION OF THE RESULTS AND DISCUSSION

As was shown above (Figs. 1 and 2), the equilibrium charges of ions moving through a nonthermal plasma differ appreciably from the purely Maxwellian case only if a large number of nonthermal particles are added ($\delta = 0.1-0.2$). With a sufficiently high number density of nonthermal particles, it is possible to explain the temperature discrepancy found earlier using a uniform model for the acceleration of ions in solar flares if we assume different elements have different ratios τ_a^0/τ_d^0 (Fig. 3).

Let us now discuss what conditions could lead to different values of τ_a^0/τ_d^0 for different elements, with $\tau_a^0 N$ fixed. This is equivalent to fixing the value of τ_a^0/τ_c^0, where τ_c^0 is the characteristic time for Coulomb energy losses for protons with energies of 1 MeV/nucleon; for ions, $\tau_c(Q, E) = \tau_c^0 A/(Q^2 N) \times E_i^{3/2}$. According to the analysis of [31], the condition that τ_a^0/τ_c^0 be fixed can be rewritten in terms of the parameters of the acceleration region:

$$\frac{\tau_a^0}{\tau_c^0} \propto \tau_a^0 N \propto \frac{N^2\sqrt{l}}{B^{5/2}(W_T/W_B)} \qquad (8)$$
$$= \frac{N^2\sqrt{l}}{B^{1/2}W_T} = \text{const},$$

where $W_B = B^2/8\pi$ and W_T are the energy densities of the regular magnetic field and the turbulence associated with Alfven waves, B is the magnetic-field strength, and l is the main scale for the turbulence ($l \sim 0.1 \times L$, where L is the geometric size of the acceleration region). Let us now write τ_a^0/τ_d^0 in terms of the physical parameters of the plasma in the acceleration region and separate out the constant ratio τ_a^0/τ_c^0, writing it in the form of a coefficient:

$$\tau_a^0/\tau_d^0 \propto \frac{Nl}{L^2B^2(W_T/W_B)^2} \qquad (9)$$
$$= \frac{N^2\sqrt{l}}{B^{5/2}(W_T/W_B)} \frac{\sqrt{l}B^{1/2}}{N(W_T/W_B)L^2}$$
$$= \left(\frac{N^2\sqrt{l}}{B^{1/2}W_T}\right) \frac{\sqrt{l}B^{5/2}}{NW_TL^2}.$$

Variations in τ_a^0/τ_d^0 are now due only to variations in the second factor. Let us consider several possible

scenarios leading to variations in τ_a^0/τ_d^0. Let us suppose that the turbulent energy density W_T grows by a factor of 9. Then, according to (8), to compensate for this variation, for example, by increasing the number density, the value of N must be increased by a factor of 3. Overall, this leads to a decrease in τ_a^0/τ_d^0 by about a factor of 30. Explaining the mean charges of heavy ions for the ACE/SEPICA data [13] based on these estimates would suggest that the Fe ions were accelerated in a denser (by about a factor of 5) arc with a higher degree of turbulence (by a factor of 20−30) than were the O and Si ions. Alternatively, the different ions were accelerated in a single arc, but at different times, with the plasma parameters varying during the evolution of the solar flare in the corresponding way. Finally, these data can also be explained by supposing variations in the strength of the regular magnetic field (say, an increase by a factor of 4), which is compensated by variations in the turbulent energy (a decrease by a factor of 2) or by a corresponding (factor of 1.4) increase in the plasma density in the acceleration region. In the first case, this leads to an increase in τ_a^0/τ_d^0 by a factor of ~ 60, and in the second case, to an increase by a factor of ~ 20. Applied to the acceleration of the ions considered, this would mean that O and Si were accelerated, for example, in regions with a stronger magnetic field and a lower level of turbulence than were Fe, Mg, and Ne.

Let us now consider another possible explanation for the observed ion charges by fitting the parameter $\tau_a^0 N$ for various elements while fixing the ratio τ_a^0/τ_d^0. According to [31], by analogy with the previous case, it is not difficult to obtain

$$\tau_a^0 N \propto \left(\frac{B\sqrt{N}\sqrt{l}}{LW_T}\right)\frac{N^{3/2}L}{B^{3/2}}, \qquad (10)$$

where the factor in parentheses, which is proportional to the square root of τ_a^0/τ_d^0, remains constant. Then, if $\tau_a^0 N$ is increased by a factor of 10 (due, for example, to an increase in the number density N by a factor of ~ 4.7), this should be compensated by an increase in the size of the arc L or the turbulent energy density by a factor of ~ 2.2 (see (10)). In this case, again, the Fe, Mg, and Ne ions should be accelerated in larger regions (or in regions with a higher turbulent energy density) than are the O and Si ions. On the contrary, if a growth in $\tau_a^0 N$ is due to a decrease in the magnetic-field strength, then a corresponding decrease in the size of the arc and/or the turbulent energy density is required to enable τ_a^0/τ_d^0 to remain constant. In this case, the characteristics of the acceleration regions for the Ne, Mg, and Fe ions, on the one hand, and the O and Si ions, on the other hand, would show trends

opposite to those presented above. This is probably less likely, given that impulsive events are enriched in Fe [32].

A joint analysis of the mean charges together with the energy spectra of elements or the energy dependences of the mean charges could probably shed light on this situation by decreasing the number of free parameters in our model in a natural way. However, due to the weakness of the particle fluxes from impulsive events, the accumulation of data on the charge states of heavy ions participating in these events is only beginning [12, 13], although some results concerning the energy spectra (averaged over several events) have been known since the beginning of the 1980s [33].

Let us now discuss the possible presence of a nonthermal population of particles in the solar-flare plasma in the required quantities indicated by our analysis ($\delta \sim 0.2$). There have been many studies in which the densities of energetic electrons have been determined based on observations of microwave and/or hard X-ray outbursts. For example, in his analysis of data on type III radio outbursts, Huang [34] estimated that the ratio of nonthermal to thermal electrons was of the order of 10^{-6}; an even lower value was obtained in [35]. Somewhat more "optimistic" conclusions follow from analyses of the hard X-ray radiation of solar flares. For example, a total number of energetic electrons of $4 \times 10^{28} - 10^{33}$ was obtained in [36]. If we suppose that impulsive events occur in regions with characteristic volumes $10^{26} - 10^{27}$ cm^3 [37] and particle densities of 10^9 cm^{-3} or higher, we find the ratio of the numbers of energetic and thermal electrons turns out to be in the range $10^{-7} - 10^{-2}$. Thus, the maximum value for δ is ~ 0.01. The value for this ratio derived from radio data and Yohkoh soft X-ray data for 44 flares is $0.005 - 0.127$ [38].

Miller and Vinas [14] estimated the number density of nonthermal electrons based on the observed fluxes of ^{3}He in solar cosmic rays. They supposed that the energetic electrons are generated by turbulence, on which the ^{3}He is also subsequently accelerated. As in our work, the value of δ they obtained was very high—of the order of 0.1. This value of δ unavoidably requires more complex (nonlinear) descriptions of the acceleration and propagation of particles in the flare plasma, since the quasi-linear approximations that are usually used to describe the dynamics of the electrons and the turbulence are evidently no longer valid in this case.

5. CONCLUSIONS

Thus, the presence of a nonthermal population of ionizing particles in the flare region cannot explain the observed mean charges for various elements if they are simultaneously accelerated in a homogeneous plasma. It is probably more reasonable to suppose that different elements are accelerated in different parts of arcs (or in different arcs), or in an arc whose parameters (the regular magnetic field, turbulent energy density, etc.) evolve during the flare process. Note that a similar tendency was found for gradual events (in the framework of thermal models for the formation of the mean charges of various elements): based on data for Si and O, the characteristic temperature of these ions was somewhat lower than the Ne and Mg temperature [10].

Therefore, inhomogeneity and/or nonstationarity of the regions with regard to the acceleration of different ions is probably a general characteristic of the generation of heavy particles on the Sun. We hope that information about the energy dependence of the mean charges of ions accelerated in impulsive events, as well as data on the locations of regions of particle acceleration, will help elucidate these questions. The data currently being obtained by the HESSI spacecraft are very important [39] in this connection.

Photoionization probably does not significantly influence the charge states of heavy ions accelerated in impulsive events. The same conclusion was drawn in connection with the role of photoionization in most gradual events [40].

ACKNOWLEDGMENTS

REFERENCES

1. D. V. Reames, Adv. Space Res. **13**, 331 (1993).
2. D. V. Reames, Astrophys. J. **540**, L111 (2000).
3. A. J. Tylka, D. V. Reames, and C. K. Ng, Geophys. Res. Lett. **26**, 2141 (1999).
4. A. J. Tylka, J. Geophys. Res. **106**, 25333 (2001).
5. C. K. Ng, D. V. Reames, and A. J. Tylka, Geophys. Res. Lett. **26**, 2145 (1999).
6. H. V. Cane and D. V. Reames, Astrophys. J. **325**, 895 (1988).
7. R. Ramaty and N. Mandzhavidze, in *High Energy Solar Physics*, Ed. by R. Ramaty *et al.*, AIP Conf. Proc. **374**, 533 (1996).
8. N. Mandzhavidze, R. Ramaty, and B. Kozlovsky, Astrophys. J. **518**, 918 (1999).
9. N. Mandzhavidze and R. Ramaty, in *High Energy Solar Physics: Anticipating HESSI*, Ed. by R. Ramaty and N. Mandzhavidze, ASP Conf. Ser. **206**, 64 (2000).

10. A. Luhn, D. Hovestadt, B. Klecker, *et al.*, in *Proceedings of the 19th ICRC,* La Jolla, 1985, Vol. 4, p. 241.

11. A. Luhn B. Klecker, D. Hovestadt, and E. Mobius, Astrophys. J. **317**, 951 (1987).

12. E. Möbius, B. Klecker, M. A. Popecki, *et al.*, in *Proceedings of the 26th ICRC*, *Salt Lake City*, *1999*, Talk SH 1.4.07.

13. E. Möbius, B. Klecker, M. A. Popecki, *et al.*, in *Acceleration and Transport of Energetic Particles Observed in the Heliosphere,* Ed. by R. A. Mewaldt, J. R. Jokipii, M. A. Lee, E. Möbius, and T. H. Zurbuchen, AIP Conf. Proc. **528**, 131 (2000).

14. J. A. Miller and A. F. Vinas, Astrophys. J. **412**, 386 (1993).

15. A. Luhn and D. Hovestadt, Astrophys. J. **317**, 852 (1987).

16. L. Kocharov, G. A. Kovaltsov, J. Torsti, and V. M. Ostryakov, Astron. Astrophys. **357**, 716 (2000).

17. V. M. Ostryakov, Yu. Yu. Kartavykh, D. Ruffolo, *et al.*, J. Geophys. Res. **105** (A12), 27315 (2000).

18. A. A. Kharchenko and V. M. Ostryakov, in *Proceedings of the 20th ICRC*, *Moscow*, *1987*, Vol. 3, p. 248.

19. Yu. Yu. Kartavykh, V. M. Ostryakov, I. Yu. Stepanov, and M. Ioshimori, Kosmich. Issled. **36** (5), 465 (1998).

20. Yu. Yu. Kartavykh, V. M. Ostryakov, E. Möbius, *et al.*, in *Proceedings of the 27th ICRC*, *Hamburg*, *2001*, Vol. 8, p. 3091.

21. D. J. Mullan and W. L. Waldron, Astrophys. J. **308**, L21 (1986).

22. Yu. Yu. Kartavykh, S. Wannawichian, D. Ruffolo, and V. M. Ostryakov, Adv. Space Res. **30**, 119 (2002).

23. E. Dzifcakova, Solar Phys. **140**, 247 (1992).

24. R. Esser and R. J. Edgar, Astrophys. J. **532**, L71 (2000).

25. G. D. Fleishman and S. G. Yastrebov, Astron. Zh. **71**, 531 (1994) [Astron. Rep. **38**, 468 (1994)].

26. D. Porquet, M. Arnaud, and A. Decourchelle, Astron. Astrophys. **373**, 1110 (2001).

27. J. C. Brown, M. Karlicky, N. Mandzhavidze, and R. Ramaty, Astrophys. J. **541**, 1104 (2000).

28. G. M. Mason, M. E. Wiedenbeck, J. A. Miller, *et al.*, Astrophys. J. **574**, 1039 (2002).

29. D. A. Verner and D. G. Yakovlev, Astron. Astrophys., Suppl. Ser. **109**, 125 (1995).

30. Yu. E. Charikov and P. B. Dmitriev, in *Nuclear Astrophysics*, Ed. by G. E. Kocharova (FTI im. Ioffe Ross. Akad. Nauk, St. Petersburg, 1991), p. 146.

31. V. M. Ustryakov, Yu. Yu. Kartavykh, and G. A. Koval'tsov, Pis'ma Astron. Zh. **26** (2), 153 (2000) [Astron. Lett. **26**, 122 (2000)].

32. J. A. Miller and D. V. Reames, in *Proceedings of the 25th ICRC*, *Durban*, *1997*, Vol. 1, p. 141.

33. E. Möbius, M. Scholer, D. Hovestadt, *et al.*, Astrophys. J. **259**, 397 (1982).

34. G.-L. Huang, Astrophys. J. **498**, 877 (1998).

35. V. N. Mel'nik, V. Lapshin, and E. Kontar, Solar Phys. **84**, 353 (1999).

36. M. J. Aschwanden, A. O. Benz, B. R. Dennis, and R. A. Schwarz, Astrophys. J. **455**, 347 (1995).

37. R. Pallavichini, S. Serio, and G. S. Vaiana, Astrophys. J. **216**, 108 (1977).

38. M. J. Aschwanden and A. O. Benz, Astrophys. J. **480**, 825 (1997).

39. R. P. Lin *et al.* (The RHESSI Team), in *Proceedings of the 10th European Solar Physics Meeting on Solar Variability: From Core to Outer Frontiers* (Prague, 2002), Vol. 2, p. 1035.

40. V. M. Ostryakov and M. F. Stovpyuk, in *Proceedings of the 10th European Solar Physics Meeting on Solar Variability: From Core to Outer Frontiers* (Prague, 2002), Vol. 1, p. 355.

Translated by D. Gabuzda

Fine Structure of Convective Motions in the Solar Photosphere: Observations and Theory

R. I. Kostyk and N. G. Shchukina

Main Astronomical Observatory, Kiev, Ukraine
Received November 18, 2003; in final form, January 9, 2004

Abstract—The granulation brightnesses and convective velocities in the solar photosphere between the levels of formation of the continuum radiation and the temperature minimum are examined. Spectral images of the granulation observed in lines of neutral and ionized iron with high spatial ($0.5''$) and temporal (9 s) resolutions were obtained using the German Vacuum Tower Telescope in Izana (Tenerife, Spain). A correlation analysis shows that the granules and intergranules change their relative brightness at a height near 250 km, and a general reversal of the velocity occurs near a height of 490 km, where the material above granules begins to predominantly descend, and the material above intergranules, to ascend. The maximum correlation coefficient between the velocity and the line brightness does not exceed 0.75. The properties of the brightness and velocity are analyzed in a sixteen-column model. Four sorts of motions are most typical and efficient. In the first two, only the sign of the relative contrast of the material changes (an efficiency of 46%). This occurs, on average, at a height of 270 km. In the last two motions, both the sign of the contrast and the direction of the motion are reversed near a height of 350 km (an efficiency of 28%). All the observed dependences are compared with theoretical relations obtained in a three-dimensional hydrodynamical model, with deviations from local thermodynamic equilibrium included in the calculation of the spectral-line profiles. This model can satisfactorily reproduce all the basic features of the convective velocities and intensities. It is concluded that the convective motions maintain their column structure throughout the photosphere, right to the level of the temperature minimum. This makes a separation of the photosphere into two regions with different granulation brightnesses and convective motions unjustified.
© *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In 1908, the Pulkovo Observatory astronomer Hansky [1] presented the first quantitative studies of the solar granulation, including the lifetimes, sizes, and horizontal velocities of the granules. The theoretical work of Unsold [2] suggested that a convective zone located directly below the photosphere was the origin of the granulation, and the granulation pattern has been identified with turbulent convection. Since 1949, when Richardson and Schwarzschild [3] obtained the first granulation spectrogram, the velocities of the granular and intergranular motions directed along the line of sight have been determined regularly. The material in granules is predominantly ascending. However, in their analysis of granulation spectrograms, Krat and Shpital'naya [4] confidently detected granules in which comparatively hot material was moving downward [4]. This finding contradicted all commonly accepted theories for the granulation. Karpinsky [5, 6] was the first to construct a morphological model for the granulation brightness using extensive observational data with very high resolution obtained beyond the Earth's atmosphere. He showed that the brightness and velocity structures

coincide only to a height of 100 km above the level of formation of the continuum (which we will call the continuum height), while velocities in the form of vertical cylindrical columns are maintained to a height $H \approx 250$ km. This column structure is sharply disrupted above these heights.

In subsequent years, numerous researchers have studied the intensities and velocities of motions in the photosphere, using various instruments and applying various data-processing techniques and interpretations. A summary of studies presented before 1994 can be found in [7]. Following these studies, we could divide the photosphere into two regions with different physical properties and assume a boundary between them at a height $H \approx 170$ km above the continuum height. This conclusion was based on the fact that the fluctuation amplitudes for the intensities and velocities of the motions first decrease with height, reaching a minimum at $H \approx 170$ km, then increase again, with the velocities at $H < 170$ km and $H > 170$ km being uncorrelated. In the lower photosphere ($H < 170$ km), the energy is transported from large to small scales by the convective vortex-granules. Only granules exceeding $1.5''$ in size reach heights of $H \approx$

Table 1. Line parameters and observing conditions

Wavelength, nm	Element	EPL, eV	d, %	H_d, km	Δt, min	N	Date
532.418	Fe I	3.21	0.861	500	31.3	200	Aug. 10, 1996
523.462	Fe II	3.22	0.722	340	158.1	943	Aug. 26, 2001
639.361	Fe I	2.42	0.745	490	158.1	943	Aug. 26, 2001

Note: The columns present the (1) wavelength, (2) chemical element and ionization state, (3) excitation potential of the lowest level EPL, (4) central depth d of the line according to the Liège atlas of the solar spectrum [13], (5) height H_d in the solar atmosphere where the line center is formed calculated allowing for deviations from LTE, (6) duration of the observations Δt, (7) number of the exposures N, and (8) date of the observation.

170 km. In the upper photosphere ($H > 170$ km), various waves play the dominant role in the energy transport. On the other hand, Espagnet *et al.* [7] concluded that the photosphere could be divided into two regions only in terms of the intensities of the motions, since the vertical convective velocities penetrate to heights of $H \approx 550$ km. Salucci [8] arrived at a similar conclusion.

In all these studies, as well as those published subsequently [9, 10], the complex intensities and velocities of the motions were analyzed using a very simple two-column model. Here, we do not restrict our analysis to two-column models in our study of the interrelations between the velocities and intensities based on data with high spatial and temporal resolutions. We compare our results with theoretical calculations based on a three-dimensional hydrodynamic model kindly provided by Dr. M. Asplund [11].

2. OBSERVATIONAL DATA

We used spectrograms with high spatial ($0.5''$) and temporal (9 s) resolutions obtained in 1996 and 2001 on the German Vacuum Tower Telescope (VTT) in Izana (Tenerife, Spain). Information about the telescope and spectrograph are presented in [12]. Spectral observations of the center of the solar disk were carried out in an unperturbed region, accompanied by a video recording in the Hα and Ca II K lines. A CCD camera with 1024×1024 pixels connected in pairs was used. The spatial resolution of a single pair-connected pixel was $0.174''$. The entrance slit of the spectrograph "cut off" an area of $0.38 \times 98''$ on the solar disk. The spectra were recorded for a single point on the solar disk. For the 1996 observations, we chose the $\lambda 532.418$ nm neutral iron line, for which we obtained 31 min of recordings. The 2001 observations were carried out with two cameras, which recorded 158 min of data for the Fe I $\lambda 639.361$ nm and Fe II $\lambda 523.462$ nm lines. More detailed information on the line parameters and observing conditions is presented in Table 1.

After the usual preliminary reduction of the spectral data (corrections for the dark current, fluctuations in the transparency of the Earth's atmosphere, and the nonuniform sensitivity of the CCD pixels), we determined the fluctuations in the intensities and velocities of all three lines for each position along the spectrometer slit and for each moment in time at eleven heights in the solar atmosphere using the "lambdameter" technique discussed in detail in [14]. The halfwidths of the spectral lines (the width between the center and the line wing) and heights in the solar atmosphere calculated allowing for deviations from local thermodynamic equilibrium (so-called NLTE heights) are presented in Table 2 for the intensities and velocities detected.

The spatial and temporal variations in the intensities and velocities are due to convective and wave motions. To separate these different types of variations, we constructed a diagnostic diagram that presents the power as a function of the temporal and spatial frequencies. The procedure used to separate the intensities and velocities into convective and wave components has been discussed in more detail previously [15−17]. Our further analysis deals only with the convective components. We have taken motions toward the observer to be positive.

3. RESULTS OF THE OBSERVATIONS

The upper left-hand plot in Fig. 1 shows the correlation $\langle V_C, V_H \rangle$ averaged over space and time between the velocity V_C at the continuum height ($H \approx 10$ km) and the velocities at all other heights V_H. This correlation gradually decreases from unity at $H \approx 10$ km to zero at $H \approx 500$ km. This observational fact has already been noted in numerous papers. Note that the results obtained for the ionized iron are consistent with the general dependence, although the correlation coefficient displays a smaller gradient with height. Note also the reversal in the sign of the correlation for the Fe I $\lambda 639.3$ nm line at $H \approx 500$ km. Since the negative correlation is very weak and reversals of the velocity direction in the upper photosphere have not

Table 2. Half-widths $\Delta\lambda_i$ of the spectral lines and corresponding heights $H(\Delta\lambda_i)$ where the granular and intergranular intensities and velocities were detected over the entire observational period

Wavelengths, nm	Half-widths $\Delta\lambda_i$, pm										
532.418	0.5	1.5	3.0	4.0	4.9	5.5	6.0	7.2	9.5	14	17
523.462	1.0	1.9	3.3	4.2	4.9	5.5	6.1	7.2	9.3	15	20
639.361	0.5	1.5	3.0	4.0	4.9	5.5	6.0	7.2	9.5	14	18
Wavelengths, nm	Heights $H(\Delta\lambda_i)$, km										
532.418	15	40	110	160	200	220	250	300	360	460	490
523.462	-25	-23	-8	20	45	60	80	120	180	290	340
639.361	5	20	70	130	190	210	250	320	390	470	490

been discussed earlier, we studied this property of the convective velocities in more detail.

Figure 2 (left) presents for the same line the velocity at various heights in the solar atmosphere as a function of the intensity (contrast) I_C in the continuum, which is formed at $H \approx 10$ km. At heights near the continuum height, the material ascends above granules and descends above intergranules. The absolute values of the velocities above granules and intergranules gradually decrease with height, approach zero at $H \approx 390$ km, and then increase again, with upward motions dominating above intergranules, and downward motions, above granules. However, this velocity-sign reversal occurs only above structures which display contrasts in the continuum higher than $6-7\%$ in absolute value. Note that the classical definition of granules and intergranules is valid only to the continuum height. This definition loses its original meaning at any other heights.

Let us return to Fig. 1. The middle left plot shows the correlation $\langle I_C, I_H \rangle$ between the contrast I_C at $H \approx 10$ km and the intensity I_H at all other heights in the solar atmosphere. The absolute value of this correlation decreases from unity at the continuum height to zero at $H \approx 250$ km, then increases again; that is, the brightness correlation changes from positive to negative 240 km below the height where the velocity correlation reverses. In other words, the material above granules already becomes cooler than its environment at heights $H > 250$ km. The sign reversal of the intensity with height above granules and intergranules is illustrated in Fig. 3 (left), which shows the intensity in the Fe I $\lambda639.3$ nm line for various heights as a function of the contrast in the continuum. Note that the intensity is the same as the relative contrast, since it is determined relative to the average value for each height.

The bottom left plot in Fig. 1 shows the correlation $\langle V_H, I_H \rangle$ between the velocity V_H and intensity I_H

at each height. This correlation first decreases with height, reversing its sign at $H \approx 200$ km, after which its magnitude increases to the height $H \approx 400$ km and then decreases again to zero at $H \approx 500$ km. This complex behavior of the correlation $\langle V_H, I_H \rangle$ results from the behavior of the correlations $\langle I_C, I_H \rangle$ and $\langle V_C, V_H \rangle$. Note that, if the correlation $\langle V_C, V_H \rangle$ is positive, the sign change of the correlation $\langle V_H, I_H \rangle$ at height $H \approx 200$ km does not contradict the sign reversal of the correlation $\langle I_C, I_H \rangle$ at height $H \approx 250$ km, since the correlation $\langle V_H, I_H \rangle$ corresponds to a single height, whereas the correlations $\langle I_C, I_H \rangle$ and $\langle V_C, V_H \rangle$ correspond to different heights. Note that the correlation $\langle V_H, I_H \rangle$ does not exceed 0.75, even at the continuum height. It appears that total coincidence of the intensity and velocity maxima is rare for the solar photosphere.

Thus far, we have examined the intensities and velocities in a two-column model. It is of interest to study the motions in a model with a larger number of degrees of freedom. If we assume that relatively hot or cool material at two heights $H \approx 10$ km and $H \approx 500$ km can move both upward and downward, we find 16 types of convective motions, shown schematically at the bottom of Fig. 4. Ascending material is denoted by a plus sign, and descending material, by a minus sign. We also used plus and minus signs to denote material whose temperature is higher than or lower than the average temperature for a given height. It turns out that all 16 types of convective motions are indeed observed in convective columns on the Sun, even such, at first glance, improbable motions as motion 13, with cool material ascending and becoming hot at the temperature minimum. The upper plot in Fig. 4 shows the number of cases (solid curve) corresponding to each of the 16 convective motions, while the second and third plots from the top show the average absolute values of the velocity $\langle V_C \rangle$ and contrast $\langle I_C \rangle$ at $H \approx 10$ km. The fourth plot in Fig. 4
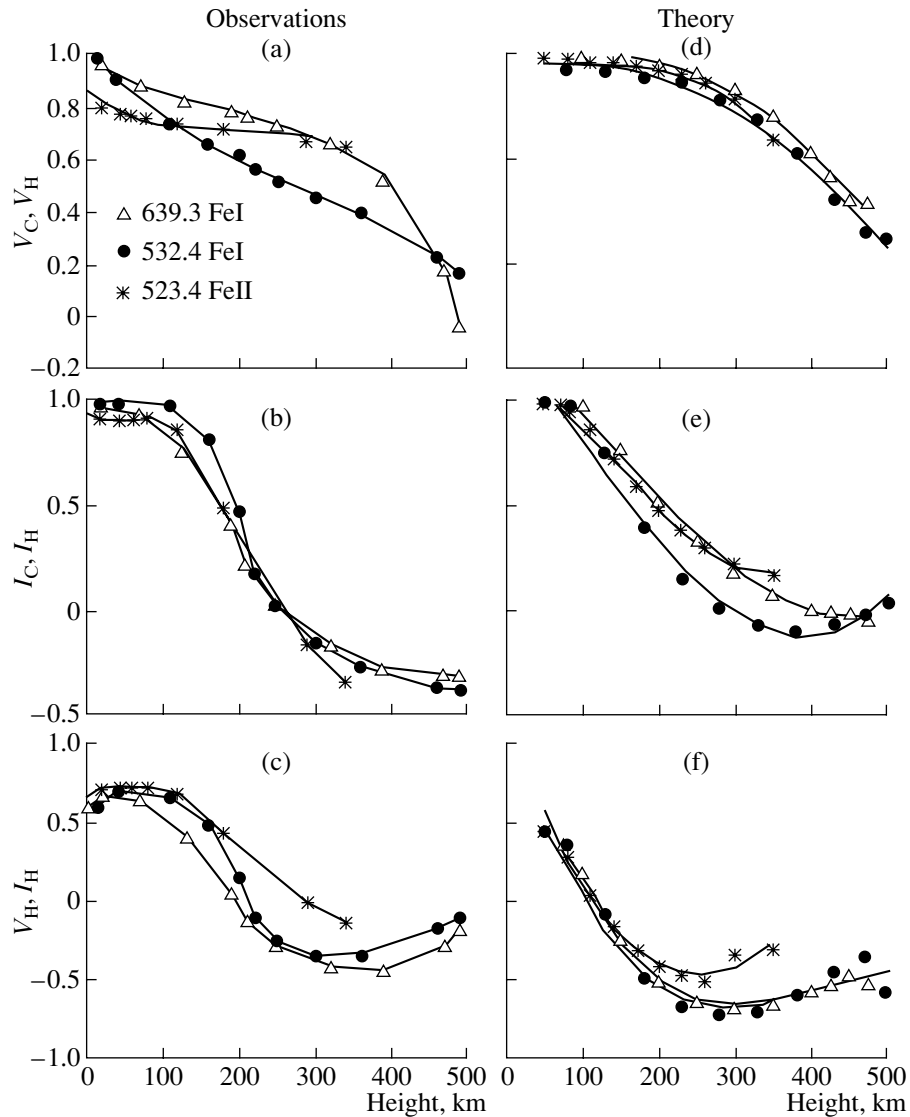
Observations

(a)

Theory

(d)

Fig. 1. The correlations $\langle V_C, V_H \rangle$, $\langle I_C, I_H \rangle$, and $\langle V_H, I_H \rangle$ obtained for various heights in the solar atmosphere using the spectral data for neutral and ionized iron. Shown are the observational data (left) and computations using the three-dimensional hydrodynamic model (right).

shows a parameter that we have called the "efficiency" (solid curve), which is the product of five quantities: (i) the number of cases for each type of convective motion, (ii) and (iii) their velocity and intensity with respect to the average values at $H \approx 10$ km, and (iv) and (v) the same two quantities for $H \approx 500$ km. This efficiency varies widely, from nearly zero (motions 3 and 15) to almost 30% (motion 10).

Four types of motions are most efficient. In two of these, the material changes only the sign of its relative contrast, while both the sign of the contrast and the direction of the motion change with height in the two others. In more detail, these four motions are the following:

(1) material that is hot at the continuum height

and cool at the temperature minimum ascending at all heights in the photosphere (motion 14, with an efficiency of 18%);

(2) material that is hot at $H \approx 500$ km and cool at $H \approx 10$ km descending at all heights (motion 10, 28%);

(3) hot material ascending at the continuum height and cool material descending at the temperature minimum (motion 8, 15%);

(4) cool material at $H \approx 10$ km descending and hot material at $H \approx 500$ km ascending (motion 7, 13%).

The total efficiency of these four motions is 74%. Thus, we can understand that two-column models
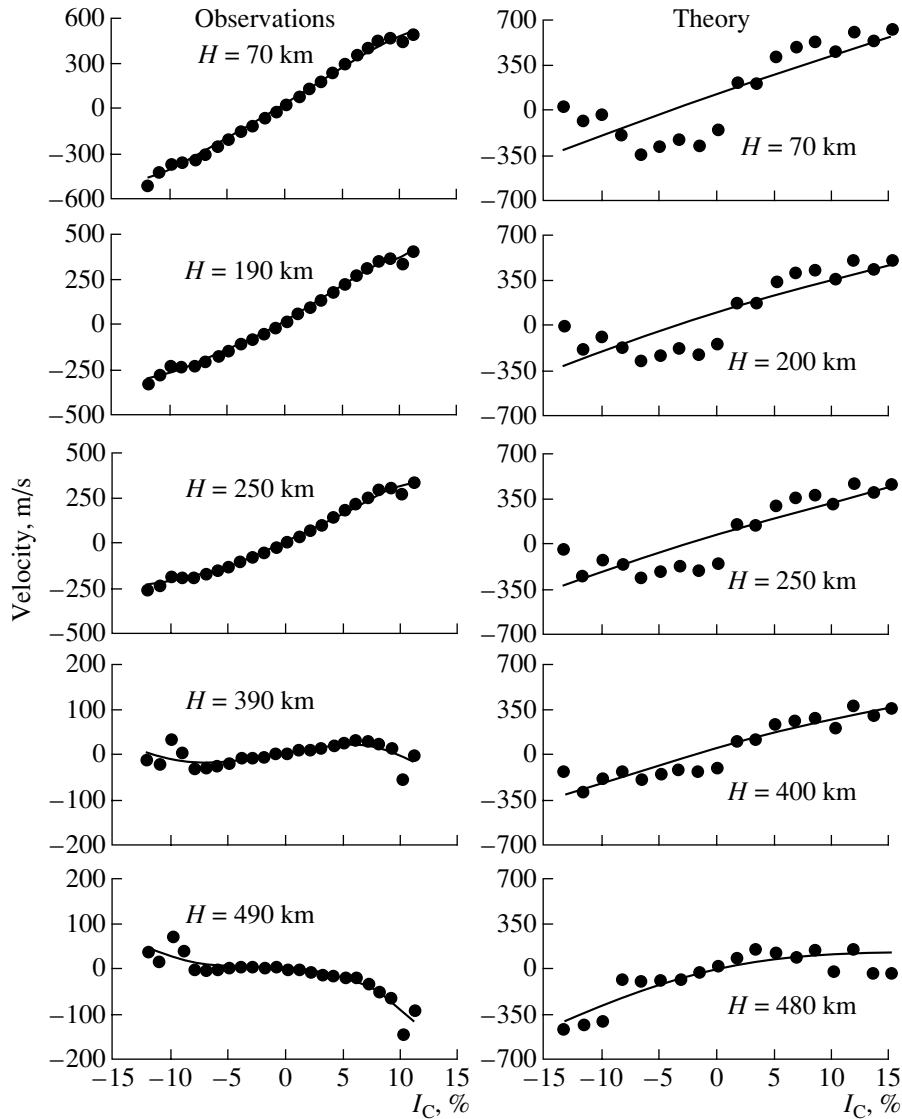
Fig. 2. The velocity of convective motions at five heights in the solar atmosphere as a function of the contrast I_C in the continuum. Shown are the observational data for the Fe I λ639.3 nm line (left) and computations using the three-dimensional hydrodynamic model (right).

for the solar photosphere with hot ascending (motion 1) and cool descending (motion 2) materials cannot explain the observed asymmetry and shifts of the Fraunhofer lines, since the total efficiency of these motions does not exceed 10%. At the same time, the semiempirical four-column model of [18], which admits both upward and downward motions of cool and hot materials, can successfully explain almost all features of the fine structure of the absorption lines observed in the solar spectrum and has predicted some features of the fine structure of the Fraunhofer lines that have been verified by observations.

Let us consider in detail these four most efficient motions. Figure 5 (left) presents heights where the material changes the sign of its relative contrast I_H

in convective columns with motions 14 and 10. We can see in Figs. 5a and 5b that the height where ascending (at all heights) material that is hot near the continuum height becomes cool (motion 14) depends only slightly on the velocity V_C and contrast I_C in the continuum. On average, this height is $H_{14}^o \approx 280 \pm 105$ km. This is also true for descending material; namely, the height where material that is hot near the temperature minimum becomes cool is almost independent of both V_C and I_C. The contrast sign reversals occur, on average, at $H_{10}^o \approx 260 \pm 100$ km (Figs. 5c, 5d).

Figure 6 presents the heights in columns with motions 8 and 7, in which there is a sign reversal of both the contrast I_H of the material and the velocity V_H of
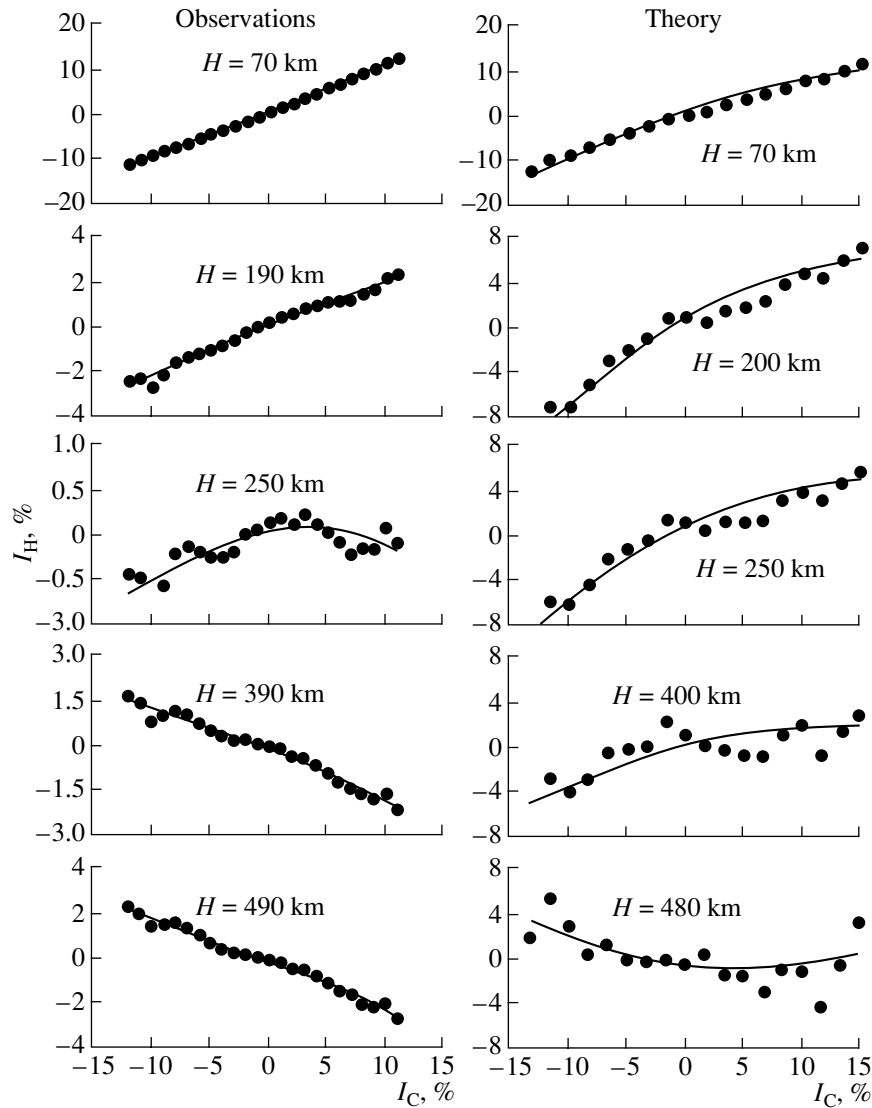
Fig. 3. Contrast of the Fe I λ639.3 nm line as a function of the intensity I_C in the continuum. Shown are the observational data (left) and computations using the three-dimensional hydrodynamic model (right) for five heights in the solar atmosphere.

the convective motions. In contrast to the previous case, the corresponding heights depend significantly on the velocity V_C at the continuum height and, to a lesser extent, on the contrast I_C in the continuum. Figure 6a shows that, the higher the velocity of ascending hot material, the higher the height of the sign reversals for the contrast and direction of motion. This height is $H_8^o \approx 360 \pm 130$ km, on average. We would expect this behavior based on general reasoning. However, it is difficult to understand the behavior shown in Fig. 6b: the lower the contrast of the granule in the continuum, the higher the height for the sign reversals of both the direction of motion and the relative contrast of the material. Unfortunately, we have not yet found any satisfactory explanations for this tendency. Figures 6c and 6d show the heights

of layers separating cool, descending material and hot, ascending material near the continuum height. These heights depend on the velocity at the continuum height and are virtually independent of the contrast in the continuum. These heights are $H_7^o \approx 350 \pm 120$ km, on average. Note that the heights H_7^o and H_8^o exceed H_{10}^o and H_{14}^o. This is quite reasonable, since we have taken into account for H_7^o and H_8^o both the sign reversal of the relative contrast I_H (as for H_{10}^o and H_{14}^o) and the sign reversal of the velocity V_H. As we have already noted, the velocity sign reversal occurs at higher heights than does the reversal of the relative contrast.

4. THEORETICAL COMPUTATIONS

Numerical modeling of the solar granulation has developed considerably over the last decade. Two-dimensional [19−21] and three-dimensional [11] hydrodynamical models that can reproduce spectral lines and their asymmetry and shifts have recently been constructed. However, we are not aware of any detailed comparisons of the calculated fine structure of the intensities and velocities with observations. In this section, we compare all the observed dependences presented in Figs. 1−6 with the results of calculations. We chose the three-dimensional model of Asplund *et al.* [11], which recommended itself highly in our previous studies [17, 22]. Note that we have data only for a single time.

The modeled region of the solar granulation is a rectangular solid 6000×6000 km in size (50×50 grid points) in the horizontal plane and with a vertical height of 3800 km. Thus, we have 50×50 one-dimensional atmospheric models describing the height variations of the temperature, density, velocity, etc., for each point of the horizontal grid. The spectral line profiles for individual models were calculated allowing for deviations from local thermodynamic equilibrium (NLTE profiles) using the NATAJA codes, which employ efficient iterative techniques for solving the radiation transport equations for multilevel atomic models in the so-called 1.5D approximation [23]. We used a realistic model for the iron atom described earlier [22, 24], which takes into account the fine structure of terms and contains about 250 levels and nearly 500 radiative bound−bound and bound−free transitions in the ultraviolet, visible, and infrared, including transitions near the FeI continuum. Thus, we obtained 2500 theoretical profiles for each spectral line. We determined the intensity and velocity fluctuations for eleven heights in the solar atmosphere by applying the lambdameter technique to these profiles, as we had for the observed profiles (Table 2).

Before we proceed with the analysis of the computations, it is interesting and helpful to compare the modeled convective velocities with the velocities reconstructed from the spectral lines. This provides a verification of all our calculations and a test for the lambdameter technique. This comparison is presented in Figs. 7a−7c. We have selected three one-dimensional models in which the calculated convective velocities increase with height (Fig. 7a), decrease with height (Fig. 7c), and change direction (Fig. 7b). The calculated quantities are shown by solid curves, and the quanities reconstructed from the spectral lines, by asterisks. In all three cases, the reconstructed velocities are in satisfactory (qualitative and quantitative) agreement with the model velocities. Any better agreement can hardly be expected,
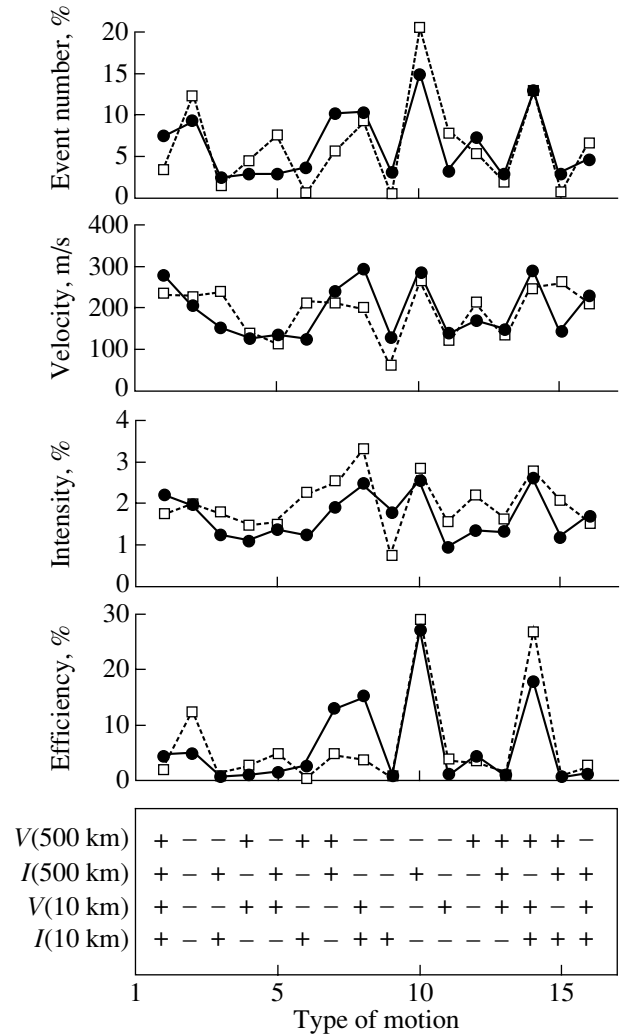

Fig. 4. Motions in the solar atmosphere according to the 16-column model. The upper plot shows the number of cases corresponding to each of the 16 types of convective motions. The second and third plots from above show the average absolute values of the velocity $\langle V_C \rangle$ and contrast $\langle I_C \rangle$ for these motions at $H \approx 10$ km. The signs of $\langle I_C \rangle$ and $\langle V_C \rangle$ are shown in the diagram at the bottom of the figure. The fourth plot shows the "efficiency" of the motions, which is the product of five quantities: (i) the number of cases, the (ii) velocity and (iii) contrast at $H \approx 10$ km, and the (iv) velocity and (v) contrast at $H \approx 500$ km.

since the discrepancies are associated with the simple method used to determine the height where the emission at the line frequency is formed, rather than with any disadvantages of the lambdameter technique or with computational errors. The point here is that the velocities shown by the asterisks in Fig. 7 corresponded to the heights at which the optical depths of the corresponding portions of the line profiles are unity. In reality, the emission at each frequency is formed in a fairly extended layer of the photosphere
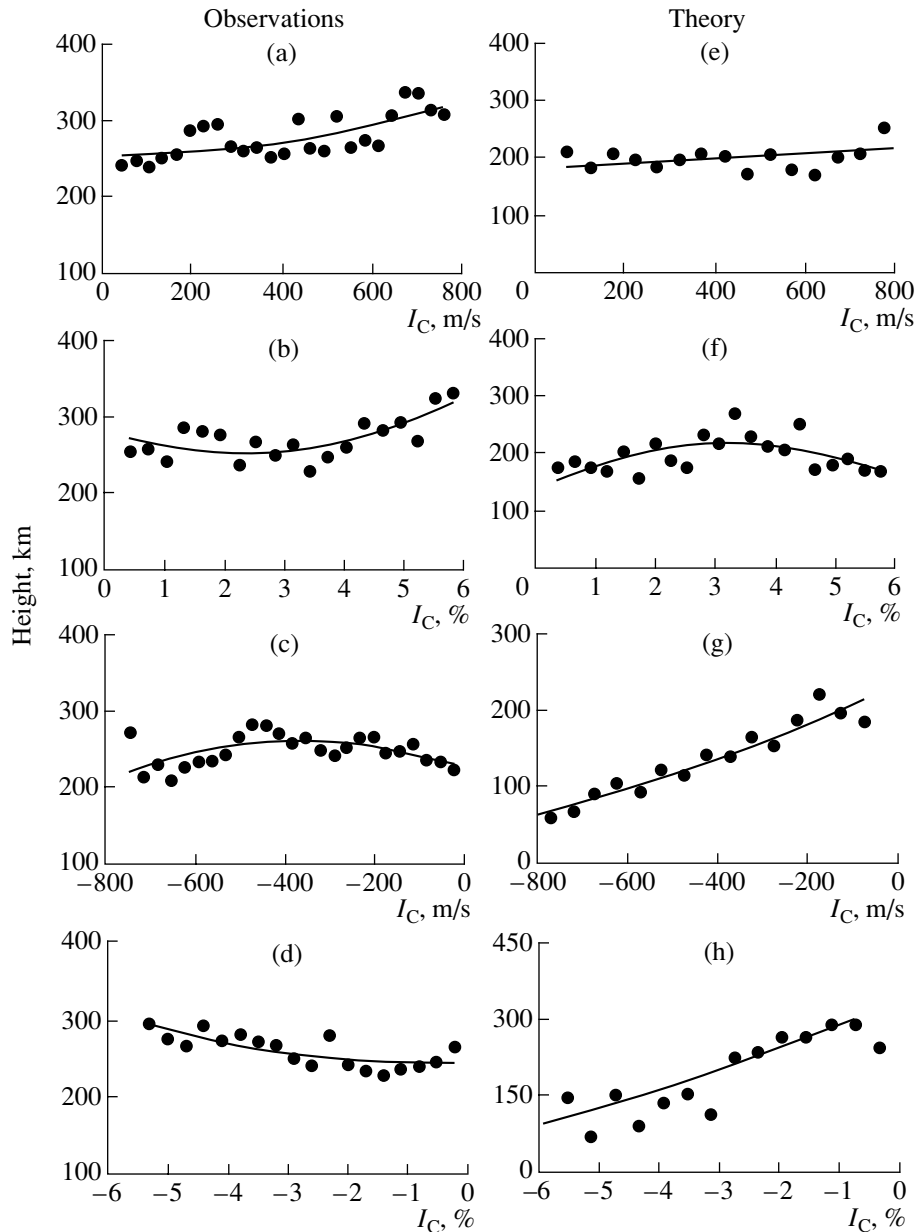
Fig. 5. Observed (a, b) and calculated (e, f) heights at which ascending material that is hot near the continuum height becomes comparatively cool. Observed (c, d) and calculated (g, h) heights at which descending material that is hot at the temperature minimum ($H \approx 500$ km) becomes comparatively cool.

(thicker than 100 km), with the convective velocity varying appreciably within this layer. Therefore, the reconstructed velocities are weighted in accordance with their contribution to the emission and, naturally, differ from the model velocities for a specific height.

We would like to draw your attention to the agreement between the observed [13] and spatially averaged theoretical profiles (Figs. 7d–7f), which provides further verification of the reliability of our computations using the three-dimensional model of the solar granulation.

Let us proceed to a comparison of the computations with the observations. Figure 1d shows the correlation $\langle V_C, V_H \rangle$. All three lines demonstrate similar behavior and satisfactory agreement with the observations. However, the Fe I λ639.361 nm line shows no sign reversal for the correlation at $H \approx 500$ km, in contrast to Fig. 1a, which presents the observational data. Similar features can be seen in Fig. 2 (right). In addition, the calculated correlation coefficients systematically exceed the observed coefficients, especially for the Fe II λ523.462 nm line. For convenience, we have reduced the spatial resolution of
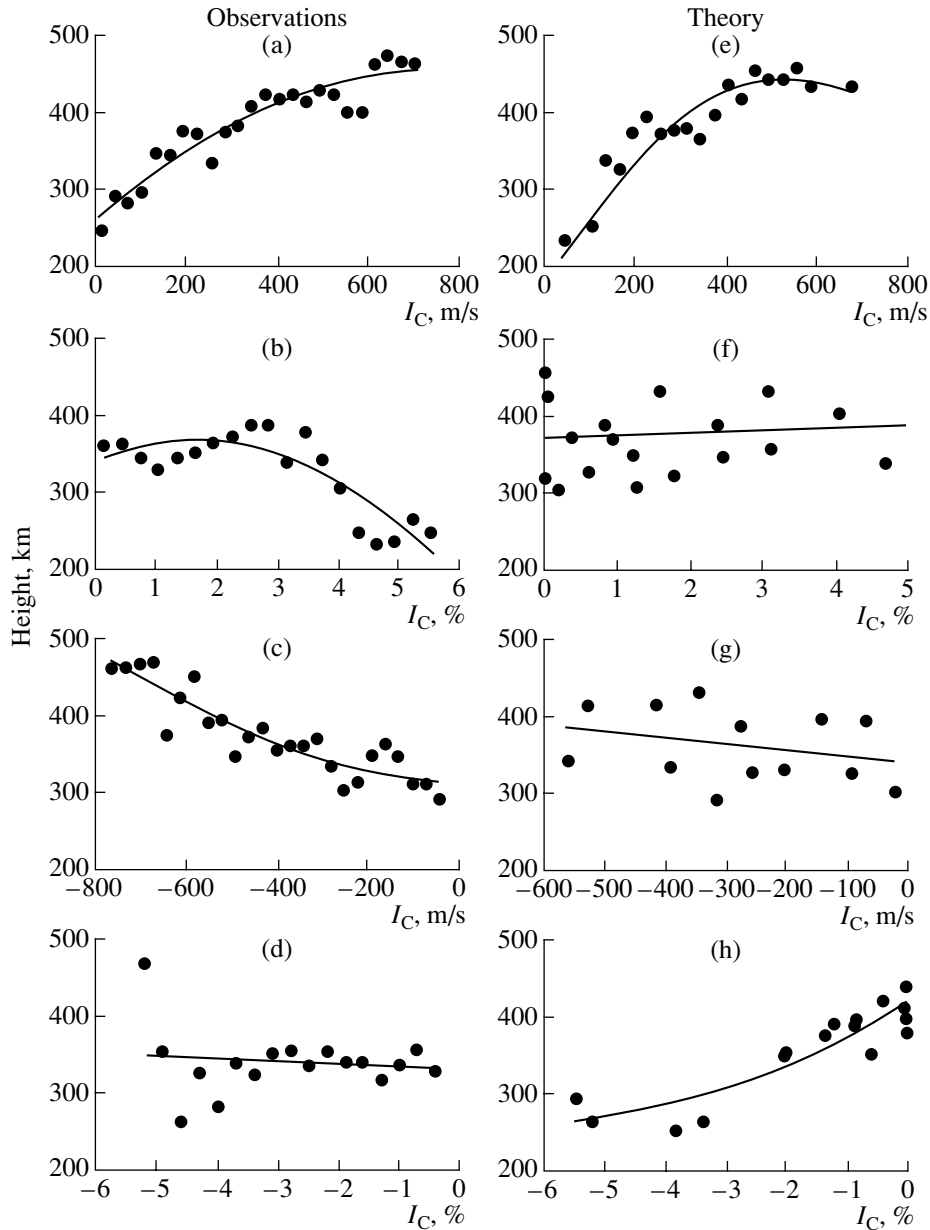
Fig. 6. Observed (a, b) and calculated (e, f) heights separating hot, ascending material and cool, descending material near the continuum height. Observed (c, d) and calculated (g, h) heights separating cool, descending material and hot, ascending material near the continuum.

calculated parameters (the contrast and velocity) to the observational resolution.

Note that the correlations obtained from the observed (Figs. 1a–1c) and theoretical (Figs. 1d–1f) spectral-line profiles correspond to somewhat different heights, especially near the continuum height. This is due to the different spatial resolution of the computations, which results in somewhat broader height intervals for the theoretical profiles than for the observed profiles with the same line halfwidths. Consequently, the minimum height for the calculated profiles exceeds that for the observed profiles.

The general behavior of the calculated correlation $\langle I_C, I_H \rangle$ (Fig. 1e) is similar to the observed behavior (Fig. 1b): there is a height decrease from unity to zero, a sign reversal, and then a small increase of the negative correlation. However, clear differences are also evident. In the three-dimensional model, the sign reversal occurs, on average, at $H^t \approx 340$ km, while the observations give $H^o \approx 250$ km (see also Fig. 3, right); the minimum value for the theoretical correlation $\langle I_C, I_H \rangle$ is -0.1, while the minimum value for the observed correlation is -0.4.

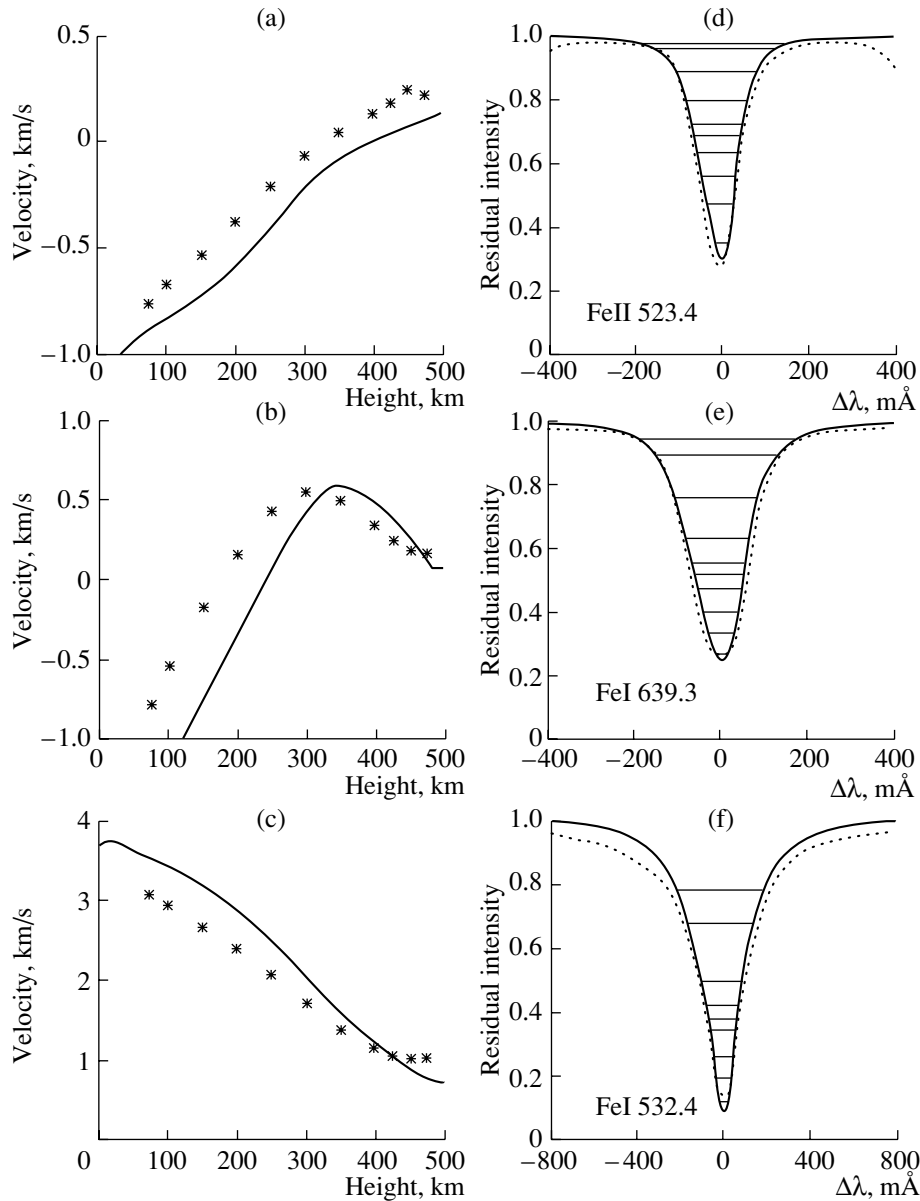The general behaviors of the observed (Fig. 1c)

Fig. 7. Convective velocities (left) for three one-dimensional models of the solar granulation (solid curves) compared with the velocities reconstructed from the calculated line profiles (asterisks). Comparison (right) of the calculated, spatially averaged spectral-line profiles (solid curves) with the observed profiles (dotted curves). The horizontal lines show the levels corresponding to the intensity and velocity fluctuations.

and computed (Fig. 1f) $\langle V_H, I_H \rangle$ correlations are also similar: a decrease in the positive correlation with height, followed by a sign reversal, an increase in the negative correlation, a minimum, and a gradual decrease of the negative correlation to zero. The most significant differences occur at heights from zero to 200 km; the observed correlation coefficients are almost constant for the first 100 km, while the calculated coefficients decrease to zero, and the theoretical correlations are negative for heights from 100 to 200 km, while the observed coefficients are positive. Consequently, the sign reversal for the theoretical

correlation occurs at a lower height (100 km) than does the observed sign reversal (200 km).

The relations obtained from the calculated spectral-line profiles are shown by the small squares in Fig. 4. This plot shows that numerous observed features of the intensities and velocities are qualitatively and quantitatively reproduced by the model, and all 16 types of convective motions are present in the three-dimensional hydrodynamical model [11].

There is satisfactory agreement between the three-dimensional model and the observations (within the

rms deviations) for the heights at which the ascending material (motion 14) reverses the sign of its contrast with respect to the environment, $H_{14}^t \approx 210 \pm 130$ km (Figs. 5e, 5f). Note that we have obtained $H_{14}^o \approx 280 \pm 105$ km from the observations. Analyzing the relations in Figs. 5e, 5f, we conclude that the height at which the ascending granules change their contrast is virtually independent of both the velocity and contrast in the continuum, in satisfactory agreement with the observations (Figs. 5a, 5b).

Figures 5g and 5h show the heights of the sign reversal for the contrast calculated for the descending material in the three-dimensional model (motion 10). In contrast to the observations presented in Fig. 5c, these heights (Fig. 5g) depend on the velocity in the continuum: the higher the descent velocity in the continuum, the lower the height at which the contrast sign changes from positive to negative. On average, the theory yields $H_{10}^t \approx 140 \pm 120$ km, whereas the observations indicate an appreciably higher height, $H_{10}^o \approx 260 \pm 100$ km. We can see a similar difference for the dependence of the height H_{10}^t on the contrast in the continuum (Fig. 5h).

Figures 6e and 6f present the heights at which the ascending granules become comparatively cool and begin to descend calculated in the three-dimensional model (motion 8). The higher the granule velocity in the continuum, the higher the height of the sign reversal. On average, this sign reversal occurs at $H_8^t \approx 380 \pm 110$ km, in good agreement with observations, which yield $H_8^o \approx 360 \pm 130$ km. On the other hand, the heights H_8^t do not depend on the contrast in the continuum, whereas the observed heights do (Fig. 6b).

Comparing Figs. 6g and 6h (motion 7) with the observations presented in Figs. 6c and 6d, we see that the dependences of the theoretical H_7^t and observed H_7^o heights on the contrast and velocity of the intergranules in the continuum are rather different, although the modeled and observed heights separating the cool, descending material from the hot, ascending material are quite similar ($H_7^t \approx 360 \pm 100$ km and $H_7^o \approx 350 \pm 120$ km, respectively). The heights H_7^t are virtually independent of the velocity in the continuum, whereas the observed heights H_7^o increase with increasing descent velocity of the intergranules in the continuum. In turn, the observed heights H_7^o do not depend on the contrast in the continuum, whereas the calculated heights H_7^t appreciably increase with decreasing contrast I_C in the continuum.

5. DISCUSSION AND CONCLUSIONS

We have studied the convective velocities and intensities in the solar photosphere at heights from the level of formation of the continuum to the height of the temperature minimum using spectral observations of ionized and neutral iron lines. Our correlation analysis has shown that the granules and intergranules change their relative brightnesses at an average height of $H \approx 250$ km. This height is in good agreement with the results presented in [25], which summarizes the data of numerous observations and publications. Note that Evans and Catalano [26], as well as Holweger and Kneer [27], were apparently the first to identify the sign reversal of the correlation $\langle I_C, I_H \rangle$ in the solar atmosphere.

A general reversal in the direction of motion occurs at $H \approx 500$ km; namely, the material above granules begins to predominantly descend, and the material above intergranules, to ascend. This behavior has apparently not been discussed earlier. However, this result must be verified, since it has been found for only a single spectral line. On the other hand, the data presented in Figs. 4 and 6 show that about 20% of the objects in the solar atmosphere reverse their motion at heights of $H \approx 350$ km.

The maximum magnitude of the correlation coefficient $\langle V_H, I_H \rangle$ does not exceed 0.75, in good agreement with the recent findings of [9]. The topologies of the intensities and velocities are apparently already different at the continuum formation height.

There are various combinations of the direction of motion and relative brightness of materials in the solar photosphere. We have analyzed 16 cases and shown that four motions are most typical and "efficient" (for spectral lines). In the first two cases, the material reverses only the sign of its relative contrast: hot, ascending material becomes relatively cool at a mean height of $H \approx 260$ km, while cool, descending material becomes relatively hot at a mean height of $H \approx 280$ km. In the last two cases, hot, ascending material at $H = 0-350$ km; cool, descending material at $H = 500-350$ km; cool, descending material at $H = 360-0$ km; and hot, ascending material at $H = 360-500$ km reverse both the sign of their contrast and their direction of motion.

We have compared all observed relations with theoretical dependences calculated using the three-dimensional hydrodynamical model of [11], taking into account deviations from local thermodynamic equilibrium when calculating the profiles of spectral lines. In our opinion, this model satisfactorily reproduces all the main features of the convective velocities and intensities. We have found all 16 types of motion considered in this model and have obtained a satisfactory agreement with the observations.

Note also the agreement of our observational data shown in Figs. 1b and 1c with the calculations presented in [28] and performed for the two-dimensional model of [20, 21].

Thus, our study shows that the convective motions maintain their column structure in the lower photosphere right to the height of the temperature minimum. Therefore, it seems unjustified to divide the photosphere into two regions with different intensities and velocities. The correlations $\langle I_C, I_H \rangle$ and $\langle V_C, V_H \rangle$ gradually vary in the solar atmosphere and reverse their signs at different heights—250 km for the intensity and 500 km for the velocity. We suggest that the zero value for the correlation $\langle V_H, I_H \rangle$ at $H \approx 200-300$ km does not imply that the convective brightness fields are disrupted above this level.

REFERENCES

1. A. P. Hansky, Pulkovo Mitt. **3**, 1 (1908).
2. A. Unsold, Z. Astrophys. **1**, 138 (1930).
3. R. S. Richardson and H. Schwarzschild, Astrophys. J. **111**, 351 (1950).
4. V. A. Krat and A. A. Shpital'naya, Soln. Dannye, No. 2, 63 (1974).
5. V. N. Karpinskii, Soln. Dannye, No. 2, 91 (1980).
6. V. N. Karpinsky, *Solar Photosphere: Structure, Convection, and Magnetic Fields,* Ed. by J. O. Stenflo (Kluwer, Dordrecht, 1990), p. 67.
7. O. Espagnet, R. Muller, Th. Roudier, *et al.* Astron. Astrophys., Suppl. Ser. **109**, 79 (1995).
8. G. Salucci, L. Bertello, F. Gavillini, *et al.*, Astron. Astrophys. **285**, 322 (1994).
9. A. Nesis, R. Hammer, M. Roth, and H. Schleicher, Astron. Astrophys. **396**, 1003 (2002).
10. A. Hanslmeier, A. Kučera, J. Rýbak, *et al.*, Astron. Astrophys. **356**, 308 (2000).
11. M. Asplund, Å. Nordlund, R. Trampedach, *et al.*, Astron. Astrophys. **359**, 729 (2000).
12. E. H. Schröter, D. Soltau, and E. Wiehr, Vistas Astron. **28**, 519 (1985).
13. L. Delbouille, L. Neven, and C. Roland, *Photometric Atlas of the Solar Spectrum from λ3000 to λ10 000* (Institut d'Astrophysique de l'Université de Liège, Liège, Belgium, 1973).
14. R. T. Stebbins and P. R. Goode, Solar Phys. **110**, 237 (1987).
15. E. V. Khomenko, R. I. Kostik, and N. G. Shchukina, Astron. Astrophys. **369**, 660 (2001).
16. R. I. Kostyk and N. G. Shchukina, Pis'ma Astron. Zh. **25**, 781 (1999) [Astron. Lett. **25**, 678 (1999)].
17. R. I. Kostyk and E. V. Khomenko, Astron. Zh. **79**, 1027 (2002) [Astron. Rep. **46**, 925 (2002)].
18. R. I. Kostyk, Astron. Zh. **62**, 112 (1985) [Sov. Astron. **29**, 65 (1985)].
19. M. Steffen, H. G. Ludvig, and A. Kruss, Astron. Astrophys. **213**, 371 (1989).
20. A. S. Gadun, Kinemat. Fiz. Neb. Tel **11** (3), 54 (1995).
21. A. S. Gadun, S. K. Solanki, and A. Johannesson, Astron. Astrophys. **350**, 350 (1999).
22. N. G. Shchukina and J. Trujillo Bueno, Astrophys. J. **550**, 970 (2001).
23. L. H. Auer, P. Fabiani Bendicho, and J. Trujillo Bueno, Astron. Astrophys. **292**, 599 (1994).
24. N. G. Shchukina and J. Trujillo Bueno, Kinemat. Fiz. Neb. Tel **14** (4), 315 (1998).
25. A. Kučera, J. Rýbak, and H. Wöhl, Astron. Astrophys. **298**, 917 (1995).
26. J. W. Evans and C. P. Catalano, Solar Phys. **27**, 299 (1972).
27. H. Holweger and F. Kneer, *Solar and Stellar Granulations*, Ed. by R. J. Rutten and G. Severino (Kluwer, Dordrecht, 1989), p. 173.
28. A. S. Gadun, A. Hanslmeier, A. Kučera, *et al.*, Astron. Astrophys. **363**, 289 (2000).

Translated by V. Badin

The Magnetic-Field Geometry
in Polar Ray Structures of the Solar Corona

B. P. Filippov[1,2], Yu. V. Platov[1], A. Ajabshirizadeh[3], and D. V. Klepikov[1]

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

[2]*Research Institute for Astronomy and Astrophysics of Maragha, P.O. 441, Maragha, Iran*

[3]*Department of Theoretical Physics and Astrophysics, Faculty of Physics, University of Tabriz, Tabriz, Iran*

Received February 4, 2004; in final form, March 15, 2004

Abstract—Variations in the positions of the intersection points of tangents to ray structures in the polar corona of the Sun during the solar cycle are considered. At first glance, the decrease in the distance q between the tangent intersection point and the center of the solar disk during activity maximum contradicts harmonic analyses that indicate that the relative weight of higher harmonics in the global field increases during this period. Indeed, the higher the harmonic number in an axisymmetric field, the closer the intersection point of the field-line tangents (the magnetic focus) to the solar surface. It is shown that q for a field composed of two harmonics with opposite polarities at the poles can be smaller than q for either of them taken alone. A simple model representing the global field using the third and seventh harmonics is analyzed; this model can reproduce quite satisfactorily the observed dynamics of magnetic foci of the polar field. © 2004 MAIK "Nauka/Interperiodica".

1. INTRODUCTION

The magnetic field of the polar regions of the Sun is of considerable interest from the standpoint of investigations of the nature of solar activity, cyclic laws, and dynamo theory. Polar zones can be regarded as relatively quiescent regions of the solar surface: sunspots and flares are never observed there, and prominences and filaments penetrate into these regions only during epochs of activity maximum, before the polarity reversals of the global magnetic field; no differential rotation is present in the polar regions; the angular momentum of the polar zones is very small. On the other hand, large-scale and long-period variations are clearly pronounced against the background of this quiescence. They seem to reflect processes occurring in the convection zone—a region of magnetic-field generation.

Since the rotational axis of the Sun is almost normal to the plane of the ecliptic, polar regions cannot face the Earth and are always located near the limb, where Zeeman magnetography is virtually impossible. For this reason, indirect techniques for estimating the polar magnetic fields are important. For example, the position of the polar crown prominence ring outlines the polarity-inversion line of the radial magnetic-field component and provides the most adequate information about the epoch of the polarity reversal [1–5]. Important information can also be gained from the geometry of the magnetic field lines. Coronal ray structures follow the shape of the field lines quite closely, since, in the case of strong magnetic fields (small plasma parameters β), the coefficients for transport along the field are much larger than those for transport across the field, and any density inhomogeneities are stretched along the magnetic field lines. It is the similarity between the ray structures observed near the solar poles during eclipses and the pattern of field lines near the poles of a magnet that first raised suspicions that there were magnetic fields on the Sun [6]. The shapes of the polar rays and dipolar field lines have subsequently been compared many times and found to be in satisfactory agreement [7, 8].

Analysis of the geometry of the distribution of magnetic-field lines can give us some information about the field without the need to invoke other data. For example, a denser pattern of field lines implies an increased field strength. The shape of the envelope of the family of tangents to the field lines (penumbral filaments) on a circle concentric to a sunspot was used to find the relationship between the potential and vortical parts of the field, its moments [9]. As a rule, the tangents to polar rays (Fig. 1) intersect at a single point [10–18]. This property is shared by all axisymmetric spherical harmonics when their angular coordinates ϑ are small. In other words, this fact indicates that the global solar magnetic field is nearly potential. The position of the tangent intersec-
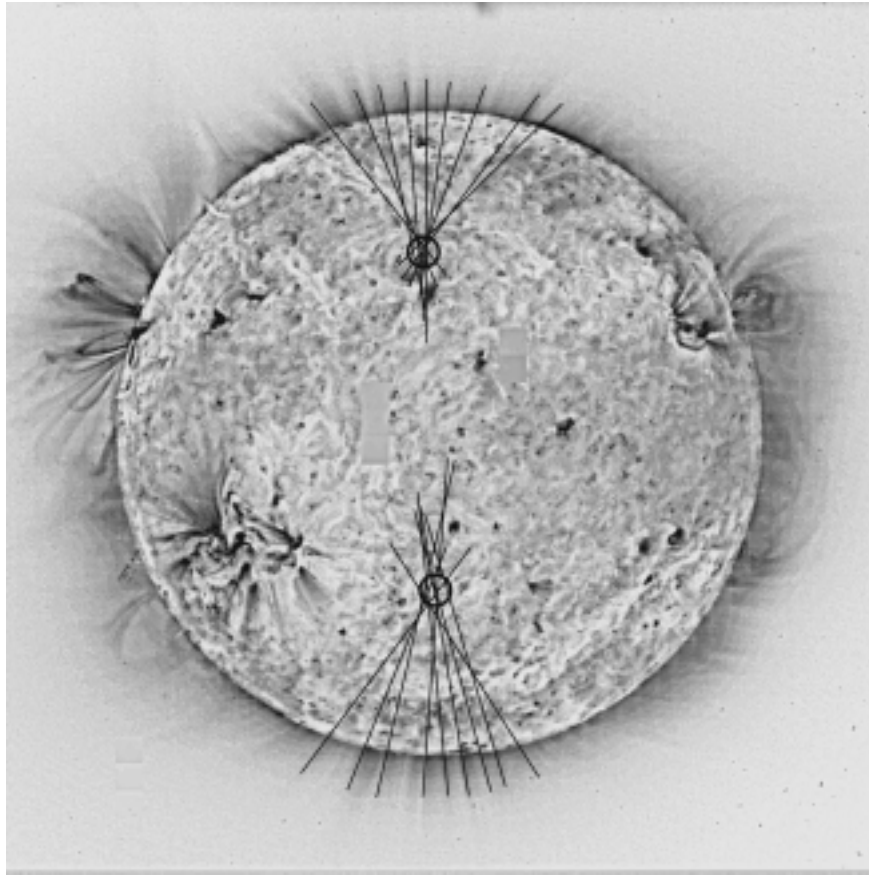
Fig. 1. Negative UV image of the Sun in 171-Å FeIX/X emission lines obtained on December 23, 1997 (13:54 UT), using the SOHO EIT, with tangents to the polar ray structures drawn. Regions of tangent intersections (magnetic foci) are marked with circles. Half the distance between the foci, q, is 0.52 $R_\odot$. (SOHO is a joint ESA−NASA project.)

tion point—magnetic focus—varies with the activity cycle [14, 18]. At activity minimum, the distance q between the center of the solar disk and the magnetic focus is maximum and is equal to approximately 0.65 $R_\odot$. At activity maximum, it reaches a minimum value of about $q = 0.45$ $R_\odot$. There has been no satisfactory explanation for this behavior of the magnetic foci. Moreover, according to our current understanding of the evolution of the global field, the variations in the positions of the foci should follow the opposite pattern. The relative weight of higher spherical harmonics increases during activity maximum, and the higher the harmonic number, the closer the magnetic focus to the surface. We will show here that the behavior of the magnetic foci reflects the complex evolution of the global magnetic field in the course of the cycle, with the polarity reversal occurring when the field structure is most complex. We will attempt to relate the motion of the foci to the dynamics of spherical harmonics known both from direct measurements and from analyses of synoptic maps. Allowing for interactions between the harmonics that can tentatively be associated with the preceding and the following

cycles make it possible to obtain a cycle-phase dependence for the depths of the foci that is close to the observed pattern.

2. MAGNETIC FOCI OF SPHERICAL HARMONICS

We can see in Fig. 2 that the distance q between the intersection point of the tangents to the surface magnetic field and the rotational axis of the Sun at the latitude $\varphi = 90° - \vartheta$, where ϑ is a coordinate in the regular spherical coordinate system, is determined by the inclination of the field line to the radius:

$$\tan \alpha = \frac{B_\vartheta}{B_r}. \tag{1}$$

According to the law of sines,

$$\frac{R_\odot}{\sin(\pi - \alpha - \vartheta)} = \frac{q}{\sin \alpha}, \tag{2}$$

so that

$$q = \frac{R_\odot}{\sin \vartheta \cot \alpha + \cos \vartheta}, \tag{3}$$

or, in view of (1),

$$q = \frac{R_\odot}{\dfrac{B_r}{B_\vartheta} \sin \vartheta + \cos \vartheta}. \qquad (4)$$

An axisymmetric potential field can be represented as a set of zonal spherical harmonics:

$$B_r^{(n)} = g_n(n+1) \left(\frac{R_\odot}{r} \right)^{n+2} P_n(\cos \vartheta), \qquad (5)$$

$$B_\vartheta^{(n)} = -g_n \left(\frac{R_\odot}{r} \right)^{n+2} \frac{dP_n(\cos \vartheta)}{d\vartheta} \qquad (6)$$

$$= g_n \left(\frac{R_\odot}{r} \right)^{n+2} P_n^1(\cos \vartheta),$$

where g_n is the weight of the corresponding harmonic and P_n^l are the associated Legendre polynomials [19, 20]. Substitution of these expressions into (4) yields

$$q^{(n)} = \frac{R_\odot}{\dfrac{(n+1)P_n}{P_n^1} \sin \vartheta + \cos \vartheta}. \qquad (7)$$

Using the behavior of the Legendre functions in the vicinity of the singular point $\cos \vartheta = 1$ [21],

$$P_n \approx 1, \qquad (8)$$

$$P_n^1 \approx \frac{1}{2} n(n+1) \sin \vartheta,$$

we find for the (northern) polar region, $\vartheta \sim 0$, that

$$q^{(n)} = R_\odot \frac{n}{n+2}. \qquad (9)$$

Formula (9) was apparently first derived in [22], where it was considered additional evidence for nondipolar structure of the global solar magnetic field at the photospheric level near both activity maximum and activity minimum and for the predominance of the higher harmonics of this global field. Indeed, for a dipole ($n = 1$), formula (9) yields $q^{(1)} = 0.33\, R_\odot$, which is lower than the values observed during any part of the cycle. A harmonic analysis of the photospheric field [23] demonstrates that the third, octupole harmonic dominates at activity minimum. The value $q^{(3)} = 0.6\, R_\odot$ is consistent with the observed value. However, as the activity grows, even higher harmonics begin to acquire larger weights ($n = 6$, $n = 9$); their $q^{(n)}$ values increase further, while the magnetic foci determined using the tangents method descend to lower depths. This contradiction has remained unaccounted for until the present.
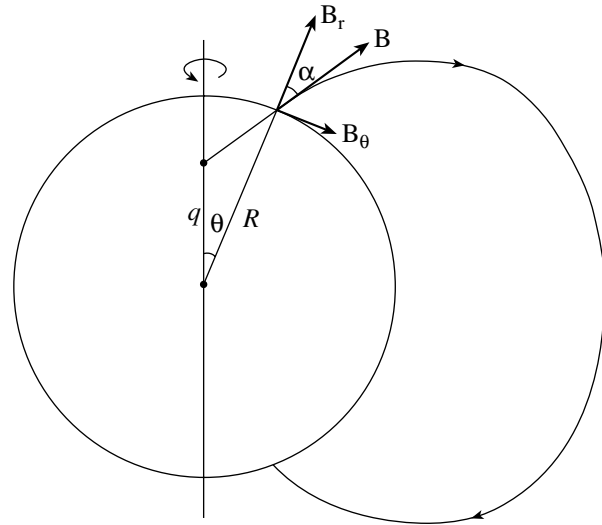
Fig. 2. Schematic showing the definition of q—the distance from a magnetic focus to the center of the solar disk.

3. MAGNETIC FOCI AND THE RELATIONSHIP BETWEEN HARMONICS DURING A POLARITY REVERSAL

The analogue to (9) for a field represented by only two harmonics is

$$q^{(n+m)} = R_\odot \frac{n}{n+2} \left(1 + \frac{\dfrac{m(n+2)}{n(m+2)} - 1}{\dfrac{(n+1)(n+2)}{C(m+1)(m+2)} + 1} \right), \qquad (10)$$

where C is the ratio of the harmonic coefficients g_m and g_n. When C is positive, $q^{(n+m)}$ varies from $q^{(n)}$ to $q^{(m)}$, depending on the relative weight of the corresponding harmonic. When C is negative, a C_2 singularity is present, where the denominator in (10) vanishes and

$$C_2 = -\frac{(n+1)(n+2)}{(m+1)(m+2)}; \qquad (11)$$

i.e., q approaches $-\infty$ on the right and $+\infty$ on the left of C_2. Let us note two other C values, viz., the C_1 value for which $q^{(n+m)} = 0$ and the C_3 value for which $B_r = 0$:

$$C_1 = -\frac{n(n+1)}{m(m+1)}, \qquad (12)$$

$$C_3 = -\frac{(n+1)}{(m+1)}. \qquad (13)$$

It is obvious that $C_1 > C_2 > C_3$ when $m > n$. It is important that, in the interval $C_1 < C < 0$, the values
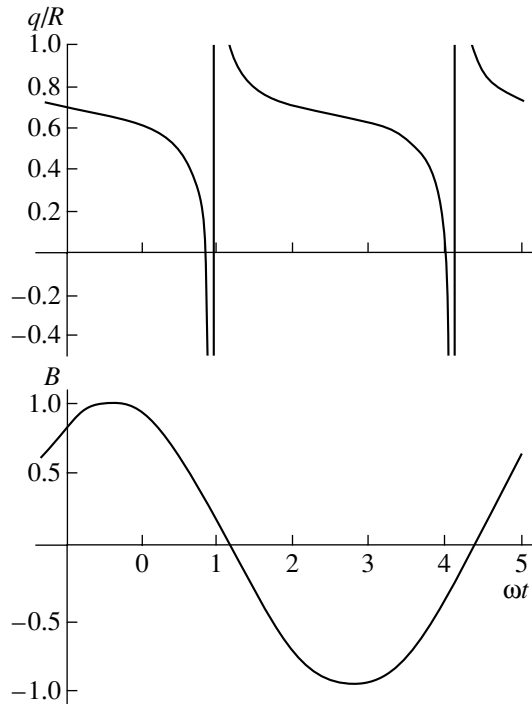
Fig. 3. Time dependence of q in a field composed upper panel of two zonal harmonics, $n = 3$ and $m = 7$, which vary according to cosine laws with a phase difference of $\varphi = \pi/2$ and an amplitude ratio of $A = 0.2$. An interval corresponding to a full magnetic cycle, or two 11-yr cycles, is shown. The variation of the radial magnetic field at the poles is shown at the bottom. The values of the field amplitudes are reached at activity minimum. The lower panel shows a plot of variations in the radial magnetic field at the pole. The amplitude field values are reached at minimum activity.

Fig. 4. Same as Fig. 3 but for the negative phase difference $\varphi = -\pi/2$.

of $q^{(n+m)}$ are smaller than q for each harmonic individually. Thus, as the relative weight of the harmonic m with the opposite polarity increases, the magnetic focus of the total field is lowered compared to the focus of any individual harmonic until a time slightly preceding the change of the field sign at the poles (the polarity reversal). This could be the origin of the observed lowering of the magnetic foci during activity maxima to levels below those corresponding to the harmonics that dominate in the global field at activity maximum.

The solar magnetic cycle is a complex process that is far from the simple oscillation of a dipolar field coaxial with the solar rotation. The field at low latitudes does not vanish, nor does it decrease during the polarity reversal, instead being enhanced and becoming more complex. The most general features of the magnetic cycle can be modeled to some extent by a dipole rotating in a meridional plane [24]. If we consider axisymmetric models, we conclude that the harmonics have the same period in time but differ in phase. This can clearly be seen, for example, in
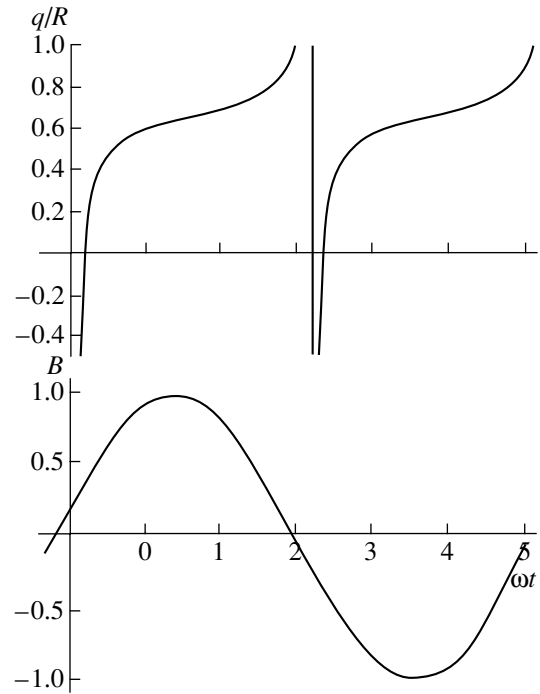
Fig. 1a in [23]: the first and third harmonics reach their maxima near the solar minimum, while the sixth and ninth harmonics reach their maxima near the solar maximum.

We will now try to construct a simple model for the dynamics of the global axisymmetric solar magnetic field that represents this field with only two odd zonal harmonics, $n = 3$ and $n = 7$, in order to describe the behavior of the parameter q in the cycle. We assume that

$$g_3(t) = \cos \omega t, \tag{14}$$
$$g_7(t) = A \cos(\omega t + \varphi),$$

where A is the ratio of the harmonic amplitudes, φ is the phase difference between the harmonics, and ω is the frequency of the full magnetic cycle, which corresponds to a period of 22 years. It follows from (11)–(13) that $C_1 = -0.21$, $C_2 = -0.28$, and $C_3 = -0.5$.

According to [23], the phase difference should be close to $\pi/2$ and, in general, either positive or negative. The modeled polarity-reversal pattern seems to agree better with the observed evolution of the global field if this phase difference is positive, i.e., the larger-scale part of the field leads: the region of old field shrinks and disappears at the pole rather than being pushed toward lower latitudes by the new field emerging at the pole. We chose the ratio of harmonic amplitudes A so that q values smaller than the q of any

harmonic were present during a substantial portion of the period.

The time dependences of $q^{(3+7)}$ for both positive and negative phase differences φ are shown in Figs. 3 and 4. In the lower part of each graph, the variation of the magnetic field at the pole is shown in relative units. The amplitude values of the field should be reached at the activity minimum, with the field value being zero at the maximum. The interval represented by the graphs corresponds to a full magnetic cycle, i.e., two 11-year sunspot cycles. The q values decrease and increase monotonically between polarity reversals for positive and negative phase differences, respectively. Near the reversal time, q approaches positive or negative infinity; we should bear in mind, however, that it is difficult to observe ray structures during such periods, if they exist at all. It is important that q changes in a jumplike fashion in our simple model, from small to large values if the phase difference is positive and from large to small values if it is negative. Thus, the values before and after the maximum should be very different. The available data [14] obtained from observations of total solar eclipses do not unambiguously demonstrate whether this effect is present. The information is most fragmentary near maxima, either due to the difficulties noted above or simply by coincidence.

Studies of the ray-structure geometry at the polar caps can now be done using solar images obtained daily in UV lines with the SOHO EIT. We plan on carrying out such work; the need for further improvements in the model will be evaluated after obtaining a more detailed picture of the q variations in the course of the cycle.

4. CONCLUSIONS

The position of the intersection point of the tangents to ray structures in the polar regions of the Sun characterizes certain general properties of the global magnetic field. For zonal spherical harmonics, the tangents to field lines on the spherical surface near a pole converge to a single point. Thus, the existence of such points indicates that the polar magnetic field is nearly potential. The distance q from the solar disk to this point (magnetic focus) varies during the solar-activity cycle; the largest q values are reached at activity minimum, and the smallest, at activity maximum. The larger the harmonic number, the larger the corresponding q; thus, the behavior of the magnetic foci is in contradiction with the dominance of lower harmonics at activity minimum and higher harmonics at activity maximum. We should bear in mind, however, that the solar magnetic cycle is not a simple synchronous variation of various spatial frequencies. The fact that the polar field vanishes and changes sign at the epoch of maximum demonstrates the existence of complex variations in the relative weights of the harmonics. For a field composed of harmonics that have opposite signs at the pole, q can be smaller than the q values for each harmonic taken alone. Thus, although the relative weight of higher harmonics grows near the maximum, the decrease in q at this time could be related to the emergence of harmonics with an oppositely directed polar field, which can tentatively be attributed to the field of the new cycle.

By way of illustration, we have described a simple model for the global field consisting of the third and seventh harmonics. Their amplitude ratio and phase difference were chosen so that the model reflects the observed dynamics of the magnetic foci of the polar field. Further development of the model requires, above all, more detailed knowledge of the behavior of q during the cycle. It is now possible to derive this information, since continuous observations of the solar corona with space-based telescopes have already been carried out over an entire cycle.

ACKNOWLEDGMENTS

REFERENCES

1. P. S. McIntosh, Rev. Geophys. Space Phys. **10**, 837 (1972).
2. V. I. Makarov and K. R. Sivaraman, Solar Phys. **85**, 227 (1983).
3. P. S. McIntosh, *The Solar Cycle*, Ed. by K. L. Harvey (Astron. Soc. Pac., San Francisco, 1992), Astron. Soc. Pac. Conf. Ser. **27**, 14 (1992).
4. V. I. Makarov, A. G. Tlatov, and K. R. Sivaraman, Solar Phys. **202**, 11 (2001).
5. B. D. Shel'ting and V. N. Obridko, *The Sun at Epochs of Magnetic-Field Sign Changes,* Ed. by V. I. Makarov and V. N. Obridko (2001) [in Russian], p. 391.
6. F. H. Bigelow, *The Solar Corona* (Washington Smithsonian Inst., 1889).
7. S. K. Vsekhsvyatskii and G. M. Nikol'skii, Astron. Zh. **32**, 354 (1955).
8. G. M. Nikol'skii, in *The Full Solar Eclipses of Feb. 25, 1952 and June 30, 1954* (Akad. Nauk SSSR, Moscow, 1959) [in Russian], p. 115.
9. M. M. Molodenskiĭ, L. I. Starkova, and B. P. Filippov, Astron. Zh. **68**, 612 (1991) [Sov. Astron. **35**, 300 (1991)].
10. H. C. van de Hulst, Bull. Astron. Inst. Netherl. **11** (410), 150 (1950).

11. E. Ya. Boguslavskaya, in *The Full Solar Eclipses of Feb. 25, 1952 and June 30, 1954* (Akad. Nauk SSSR, Moscow, 1959) [in Russian], p. 100.

12. S. K. Vsekhsvyatskii and G. M. Nikol'skii, Astron. Zh. **33**, 87 (1956).

13. K. Saito, Publ. Astron. Soc. Jpn. **10**, 49 (1958).

14. A. T. Nesmyanovich, Astron. Zh. **39**, 996 (1962) [Sov. Astron. **6**, 774 (1962)].

15. S. Koutchmy and K. Bocchialini, *Solar Jets and Coronal Plumes*, ESA SP-421 (1998), p. 51.

16. V. I. Makarov, *Proceedings of the 18th NSO Workshop*, ASPCS **140**, 83 (1998).

17. R. A. Gulyaev, *Solar Jets and Coronal Plumes,* ESA SP-421, 277 (1998).

18. V. I. Makarov, K. S. Tavastsherna, and D. K. Callebaut, *The New Solar Activity Cycle: Observational and Theoretical Aspects* (GAO, St. Petersburg, 1998) [in Russian], p. 111.

19. S. Chapman and J. Bartels, *Geomagnetism* (Oxford Univ. Press, London, 1940).

20. M. Altschuler and G. Newkirk, Jr., Solar Phys. **9**, 131 (1969).

21. H. Bateman, *Higher Transcendental Functions,* Ed. by A. Erdélyi (McGraw-Hill, New York, 1953–1955; Nauka, Moscow, 1965), Vols. 1–3.

22. F. A. Ermakov, V. N. Obridko, and B. D. Shel'ting, Astron. Zh. **72**, 753 (1995) [Astron. Rep. **39**, 672 (1995)].

23. S. Bravo, G. A. Stewart, and X. Blanco-Cano, Solar Phys. **179**, 223 (1998).

24. L. I. Starkova and L. S. Solov'ev, Pis'ma Astron. Zh. **23**, 236 (1997) [Astron. Lett. **23**, 207 (1997)].

Translated by A. Getling